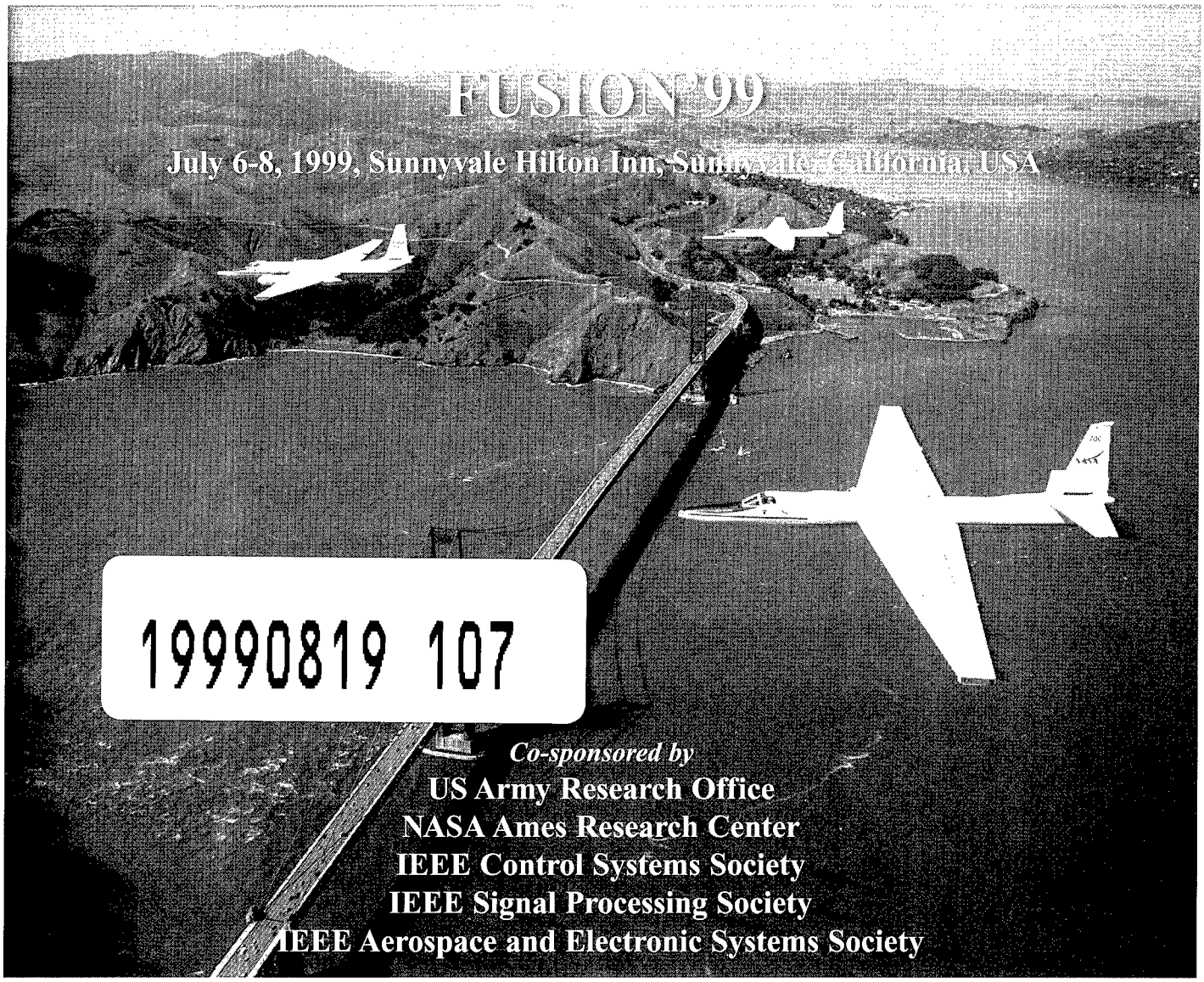
International Society of Information Fusion

ARO 40107.1-EL-CF

PROCEEDINGS OF THE SECOND INTERNATIONAL CONFERENCE ON INFORMATION FUSION

Volume II



REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		Final Report

4. TITLE AND SUBTITLE

FUSION'99 -- The Second International Conference on Information Fusion, Volumes I and II

5. FUNDING NUMBERS

DAAD19-99-1-0271

6. AUTHOR(S)

Dr. James Llinas, Principal Investigator

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)

International Society of Information Fusion
Mountain Vieew, CA 94043

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

ARO 40107.1-EL-CF

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

NO ABSTRACT PROVIDED

14. SUBJECT TERMS

15. NUMBER IF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

International Society of Information Fusion

PROCEEDINGS OF THE SECOND INTERNATIONAL CONFERENCE ON INFORMATION FUSION

Volume II

FUSION'99

July 6-8, 1999, Sunnyvale Hilton Inn, Sunnyvale, California, USA

Co-sponsored by

US Army Research Office
NASA Ames Research Center
IEEE Control Systems Society
IEEE Signal Processing Society
IEEE Aerospace and Electronic Systems Society

This set of volumes contains technical papers presented at the Second International Conference on Information Fusion (FUSION'99). Their inclusion in this publication does not necessarily constitute endorsements by the International Society of Information Fusion, the sponsors, or by the publisher.

Distributed by

International Society of Information Fusion
415 Clyde Avenue, Suite 108
Mountain View, CA 94043 USA
Tel: 650-966-8700, Fax: 650-966-8700
Email: info@inforfusion.org, URL: www.inforfusion.org

Volume I, ISBN 0-9671429-C-3-1-1
Volume II, ISBN 0-9671429-C-3-2-X
Set, ISBN 0-9671429-C-3-3-8

Photo on the cover: The Golden Gate Bridge, San Francisco, California, viewed from the U-2 airplane. Courtesy of NASA Ames Research Center, photographer Dominic Hart.

Printed by Omnipress in the U.S.A.

Preface

Dear Members of the Information Fusion Community:

It is a pleasure to report to you that the Information Fusion community continues to mature and grow, a positive reflection on all members and especially on that subgroup of the community that persists in supporting its maturation process. Thanks are due to Dongping Daniel Zhu and X. Rong Li, Belur Dasarathy, and the members of the Transitional Board of the International Society of Information Fusion (ISIF), for the attention paid to and energy expended on the wide variety of tasks and issues involved with trying to get the ISIF established. Tasks of this sort are 'yet another thing to do' for those involved but these noble, collective efforts and their results and consequences are what give identity and substance to a community. Slowly but persistently, this community is filling in the "Infrastructure gaps" it has suffered from for some time--we hope soon to have a Society, an International Journal, and an Information Analysis Center; we already have one University Research Center, which could be expanded to a Consortium framework.

The ISIF is a particularly welcome and needed infrastructure initiative in our community, but it will only be as good as the collective efforts of its membership. Being a member of any Society results in both an opportunity and an obligation; opportunity for collegiality in its fullest sense, and obligation to contribute in its fullest sense. Being among the oldest in this community, I can tell you that I have always been proud to label myself as a member of the "fusion" community since it is a distinctive, extraordinarily interesting field of specialization, and one with great promise. We welcome and encourage you to become "official" members via the ISIF, about which we will all have considerable discussion at FUSION'99 - give us your thoughts about what ISIF should be, and give us your membership; see http://www.inforfusion.org for more information.

In recent visits I have had the opportunity to interact with and learn from Information Fusion researchers in Australia, in Spain, and in Norway, and last year I was involved in a technology planning task in Sweden. In all cases I was impressed with both the nature of the work and the talented people involved in it. I think I can say without reservation that all of the people involved in these IF efforts, as well as the cognizant organizational leaders and managers are anxious for interaction, and technology and knowledge-sharing, and for a forum to periodically share ideas. Inspired by this, I have motivated a session on "International Collaboration in IF" for this year's conference which I hope will be a standing session for future conferences, and which I hope will be one focused forum in which people can both understand what options for collaboration may exist and also to act on them. Of course the "FUSION'XX" conferences serve this purpose in the large, but offering some details on the underlying mechanics regarding programs and activities specifically tailored to international collaboration won't hurt.

Welcome to FUSION'99

Jim Llinas
President, International Society of Information Fusion

International Society of Information Fusion

Initial Board of Directors

Yaakov Bar-Shalom, University of Connecticut, CT, USA
Mark Bedworth, Defence Evaluation & Research Agency (DERA), Malvern, UK
Chee-Yee Chong, Booz-Allen & Hamilton, San Francisco, CA, USA
Belur V. Dasarathy, Dynetics, Inc., Huntsville, Alabama, USA
Alfonso Farina, Alenia Defence Systems, Italy
X. Rong Li, University of New Orleans, LA, USA
Jim Llinas, State University of New York, Buffalo, NY, USA
Daniel McMichael, CSSIP, Australia
Jane O'Brien, Defence Evaluation & Research Agency (DERA), Malvern, UK
Pramod K. Varshney, Syracuse University, NY, USA
Dongping Daniel Zhu, Zaptron Systems, Inc., Mountain View, CA USA

Initial Society Officers

President: Jim Llinas, State University of New York, Buffalo, NY, USA
Secretary: Dongping Daniel Zhu, Zaptron Systems, Inc. Mountain View, CA, USA
Treasure: Chee-Yee Chong, Booz-Allen & Hamilton, San Francisco, CA, USA

Society Address

415 Clyde Avenue, Suite 108
Mountain View, CA 94043, USA
Tel: 650-966-8700, Fax: 650-966-8780
Email: info@inforfusion.org URL: http://www.inforfusion.org

Foreword

Across Las Vegas desert land, Heat waves shimmering from the sand.
A fusion caravan comes into view, Destination – Timbuktu.

If Shakespeare is correct that "What's past is prologue," then FUSION'98 should be an good introduction that brings us together again at FUSION'99 in the Silicon Valley, exactly one year later. Clearly, data fusion follows from idea fusion and people fusion

It gives us great pleasure to introduce this collection of papers presented at the Second International Conference on Information Fusion (FUSION'99), organized by the International Society of Information Fusion (http://www.inforfusion.org) on July 6 through July 8, 1999, at Sunnyvale Hilton Inn, California, USA. These papers reflect the state-of-the-art of sensor, data and information fusion, and cover architecture, algorithms and applications in many fields, ranging from target tracking and recognition to diagnostic information fusion and image fusion to biomedical and management information fusion.

Many factors have contributed to FUSION'99. First of all, we'd like to thank the conference sponsors, without their support this conference would not have been possible. These sponsors are NASA Ames Research Center*, US Army Research Office*, IEEE Signal Processing Society, IEEE Control Systems Society, and IEEE Aerospace and Electronic Systems Society.

We are fortunate to have many renowned people to provide vision and leadership to the conference. We are especially grateful to Dr. Yaakov Bar-Shalom of University of Connecticut who serves as Honorary Chairman, Franklin White of Navy SPAWAR as Steering Committee Chairman, Dr. Kenneth Ford of NASA as Advisory Committee Chairman, Mark Bedworth of DERA, UK and Dr. X. Rong Li of University of New Orleans as General Vice Chairmen, and Dr. Pramod Varshney of Syracuse University as Technical Program Chairman. We gratefully acknowledge Dr. Bill Sanders of Army Research Office for his continued inspiration and support.

We are very grateful to the many colleagues who are experts in the field and have greatly helped organize the conference. In particular, the General Chairman would like to thank all members on the Technical Program Committee, led by Dr. Pramod Varshney and Dr. Peter Willett, for their efforts in assembling a collection of quality papers, and Dr. Robert Levinson for his tireless effort in printing and publishing the Proceedings. We like to acknowledge other Executive Committee members: Dr. Chee-yee Chong for managing logistics and finance, Captain Erick Blasch for leading a successful sponsors program, Dr. Belur Dasarathy for publicizing the conference to a wide audience, and Dr. Fa-long Luo for local arrangements. Last but not the least, Society board directors and liaisons, session chairs, authors, and many others have offered valuable assistance. They all helped make the conference a success.

We also like to thank the following persons: Deborah Jean Gamble-Ly of Creation, Janny Wu, and Mike Lee of ComStar for administrative assistance, Maylene Duenas and her staff at NASA for technical support, Bob Hamm of OmniPress for publication, and the staff at Zaptron Systems for web site support.

With the success of FUSION'99, we can expect even greater successes at FUSION'2000 in the new millennium. In the words of Sir Winston Churchill: *"This is not the end, it is not even the beginning of the end, but it is perhaps the end of the beginning."*

Dongping Daniel Zhu, General Chairman
Zaptron Systems, Inc.
Robert Levinson, Publication Chair
University of California-Santa Cruz

* The views, opinions, and/or findings contained in this proceedings are those of the authors and should not be construed as an official US government or its agency's position, policy, or decision, unless so designated by other documentation.

Technical Program Chair's Message

I am delighted to welcome you to FUSION'99. We have assembled an excellent technical program consisting of 29 contributed and invited sessions. The conference attracted about 210 submissions from 22 countries. Each submission was reviewed by the technical program committee and only worthy papers were included in the final program. I was extremely pleased with the large number of submissions and their high quality. In addition to the technical sessions, we feature three plenary talks and a luncheon talk by R. Luo (Taiwan), K. Ford (USA), G. Shaw(USA) and F. White(USA). All of these speakers are widely known and have significant experience in their areas of expertise.

It is a pleasure to acknowledge the tireless effort of Peter Willett, the Technical Program Vice Chair. He reviewed each and every submission and was instrumental in putting the sessions together. I would like to thank the members of the Technical Program Committee for their assistance with reviewing: M. Alford (USA), B. Dasarathy (USA), D. McMichael (Australia), J. O'Brien (UK), E. Shahbazian (Canada), and P. Svensson (Sweden).

The efforts of the following persons in organizing invited sessions are greatly appreciated: C. Anken, E. Blasch, R. Blum, O. Drummond, K. Goebel, M. Kokar, M. Larkin, R. Liuzzi, J. Llinas, G. Rogova, S. Shah, A. Stoica, and D. Zhu.

This is the second year for this conference and we have made great strides in this short period. I am confident that the conference will continue to grow both in terms of size and quality. Thank you all for making this conference a success.

Pramod E. Varshney
Technical Program Chair
Professor
Syracuse University
NY, USA

Fusion'99 Executive Committee

Honorary Chairman:
Yaakov Bar-Shalom, University of Connecticut, CT, USA

Steering Committee Chairman:
Franklin E. White, SPAWAR, San Diego, CA, USA

Advisory Committee Chairman:
Kenneth M. Ford, NASA Ames Research Center, Moffet Field, CA, USA

General Chairman:
Dongping Daniel Zhu, Zaptron Systems, Inc., Mountain View, CA, USA,

General Vice Chairmen:
Mark Bedworth, Defense Evaluation & Research Agency, UK
X. Rong Li, University of New Orleans, LA, USA

Technical Program Committee Chair:
Pramod K Varshney, Syracuse University, NY, USA

Technical Program Committee Vice Chair:
Peter Willett, University of Connecticut, CT, USA

Publicity Committee Chair:
Belur V. Dasarathy, Dynetics, Inc. Huntsville, AL, USA

Publication Committee Chair:
Robert Levinson, University of California at Santa Cruz, CA, USA

Sponsors Program Chair:
Erick Blasch, Air Force Research Lab/SNAT, OH, USA

Financial Committee Chair:
Chee-Yee Chong, Booz-Allen & Hamilton, San Francisco, CA, USA

Local Arrangement Committee Chair:
Fa-long Luo, Resound Corporation, Redwood City, CA, USA

Technical Program Committee

Pramod Varshney (Chair), Syracuse University, USA
Peter Willett (Vice Chair), University of Connecticut, USA
Mark Alford, Air Force Research Laboratory at Rome, USA
Belur Dasarathy, Dynetics Corporation, USA
Daniel McMichael, University of South Australia, Australia
Jane O'Brien, Defense Evaluation Research Agency, U.K.
Elisa Shahbazian, Lockheed-Martin, Canada
Per Svensson, FOA, Sweden

Fusion'99 International Committee

J. K. Aggarwal, University of Texas-Austin, USA
C. Anken, AFRL, USA
Alain Appriou, French National Establishment for Aerospace Research, Chatillon, France
Steve Bishop, Army Night Vision Lab, Virginia, USA
Rick S. Blum, Electrical Engineering and Computer Science Dept., Lehigh University, USA
George Chapline, Lawrence Livermore Labs, CA, USA
Nianyi Chen, Chinese Academy of Sciences, Shanghai, P. R. China
Oliver Drummond, Consulting Engineer, Culver City, CA, USA
Alfonso Farina, ALENIA, Italy
Mohamad Farooq, Royal Military College of Canada, Ontario, Canada
A. Fatholahzadeh, SUPELEC: Ecole Superieure d'Electricite, France
Kai Goebel, Information Technology Laboratory, General Electric, NY, USA
Chris Harris, University of Southampton, UK
M. Hinman, AFRL, USA
R.A. Hogendoorn, National Aerospace Laboratory, The Netherlands
Masatoshi Ishikawa, University of Tokyo, Japan
Vishrut Jain, National University of Singapore, Singapore
Ivan Kadar, Northrop Grumman Corp, Advanced Systems and Technology, NY, USA
T. Kirubarajan, University of Connecticut, USA
M. Kokar, Northeastern University, MA, USA
P. Korpisaari, Tampere University of Technology, Finland
Rudolf Kruse, Universitaet Magdeburg, Germany
Michael J. Larkin, Naval Undersea Warfare Center-Newport Division, Newport, RI, USA
James Llinas, State University of NY at Buffalo, Buffalo, NY, USA
R. Lynch, Naval Undersea Warfare Center, RI, USA
Ren C. Luo, National Chung Cheng Univesity, Taiwan
Rabinder N. Madan, Office of Naval Research, Arlington, VA, USA
Peter Wide, Department of Technology and Science, Oerebro University, Oerebro, Sweden
R. Liuzzi, AFRL, USA
Shozo Mori, Raytheon TI Systems, San Jose CA, USA
S. Musick, AFRL, USA
Akira Namatame, Dept.of Computer Science, National Defense Academy, Japan
Theo van Niekerk, Port Elizabeth Technikon, South Africa
Nageswara S. V. Rao, Oak Ridge National Laboratory, TN, USA
G. Rogova, Calspan/CUBRC, NY, USA
Peter Wide, Orebro University, Sweden
Bill A. Sander, Army Research Office, NC, USA
S. Shah, Wayne State University, MI, USA
E. Shahbazian, Lockheed-Martin, Canada
Alan Steinberg, ERIM, MI, USA
Adrian Stoica, NASA Jet Propulsion Laboratory, CA, USA
Kamyshnikov Vladimir, Department of Economy, Tomsk State Architectural University, Russia
Zhongtuo Wang, Business School, Dalian University of Technology, P. R. China
Vitaly Yaschenko, Institute of Mathematical Machines and Systems, Ukraine

1 Plenary Speech I: "Multisensor Fusion and Integration Issues, Approaches and Opportunities"

Dr. Ren C. Luo, Professor and Dean College of Engineering National Chung Cheng University, Taiwan and General Chair of MFI'99 - IEEE International Conference on Multisenor Fusion and Integration for Intelligent Systems

1.1 ABSTRACT

Interest has been growing in the use of multiple sensors to increase the capability of intelligent systems. In this presentation, the issues, approaches in dealing with multisensor fusion and integration (MFI) will be discussed. The applications and potential opportunities for the implementation of MFI will also be included. The issues involved in integrating multiple sensors into the operation of a system are presented in the context of the type of information these sensors can uniquely provide. The advantages gained through the synergistic use of multisensory information can be decomposed into a combination of four fundamental aspects: the redundancy, complementarily, timeliness, and cost of the information can then defined as the degree to which each of these four aspects is present in the information provided by the sensors.

In general, sensory fusion can be accomplished at different levels: data fusion, feature fusion and decision fusion. More commonly known is data fusion level, Example of this type of fusion are fusion of multiple ultrasonic data, and fusion of images from different imaging sensors. In feature fusion level, features are extracted from the raw measurements that are then combined in a quantitative or qualitative manner. For example, feature fusion can be used to fuse information from imaging and a non-imaging sensor. Decision fusion level can be employed when the sensors available are not compatible or be applicable to many pattern recognition problems.

Typical of the applications that can benefit from the use of multiple sensors are industrial tasks like assembly, military command and control for battlefield management, mobile robot navigation, multitarget tracking, and aircraft navigation. Common among all of these applications is the requirement that the systems intelligently interact with and operate in an unstructured environment without the complete control of a human operator. Advances in hardware, software and algorithm have made it possible to employ multiple data sources for information gathering and to develop more complex multisensor fusion and integration system. An example of applying MFI system in an automations mobile robot/intelligent wheelchair system with video demonstration will also be presented.

1.2 Short Biographical Sketch

Ren C. Luo (IEEE M'82 - SM'87 - F'92), is currently a Professor and Dean of College of Engineering at National Chung Cheng University, he also served as Director of Automation Technologies Program at National Science Council and Advisor of Ministry of Economics Affairs in Taiwan, R.O.C. He was a Professor in the Department of Electrical and Computer Engineering and the Director of the Center for Robotics and Intelligent Machines at North Carolina State University in Raleigh, North Carolina, USA. He received his Ph.D degrees from Technische Universitaet Berlin, Berlin, Germany in 1982.

¿From 1983 to 1984, he was an Assistant Professor in the Department of Electrical Engineering and Computer Science at the University of Illinois at Chicago. From 1984 to 1990, he was Assistant, Associate Professor and became Professor since 1991 in the Department of Electrical Computer Engineering at North Carolina State University, Raleigh, NC. From 1992 to 1993, he was Toshiba Chair Professor at University of Tokyo, Japan.

Dr. Luo's research interests include: sensor-based intelligent robotics systems, multisensor fusion and integration, computer vision, rapid prototyping and advanced manufacturing systems. Dr. Luo has published over 170 technical journals, proceedings, and patents in the above-mentioned areas. He authored a book, Multisensor Fusion and Integration (Ablex, 1995); and was editor of the book, Robotics and Vision (IEEE, 1988). Dr. Luo was also guest editors for the Journal of Robotics Systems (John Wiley and Sons. Vol. 7, 3, 1990), IEEE Transactions on Industrial Electronics in special issues on the topics of multisensor fusion and integration for intelligent machines, and editor of IEEE/ASME Transactions on Mechatronics.

2 Plenary Speech II: "AI and Space Exploration"

Dr. Kenneth M. Ford Associate Center Director for Information Technology and Director of NASA's Center of Excellence in Information Technology, NASA Ames Research Center, Moffet Field, CA, USA

2.1 ABSTRACT

Humans are quintessentially explorers and makers of things. These traits, which identify us as a species and account for our survival, are reflected with particular clarity in the mission and methods of space exploration. The romance associated with the Apollo project is being replaced with a different vision, one where we make tools to do our exploring for us. We are building computational machines that will carry our curiosity and intelligence with them as they extend the human exploration of the universe.

In order to succeed in places where humans could not possibly survive, these "remote agents" must take something of us with them. They must be self-reliant, smart, adaptable and curious. Our mechanical explorers cannot be merely passive observers or puppets dancing on tenuous radio tethers from earth. They simply will not have time to ask us what to do: the twin constraints of distance and light-speed would render them helpless while waiting for our instructions, even if we knew what to tell them. AI plays a central role in space exploration because there is, literally, no other way to make it work. Our bodies cannot fly in the tenuous Martian atmosphere, endure Jupiter's gravity or the electromagnetic turbulence of Saturn's rings; but our machines can, and we will send them there. Once at distant worlds, however, they must deal with the details themselves. The only thing we can do is to make them smart enough to cope with the tactics of survival.

How clever will these agents of human exploration need to be? Certainly, cleverer then we can currently make them. It will not be enough to be situated and autonomous: they will need to be intelligent and inquisitive and thoughtful and quick. NASA is committed to integrating intelligent systems into the very center of our long-range strategy to explore the universe.

In this talk, I will describe the current and future research directions of NASA's expanding information technology effort with a particular emphasis on intelligent systems.

2.2 Short Biographical Sketch

Kenneth M. Ford is the Associate Center Director for Information Technology at NASA Ames Research Center and Director of NASA's Center of Excellence for Information Technology. In these roles, Dr. Ford has had the honor and responsibility of helping shape NASA's IT research effort (about 200M dollars effort at Ames, but much larger Agency wide). The Ames Research Center has about 5,000 employees, of which about a third work in IT and 700 have Ph.D degrees.

Additionally, Dr. Ford is the Director and Founder of the Institute for the Interdisciplinary Study of Human and Machine Cognition (IHMC) at the University of West Florida - a multidisciplinary research unit of the State University System. Since its founding in 1990, IHMC has rapidly grown into a well-respected research institute investigating a broad range of topics related to understanding cognition in both humans and machines with a particular emphasis on building cognitive prostheses to leverage and amplify human intellectual capacities. While at the University of West Florida Professor Ford received national and local recognition for teaching excellence and in 1997 he was awarded the University's highest research distinction, the Research and Creative Activities Award. Dr. Ford has been on a leave absence from the University to NASA for the last two years.

Dr. Ford entered computer science and artificial intelligence through the back door of philosophy. After studying epistemology as an undergraduate, he joined the Navy and wound up fixing computers among other things. When his Navy stint ended, he earned his doctoral degree in computer science from Tulane University in 1988. His research interests, among others, include: artificial intelligence, knowledge-based performance support systems, computer-mediated learning, and internet-based applications. Dr. Ford is the author of well over 100 scientific papers and the author/editor of five books.

Dr. Ford is the Editor-in-Chief of AAAI/MIT Press, Executive Editor

of the International Journal of Expert Systems, Associate Editor of the Journal of Experimental and Theoretical Artificial Intelligence, and is a Behavioral and Brain Sciences (BBS) Associate.

3 Plenary Speech III: "Music Enhances Learning: Keeping Mozart in Mind"

Dr. Gordon Shaw Professor Emeritus, Elementary Particle Theory Theoretical Neurobiology Department of Physics and Center for the Neurobiology of Learning and Memory University of California - Irvine CA, USA

3.1 ABSTRACT

Theoretical studies [Leng and Shaw, 1991], the "Mozart effect," based on the trion model [Shaw et al., 1985] predicted that music would enhance spatial-temporal reasoning (the ability to mentally image and transform patterns in space and time). Recent supporting experiments involving the Mozart Sonata for Two Pianos in D Major-K.448 are: behavioral studies showed that listening to it enhanced spatial-temporal reasoning in humans [Rauscher et al., 1993, 1995; Johnson et al., 1998] and in rats [Rauscher et al., 1998]; EEG studies [Sarnthein et al., 1997] showed that listening to it results in increased coherence lasting several minutes; exposure to it reduced pathological activity in comatose epileptic patients [Hughes et al., 1998]. MRI studies [Muftuler et al., 1999] showing excitation of cortex relevant to spatial-temporal reasoning. Studies relevant to education are: We [Rauscher et al., 1997] showed that preschool children who were given 6 months of piano keyboard training improved dramatically on spatial-temporal reasoning. Second grade children (in the inner-city 95 St. School in Los Angeles) given 4 months of piano keyboard training as well as training on Peterson's math video software scored striking higher [Graziano et al., 1999] on proportional math and fractions. Support for the trion model from cortical data [Bodner et al., 1997] show families of firing patterns related by symmetries. Implications for education, basic neuroscience, clinical medicine, and technology are discussed.

3.2 Short Biographical Sketch

Professor Shaw earned his B.S. from Case Institute of Technology in 1954 and his Ph.D. in Theoretical Physics from Cornell University in 1959. He had post-doctoral positions at Indiana University and the University of California, San Diego, and a teaching position at Stanford University before joining the new UCI campus in 1965. In addition to his research in elementary particle theory, he started working on brain theory in 1974. He is a member of the UCI Center for the Neurobiology of Learning and Memory.

4. Plenary Speech IV: "International Fusion: Changes and Approaches"

Franklin E. White, Jr.
Director, Program Development, Navy SPAWAR Systems Center,
Code D101, San Diego, CA, USA, Email: whitefe@spawar.navy.mil

4.1 ABSTRACT

In the current Information age, the potential for overwhelming availability of data largely without meaning has become a reality. Everywhere individuals and organizations are drowning in data and information and starved for knowledge and understanding. This is a problem that has become apparent worldwide in developed and developing countries. One of the keys to addressing this is data and information fusion. Fusion has long been the domain of a relatively small number of practioners in a largely classified endeavors within nations. This speech will address the changes in this world view that are coming about and discuss the burgeoning exchange of information about fusion on an increasingly global basis. It will also suggest some discipline and approaches essential to making fusion tools useful, and discuss some of the needed mechanisms and pitfalls as an international community comes together.

4.2 Short Biographical Sketch

Franklin E. White Jr. has spent 30 years with Navy as an officer and scientist. He has focused on integration and fusion efforts, has worked with Navy's Command, Control and Intelligence systems and is Chairman of the Joint Directors of Laboratories, Data Fusion Group. Mr. White has long term experience with Top Level architectures, serving on the team that developed the Copernicus Architecture and spending two years on detail to the Intelligence Community Management Staff (CMS) where he chaired the working group that developed the INTELINK information sharing concept. He has long been a supporter of international cooperation serving for 2 years at RAF Brawdy Wales, UK and temporarily at many European sites and is active in many international programs. He has spoken at international CIS symposia and AFCEA meetings. He is a long time member of AFCEA , SASA, The Naval Institute and Naval Intelligence Professionals and is currently the Director of Program Development at SPAWAR Systems Center San Diego.

Tuesday July 6

Session TA1 Naval Applications
Chair: Michael Larkin, Naval Undersea Warfare Center

Session TA2 Medical Applications
Chair: Robert Levinson, University of California at Santa Cruz

Session TA3 Decentralized Detection Systems
Chair: Jane O'Brien, Defense Evaluation & Research Agency, U.K.

Session TA4 Formal Methods for Information Fusion
Chair: Mitch Kokar, Northeastern University

Session TB1 Radar and Communication Applications of Fusion
Chair: Rick Blum, Lehigh University

Session TB2 Image Fusion I
Chair: Elisa Shahbazian, Lockheed-Martin, Canada

Session TB3 Fusion for Target Tracking I
Chair: Mohamad Farooq, Royal Military College of Canada

Session TB4 Classification I
Chair: Nageswara S. V. Rao, Oak Ridge National Laboratory

Session TC1 International Collaboration in Fusion Research and Development
Chair: James Llinas, State University of New York at Buffalo

Session TC2 Diagnostic Information Fusion
Chair: Kai Goebel, Information Technology Laboratory, General Electric

Session TC3 Fusion of Fuzzy Information
Chair: Daniel McMichael, University of South Australia

Wednesday July 7

Session WA1 Image Fusion II
Chair: P. Svensson, Defence Research Establishment, Sweden

Session WA2 Fusion Architecture & Management I
Chair: Alan Steinberg, Environment Research Institute of Michigan

Session WA3 Fusion for Target Tracking II
Chair: S. Musick, US Air Force Research Laboratories

Session WA4 Biological and Linguistic Models for Fusion
Chair: George Chapline, Lawrence Livermore National Laboratory

Session WB1 Knowledge-Based Techniques for Information Fusion and Discovery
Chair: Ray Liuzzi and Craig Anken, US Air Force Research Laboratories

Session WB2 Hardware for Information Fusion
Chair: Adrian Stoica, NASA Jet Propulsion Laboratory

Session WB3 Multisensor-Multisource Fusion for Object Tracking and Recognition
Chair: Shishir Shah, Wayne State University

Session WB4 Fusion for Target Tracking III
Chairs: X. Rong Li, University of New Orleans and T. Kirubarajan, University of Connecticut

Thursday July 8

Session RA1 Image Fusion III
Chair: Robert S. Lynch, Naval Undersea Warfare Center, RI, USA

Session RA2 Fusion Architecture and Management II
Chair: Ivan Kadar, Consultant, Interlink and Northrup Grumman Corp., USA

Session RA3 Information Fusion for Decision Support
Chair: Galina Rogova, Calspan/CUBRC, NY, USA

Session RA4 Fusion for Fault Detection and Diagnosis
Chair: Mark Alford, Air Force Research Laboratory-Rome, NY, USA

Session RB1 Management and Business Information Fusion
Chair: Nianyi Chen, Chinese Academy of Sciences, Shanghai, P. R. China

Session RB2 Multisensor Target Tracking and Recognition of Small-to-Medium Sized Targets
Chair: Oliver Drummond, Consulting Engineer, Culver City, CA, USA

Session RB3 Emerging Applications I
Chair: Xue-Gong Zhang, Tsinghua University, Beijing, P. R. China

Session RC1 Emerging Applications II
Chair: Peter Wide, Orebro University, Sweden

Session RC2 Classification II
Chair: M. Hinman, Air Force Research Laboratory, USA

Session RC3 Sensor Fusion for Automatic Target Recognition
Chair: Erik Blasch, Air Force Research Laboratory, USA

* Manuscript not received or withdrawn.

Session WB3
Multisensor-Multisource Fusion for Object Tracking and Recognition
Chair: Shishir Shah
Wayne State University, MI, USA

COMPUTER-ASSISTED MULTISENSOR SYSTEM FOR SURVEILLANCE (CAMS)

Bozzo Alessandro
Elettronica S.p.A.
Via Tiburtina V. Km. 13.700
00135 Roma
ITALY
Email: alexboz@tin.it

Menegotti Ubaldo
Elettronica S.p.A.
Via Tiburtina V. Km. 13.700
00135 Roma
ITALY
Email: eltyb@tin.it

Pillinini Paolo
Elettronica S.p.A.
Via Tiburtina V. Km. 13.700
00135 Roma
ITALY
Email: pablinho@tiscalinet.it

Abstract

Surface areas, such as airports, harbors, subject to illegal activities, violations of navigation laws and possible accidents require a constant and effective surveillance effort. This can be achieved through a ground-based surveillance network consisting of various types of sensors managed by suitable control centers; the objective being to provide a prompt detection of unusual or unexpected events, to optimize the available resources and to support the selection and implementation of pre-defined emergency programs. This paper presents a suitable network structure, with its technical characteristics, related to existing equipment used in civilian applications.

Keywords

Multisensor, Data fusion, Surveillance.

1. Introduction

A wide and effective surveillance of critical surface areas (harbours, straits, etc.) is essential to:

- Providing maritime traffic control in order to prevent collisions (e.g., between ships, running aground, striking reefs and structures, etc.);
- Enforcing anti-pollution laws;
- Enforcing navigation laws.
- Organizing and supporting search and rescue operations (ship wrecks, accidents, etc.);
- Preventing accidental environment pollution and supporting any restoration efforts should such events occur;
- Detecting and countering illegal activities (smuggling, narcotics, illegal immigration, etc.).

The above activities should be exploited in any in critical weather situations.

The continuous surveillance of extensive areas will not be effective if assigned solely to naval and airborne patrol units, but must rely on the support of an integrated network of diversified ground-based multi-sensor elements.

2. System Composition

In order to guarantee its effectiveness in any critical conditions (of weather and traffic) and to monitor both cooperative and non-cooperative units, the system needs to receive data from different kinds of sensors. Moreover the data received have to be collected and managed by an integrated system.

As shown above, the CAMS system basically exploits the following elements:

- differential GPS-based location system,
- DF network,
- radar,
- weather station,
- set infrared sensors
- video-camera
- Control Center Unit.

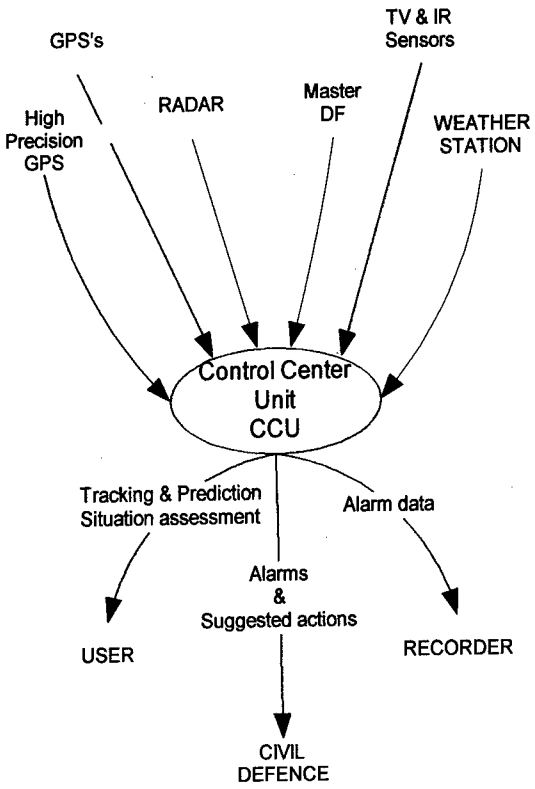


Figure .1 CAMS System Block Diagram

Now we can see a brief description of each element.

2.1. Differential GPS (DGPS)

The DGPS-based technology is very useful for equip all mobile units permanently deployed in the controlled area. Also the mobile units not located permanently deployed in the area, could be temporarily equipped with a portable GPS radio link. The Control Center Unit is supplied with a high performance GPS receiver that receives information from each GPS equipment and processes it using differential algorithms. Moreover, each GPS performs (real time) the adjustments based on the well-known position of the reference GPS receiver and other information (ephemeredes).

This kind of technology can guarantee location errors below several centimeters.

2.2. DF Stations

The main functions of the DF stations are to intercept, determine DOA, monitor and, if requested, record (audio) selected emissions of interest, even of short duration, in the V/UHF bands and, if necessary, also in the HF band (communication channels used by large naval units).

Emitter fixing is also possible, provided the DF Stations co-operate. For this purpose, the DF stations are grouped together, and each group is connected to a Master Station. The Master Station is linked to the Control Center Unit.

2.3. Radar

This sensor operates in the S and X bands and is used for detecting, locating and tracking targets in the assigned operations area (even if such targets maintain total radio/radar silence). With an antenna height of 100m, this radar can ensure detection ranges of 45- 50 Km in the S band for average size naval units (RCS equal to 1000 sqm), and 20 - 25 Km in the X band for small size naval units (RCS equal to 10 sqm).

2.4. Weather station

The weather station supplies the system with data relating to sea, wind and visibility conditions. This kind of data is very useful both to determine safety range to control critical parameter and to record weather conditions for statistical analysis of critical events.

2.5. Infrared sensors

The operator uses the infrared sensors for the purpose of sorting tracks (i.e. mobile unit discrimination), for accurate target identification (optical fingerprinting) and in order to reveal any illegal activities. In fact, the very high angular discrimination afforded by this equipment can be an invaluable asset in separating close targets, in critical situations not easily distinguished by other sensors, and in identifying target shape and features. Finally, more generally, these kinds of sensor allow a night vision control of the area.

2.6. Video-camera

Similarly to the infrared sensor, the video camera is very useful to sort tracks and for optical target identification in low light conditions.

2.7. Control Center Unit (CCU)

Control Center Unit is the heart of the system, and it is where the data are collected, fused and stored and where almost all the processing is performed. Further details are given in the following. It consists of:

- Processing Unit
- Operator Console (with display, keyboard, trackball/mouse)
- TV Monitor and controls
- IR Monitor and controls.

3. CCU measure management

All sensors are connected to the Control Center Unit (CCU) via radio link or via cable. The CCU's purpose is to process the data received from sensors in order to detect, identify, track and estimate the position and the main kinetic parameters of the mobile units, as well as supply scenario assessment and to support decision-making process.

The use of diversified sensor in surveillance systems makes it possible to compensate the weak points if some with the strong points of others and provides further redundancy. This approach increases system robustness.

To fully exploit the collected information, all incoming fragments and packets of information must be synergistically combined. Below is a brief description of the CAMS data fusion process.

Once the initialization step is completed, the data fusion integration can process the data. The assessment is basically a step process.

Each sensor supplies a different set of measurement relating to the intercepted mobile unit: radar furnishes distance and the azimuth angle, while DGPS and DF supplies latitude and longitude. A further difference is the refresh rate, which is different from a sensor to other. So the first step is to *normalize* the received measurements in terms of type and time.

All measurements are converted into standard latitude and longitude data using normal transformation formula. In particular, the measurements received from radar are converted and covariant matrix (later needed in the fusion process) is generated.

To calculate the position of each mobile unit at the CCU's refresh time, all measurements are linear extrapolated using the following formulas.

Let

$X_m(T_{ki}) = [N_m(T_{ki}), N'_m(T_{ki}), E_m(T_{ki}), E'_m(T_{ki}), D_m(T_{ki}), D'_m(T_{ki})]$

the state vector of a mobile unit "m" at T_{ki} instant where the letters N-E-D identify the position and the letters N'-E'-D' the speed.

Let T_{ki} the instant when the sensor furnished its report, with $T_k >= T_{ki}$.

Let T_k be the instant when the CCU updates all data.

The new position vector at the instant T_k, that is

$X_m(T_k) = [N_m(T_k), E_m(T_k), D_m(T_k)]$,

could be obtained using the following formulas:

$N_m(T_k) = N_m(T_{ki}) + N'_m(T_{ki}) * (T_k - T_{ki})$

$E_m(T_k) = E_m(T_{ki}) + E'_m(T_{ki}) * (T_k - T_{ki})$

$D_m(T_k) = D_m(T_{ki}) + D'_m(T_{ki}) * (T_k - T_{ki})$

All measurements received have to be used for updating the exiting tracks or for initializing new tracks. Track updating process begins with a *"gating procedure"*. This technique is used to eliminate unlikely observation-to-track-pairing. A gate is located around the predicted track position. Then, if a single measurement is within the gate, and if it is not within the gate of any other track, the measurement will be correlated with the track and used to update the track. If more than one measurement is within the gate, or worst, if it is within the gates of more than one track, further *correlation* logic is required.

There are basically two approaches to solving resolution. First, the "nearest-neighbor" (NN) approach looks for a unique pairing, so that, at the most, one measure can be used to update a given track. In the second approach, "all neighbors" (AN) allows a track to be updated with a combination of all measurements within its gate. We have followed the first approach.

First of all it is necessary to calculate a distance between measurement and the predicted position of track. Since, as described below, Kalman filter technique will be used, the covariance matrix is used to calculate this quantity in a statistical way. The *assignment* method is used to compose a matrix arranging tracks in one dimension and the measurements in the other dimension, and with non-zero elements as the appropriate distance functions. Zero elements are set as unacceptable because out of gating criteria. The optimal solution will give the maximum number of possible assignments. Even for simple cases, such as three conflicts, the enumerative method is too time consuming to be implemented. In this application the Munkres optimal assignment algorithm, modified by Burgeois and Lassalle has been used.

After the assignment process, using the well-known Kalman filter technique, is possible to *determine the present state* and *predict* the future state of each track, in terms of position and speed.

In a more complex situation *(i.e. to track mobiles units with different kinematics characteristics)* a bank of Kalman filters could represent a significant improvement that can offer unique advantages over the single Kalman filter approach. Each filter could be tuned to a particular combination of target class and operational parameters.

Finally, if possible, the tracks are *identified* by comparing the obtained data with the data stored in a set of libraries, either automatically or by operator aid.

The CCU data management process is summarized in the following diagram.

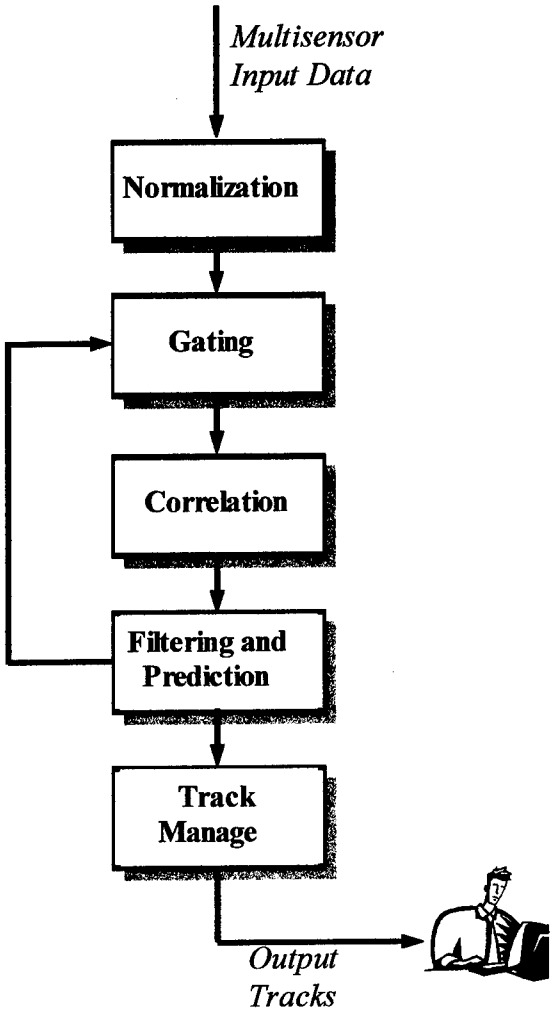


Fig.2 CCU Data Management Block Diagram

4. CAMS main features

The main characteristics of CAMS are summarized below.

- ✓ System configuration guarantees effectiveness in critical weather situations and crowded area and for cooperative and non-cooperative mobile unit. Heterogeneous types of survey allow this performance.
- ✓ The situation is represented on a geographic map display, where the operator has a clear and overall view of the controlled area.

- ✓ The operator can interact with the geographic display map using the standard tools: panning, zooming, scrolling, etc....
- ✓ CAMS is able to support query both geographic and numeric queries. So the operator can submit heterogeneous queries with geographic conditions (e.g. all dangerous area) and numeric conditions (all mobile unit with speed greater then a fixed value), related by logical operator (and, or, not).
- ✓ All resources (people, vehicle, etc..) are under control, so the system can provide aid to the needful and re-routing if necessary;
- ✓ The system is able to control resources the situation in real-time. It can analyze the values of the significant parameters, such as the distance between mobile units, distance between mobile unit and reefs or obstacles, unit velocities in a given area, to prevent accident and unattended situations.
- ✓ Based on previous analysis, the system automatically produces the safety range for controlled parameters, in relation to weather and other conditions. If one or more parameter value falls out of the calculated range, a warning or an alarm is generated.
- ✓ When a warning or alarm is generated the CAMS supports the operator in his decision-making processes in order to properly manage the situation. This is done relating the current situation with a set of libraries of standard situations and historical situations. After this diagnoses an actual-situation-score is generated and the right actions are proposed to operator to face the problem. The action selection is based on a set of libraries of planned actions.
- ✓ When an unattended situation is detected, all relevant data are automatically recorded, such as weather conditions, number of units involved, speed of each mobile unit, mutual distance and so on. Data are also recorded under operator command.
- ✓ Recorded data contribute to building historic archive. The CAMS supports a statistical analysis process of critical situations. It can sort data on the basis keywords: date, kind of situation, wind speed, wind direction, types of resources used and their technical features, etc... Afterwards results are shown in a tabular form, or plotted. Moreover, this data can be used to reproduce the situation of interest in an off-line system such as simulator.
- ✓ CAMS has been designed to be upgraded with an existing Route Planning Module. This module is very useful in furnishing suggestions that can improve overall effectiveness.

Figure 3 shows the main screen mask used by the operator to control the situation and interact with the system. Special attention has been assigned to the man/machine interface to assure a clear and prompt understanding in order to reduce reaction time. Other information in accessible through secondary masks echoed on the main mask.

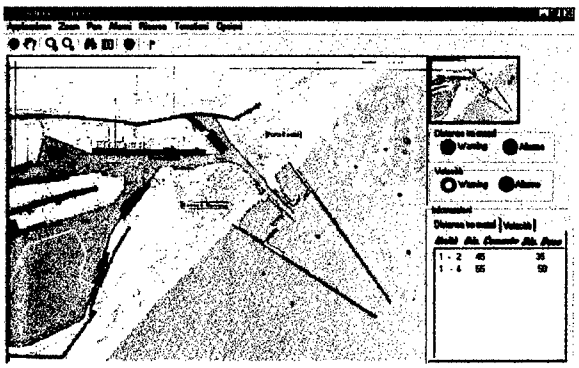


Fig.3 CAMS main mask

5. Application

The system has been designed to be used in Ravenna's harbour. This is located in the north of Italy and is characterized by very unique topographic, meteorological and traffic conditions.
The area to control is fairly wide and includes a roadstead where the ships wait for the permission to

dock and a narrow channel that leads to quay. This channel is split up into seven sections and each section requires different navigation permission depending on to ship tonnage and dimensions.

Its geographic location (in the north of Italy, near Venice) is strongly affected by the weather conditions. For most of the day during the year, the weather is foggy rainy and, generally visibility is usually reduced.

Finally another important aspect is the type of traffic characterizing the harbor.

Ravenna harbor is one of the most important commercial harbors of the Adriatic Sea, and is expected to become even more crowded. The following tables give an idea of the traffic growth in terms of tonnage, ships average tonnage and goods tonnage.

	1990-1997
Number of Ships	+16.3 %
Average Ships Tonnage	+5.5%
Goods Tonnage	+35.6%

Tab.1 Ravenna harbor traffic variations

To complete the above percentage information it is necessary to consider that the number of ships passed through the harbor during the period from 1990 to 1997 is 39.500 units.

Moreover it is very important to consider the growth of dangerous goods passing through the harbor as shown in the following table.

	1990-1997
Petroleum products	+26.4%
Chemical products	+3%
Total	+22.7

Tab.2 Dangerous goods variations

The above considerations and on the basis of the analysis of the accident occurred, illustrated in the following table, clearly highlight the need to have an integrated and efficient control system.

This table shows the number and types of accidents that occur in the harbor in the period 1990-1997.

	1990-1997
Fire, explosion	13%
Collision ship/ship	27%
Collision ship/ barrier	23%
Criminal intent	37%

Tab.3 Accident percentage

To gain a more comprehensive overview of the harbor condition, it is necessary to consider all events relating to illegal activities, violation of navigation rules, as well as violations of environment safety laws.

6. Conclusions

The use of a ground-based surveillance network, similar to the one described above, consents:

- surveillance of assigned areas, prompt reporting of any situation changes or targets of interest (continuous monitoring);
- organization of Search And Rescue (SAR) operations support;
- planning and scheduling support mobile units, to minimize for each ships waiting time in the roadstead in order to reduce costs;
- useful aid in decision-making processes during emergency conditions and providing a prompt reaction;
- reducing the number of necessary units and optimization their employment;
- reducing patrol mission duration and economizing on material and human resources;
- generating reports and official documents on various situations and events;
- communicating alarm to other departments (with different assignments).

Also, with special attention to unusual, or unexpected events, the use of computer-assisted

multisensor system for surveillance allows to process and elaborate statistical data concerning:

- available equipment (communications, radar systems) and related technical characteristics (electrical parameters)
- typical operational patterns: such as, number of units involved, mobile units reciprocal distances, distances between mobile units and reefs, or obstacles, mobile units route and speed, weather conditions (wind speed, wind direction, etc...)

References

[1] R. T. Antony "Principles of Data Fusion Automation" Artech House 1995.

[2] Y. Bar-Shalom "Multitarget-Multisensor Tracking: Advanced Applications" Artech House 1990.

[3] Y. Bar-Shalom and Xiao-Rong Li "Estimation and Tracking: Principles, Techniques, and software" Artech House 1993.

[4] S. S. Blackman "Multiple-Target Tracking with Radar Applications" Artech House 1986.

[5] F. Burgeois and J.C. Lassalle "An extension of the Munkres Algorithm for the Assignment Problem to Rectangular Matrices" Communications of the ACM, Vol 14, Dec 1971, pp 802-806.

[6] A. Farina and F.A. Studer "Radar Data Processing Vol 1 – Introduction and Tracking" Research Studies Press LTD. 1985.

[7] A. Farina and F.A. Studer "Radar Data Processing Vol 2 – Advanced Topics and Applications" Research Studies Press LTD. 1985.

[8] Knut "Fundamental Algorithms Second Edition" The art of Computer Programming, 1973.

[9] H.W. Kuhn "The Hungarian Method for the Assignment Problem" Naval Research Logistic Quarterly, N°2, 1995, pp 83-97.

[10] Jijie Zhu. "Conversion of earth-centered earth-fixed coordinates to geodetic coordinates", IEEE Transactions on aerospace and electronic system. Vol. 30, N° 3 July 1994, pp. 957-962.

Use of Hyperspectral Data with Intensity Images for Automatic Building Modeling*

A. Huertas[1], R. Nevatia[1] and D. Landgrebe[2]

[1]Institute for Robotics and Intelligent Systems
University of Southern California
Los Angeles, California 90089-0273
huertas|nevatia@iris.usc.edu

[2]Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana 47907-1285
landgreb@ecn.purdue.edu

Abstract

Geospatial databases are needed for many tasks in civilian and military applications. Automated building detection and description systems attempt to construct 3-D models using primarily PAN (panchromatic) images. These systems can make use of cues derived from other sensor modalities to make the task easier and more robust. The recent development of hyperspectral sensors such as HYDICE (HYperspectral Digital Imagery Collection Experiment) can provide reasonably accurate thematic maps. Such data, however, tends to be of lower resolution, have geometric distortions and camera models are needed to map points between the different sensors. We use the thematic map to provide cues for presence of buildings in the PAN images for accurate delineation. It is shown that such cues can not only greatly improve the efficiency of the automatic building detection system but also improve the quality of the results. Quantitative evaluations are given.

Key Words: Information Integration, Sensor Fusion, HYDICE, Hyperspectral data, 3-D Building Modeling, Thematic Map.

1 Introduction and Overview

Three-D models of man-made structures in urban and sub-urban environments are needed for a variety of tasks. The principal sensor products used for this task have been panchromatic (PAN) images acquired from an aircraft [Noronha & Nevatia, 1997, Collins et al., 1998, Grün, et al. 1997, Grün & Nevatia, 1998, Paparoditis et al., 1998]. PAN images have many advantages: they are relatively easy to acquire at high resolution (say of the order of 0.5 meters/pixel) and humans find it is easy to visualize them and to extract the needed information from them. However, their use for automatic extraction has proven to be quite difficult. One of the principal causes of this difficulty is the high density of features present in the images. PAN image pixels encode reflected light intensity that gives little information to the nature of the material reflecting it. While it is possible to apply analyses that help recover structure from image elements, the problem of segmenting aerial scenes accurately remains a challenge.

* This research was supported in part by the U.S. Army Research Office under grant No. DAAH04-96-1-0444.

In recent years, advances in the solid state electronics have made possible the construction of hyperspectral sensors with an orders of magnitude increases in the number of bands possible, while at the same time providing improved signal-to-noise ratios. One such sensor, called HYDICE collects data of 210 bands over the range 0.4-2.5 μm with a field of view 320 pixels wide at an IFOV (pixel size) of 1 to 4 m depending on the aircraft altitude and ground speed. Given the spectral detail in such data it becomes practical and effective to construct a *thematic map* of an area that shows the layout of the various types of land cover and distribution of various materials in the scene.

In this paper, we focus on the task of building detection and reconstruction with the assistance of corrected and geo-referenced thematic maps derived from HYDICE data. The complementary qualities of conventional images and HYDICE image data provide an opportunity for exploiting them in different ways to make the task of automatic feature modeling easier.

Combining the two data sources at the pixel level is difficult as there is not a one-to-one correspondences between the pixels in the two sources, in general; hyperspectral data poses major challenges in terms of geometric corrections and terrain normalization. Instead, we propose to extract information from each which is then combined and perhaps used to guide extraction of additional information. In particular, we feel that the HYDICE data is suited for *detecting* possible building locations as buildings may be characterized by their roof materials. However, hyperspectral analysis results in a label for each pixel, but does not, by itself, combine pixels into objects such as buildings. HYDICE image data tends to be of lower resolution than conventional PAN images. Object boundaries are not likely to be precise and it may be difficult to distinguish a building from other nearby objects such as roads. PAN images, with much higher resolution can provide precise delineation as well as distinguish a building from other high objects much more reliably.

In the next sections, we describe how thematic maps are derived and how useful *cues* can be extracted from the HYDICE data. Use of these cues in the building extraction process is then described. Results comparing the effects of these cues are presented in section 4. Other approaches to use of HYDICE data may be found in [Ford et el., 1998, Bea & Healey, 1998, Healey, 1999, Madhok & Landgrebe, 1999].

2 Thematic Maps from HYDICE Data

The intent for multispectral and hyperspectral image data analysis is to rapidly and inexpensively associate a ground cover label to each pixel in the image. Given the multivariate nature of such data, the process of data analysis is one of dividing up the N-dimensional feature space into M exhaustive but non-overlapping regions where M is the number of classes of materials existing in the scene. The process involves defining the M classes of interest in a quantitative fashion, such that each pixel in the scene, which exists as a discrete location in the N-dimensional space, can be uniquely associated with one of the M classes. Frequently, this is done by using a small number of samples in the scene, called design samples or training samples, to define an N-dimensional probability density function for each of the M classes. Then an unknown pixel can be evaluated in terms of the likelihood of each possible class to determine the most likely class membership.

The onset of high dimensional hyperspectral data, on the one hand, greatly increases the potential of such a process. However, it has also introduced significant new challenges to the analysis process to achieve this potential, because such high dimensional feature spaces are much more complex. Not only can a 210-dimensional probability distribution not be visualized, but even the ordinary rules of geometry of 2- or 3-dimensional space do not apply in such high dimensional spaces [Lee & Landgrebe, 1993; Jimenez & Landgrebe, 1998]. Much progress has been made in recent years in understanding such high dimensional spaces and in devising effective analysis procedures for them [Landgrebe, 1999]. The following example from Fort Hood, Texas, will serve to illustrate some of the tools available for this process.

In this case, bands in the regions where the atmosphere is opaque were not considered and 171 bands in the 0.4 to 2.45 μm region of the visible and infrared spectrum were used. This data set contains 1208 scan lines with 307 pixels in each scan line. It totals approximately 130 Megabytes. The primary intent of the analysis of this data set was to identify rooftops and other impervious materials in the scene. With data this voluminous and complex, one might expect a rather complex analysis process, however, it has been possible to find quite simple and inexpensive means to do so. The steps used and the time needed on an inexpensive personal computer for this analysis are listed in the following table and are briefly described below.

Table 1: Thematic Classification Time

Operation	CPU time	Analyst time
Display Image	15 sec.	
Define Classes		30 min.
Feature Extraction	11 sec.	
Reformat	117 sec.	
Classification	86 sec.	
Total	229 sec.	30 min.

Define Classes

A software application program called MultiSpec, available to anyone at no cost from http://dynamo.ecn.purdue.edu/~biehl/MultiSpec/, was used. The first step is to present to the analyst a view of the data set in image form so that training samples, examples of each class desired in the final thematic map, can be marked. A simulated color infrared photograph form is convenient for this purpose; to do so, bands 60, 27, and 17 are used in MultiSpec for the red, green, and blue colors, respectively. The image is shown in Figure 1. (Color versions of the figures in this paper are available at http://iris.usc.edu/home/iris/huertas/www/hydice.)

Feature Extraction

After designating the training areas, a feature extraction algorithm is applied to determine a feature subspace that is optimal for discriminating between the specific classes defined. The algorithm used is called Discriminate Analysis Feature Extraction (DAFE). The result is a linear combination of the original 171 bands to form 171 new bands that automatically occur in descending order of their value for producing an effective discrimination. From the MultiSpec output, it is seen that the first 15 of these new features should be adequate for successfully discriminating between the classes.

Reformatting

The new features defined above are used to create a 15 band data set consisting of the first 15 of the new features, thus reducing the dimensionality of the data set from 171 to 15.

Classification

Having defined the classes and the features, next a classification is carried out. The algorithm in MultiSpec used was the standard Gaussian maximum likelihood algorithm in which the mean vector and covariance matrix for each class are estimated from the training samples. These

Figure 1. Simulated infrared image using HYDICE bands 17, 27 and 60.

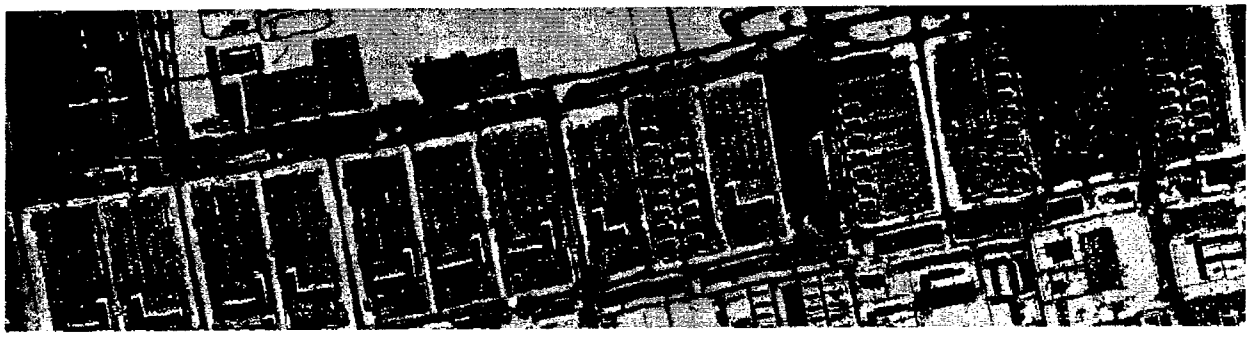


Figure 2. Thematic Map of the classes roof (black), road, lot, field (grays) and shadow (white)

estimates then allow calculating the likelihood of each class for a given pixel. The label of the most likely class is assigned to the pixel.

Hyperspectral data provides the capability to discriminate between nearly any set of classes. Research has shown that, of all the variables to the data analysis process, the most important one is the size and quality of the classifier training set. There are a number of additional steps that could be taken to further polish the result, but the current result appears to be satisfactory for the current use.

3 Integration of HYDICE and PAN Information

In order to integrate cues extracted from HYDICE data into the building detection and description system we require that the thematic map be rectified and registered to the PAN imagery as described next.

Geometric Rectification

Geometric rectification is needed to correct for the oscillations and "waviness" introduced by the nature of the HYDICE pushbroom sensor. Rectification is performed on the thematic map rather than on the hydice data directly. The method utilizes ground control points and control linear features typically found in urban scenes together with the pushbroom sensor model and a gauss-markov platform model to yield coordinate relationships between ground and image spaces. See [Lee, et al. 1999] for details. The accuracies achieved are in the 0.5 to 1 pixel range. Figure 3 shows a geometrically rectified thematic map of a portion of the Ft. Hood site. Note the straight roads. The waviness of the image boundaries gives an idea of the extent of rectification required.

Registration with PAN Images

The corrected thematic map has the geometric characteristics of an orthographic projection. The estimation of the sensor parameters, or "camera" model, associated with this overhead (nadir) viewpoint is straightforward. The camera model allows us to derive the appropriate 3D- to-2D and 2D-to-image transforms needed to register the available PAN images to the thematic map. We use these transforms to project EO 2-D and 3-D features onto the thematic map to assist and support the building detection system at various stages of processing. We describe in more detail, and illustrate these processes, with an example, below in section 4.

Cue Extraction

Figure 4 shows some of the barrack buildings in Fort Hood, Texas. The corresponding thematic map is shown in Figure 5. We first extract the roof pixels from the thematic map. These are shown in Figure 6. Many pixels in small regions are misclassified or correspond to objects made of similar materials as the roofs. The building cues extracted from this image are the connected components of certain minimum size. These components are shown

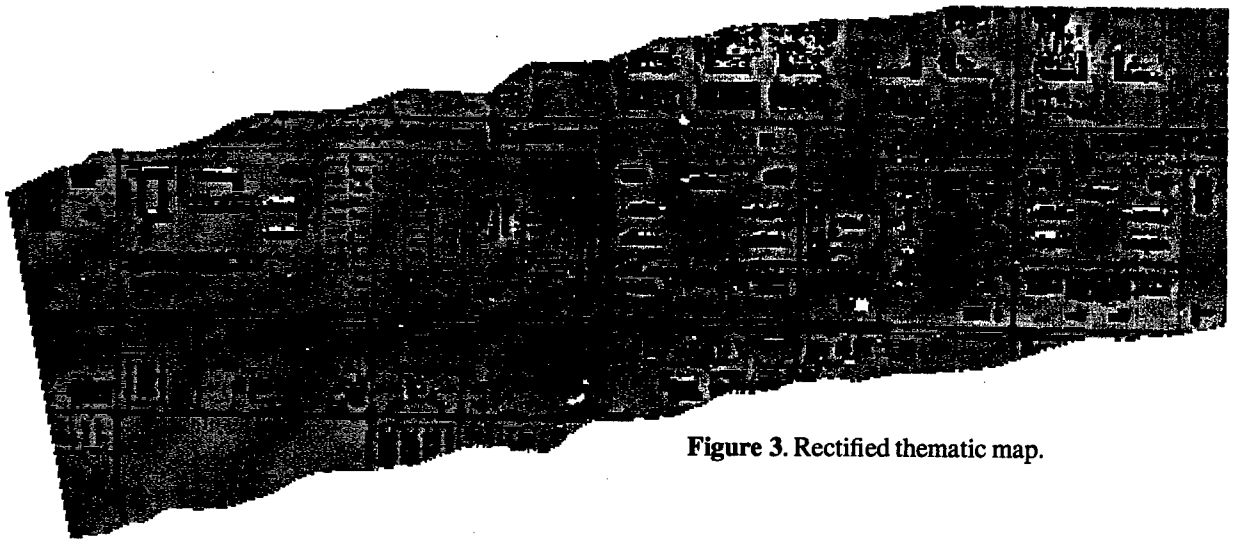


Figure 3. Rectified thematic map.

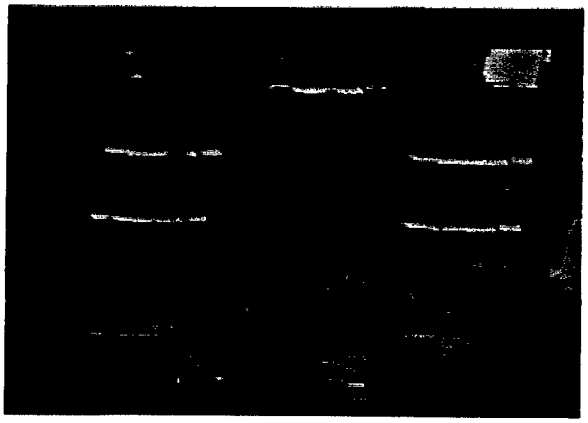


Figure 4. Barrack buildings at Fort Hood

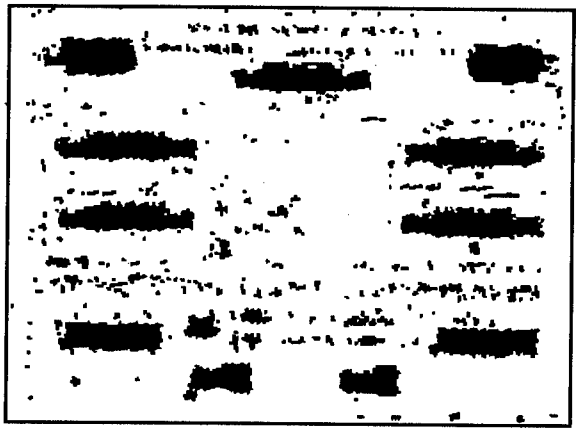


Figure 6. Roof Class from Thematic Map

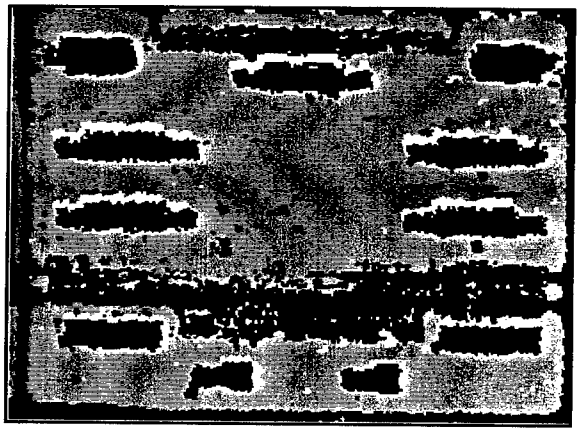


Figure 5. Thematic map.

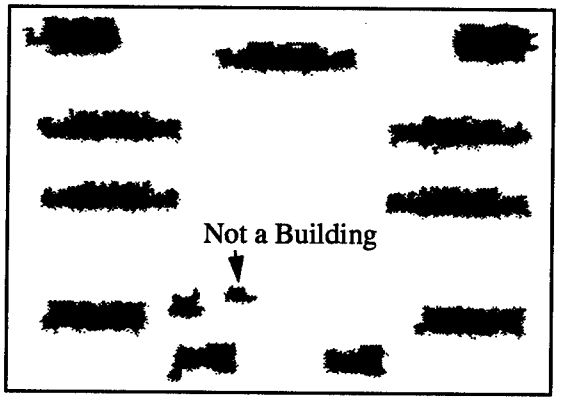


Figure 7. Building Cues

in Figure 7; Except for one region, these components correspond to building roofs.

4 Multi-View System

We next describe the use the HYDICE cues in the multi-view building detection system described in [Noronha & Nevatia, 1997]. This system has three major phases: hypothesis formation, selection and validation. This system assumes that the roofs of buildings are rectilinear though the roofs need not be horizontal (some forms of gables are allowed). Hypotheses are formed by collecting a group of lines that form a parallelogram in an image. Multiple images and matches between lines are used in the hypotheses formation stage. As line evidence can be quite fragmented, liberal parameters are used to form hy-

potheses. Properties of resulting hypotheses are used to select among the competing hypotheses. The selected hypotheses are then subjected to a verification process where further 3-D evidence, such as presence of walls and predicted shadows are examined.

The cues extracted from the HYDICE data can help improve the performance of the building description system at each of the three stages described above. We show some details and results of these processes.

Hypothesis Formation

Cues can be used to significantly reduce the number of hypotheses that are formed by only considering line segments that are within or *near* the cue regions. The 3-D location of a line segment in the 2-D PAN images is not known. To determine whether a line segment is near a HYDICE cue region we project the line onto the cue image at a range of heights, and determine if the projected line intersects a cue region. Figure 8 shows the line segments detected in the image of Figure 4 (using a Canny edge detector); Figure 9 shows the lines that lie near the HYDICE cues. As can be seen, the number of lines is reduced drastically (84%) by filtering without loosing any of the lines needed for forming building hypotheses.

This not only results in a significant reduction in computational complexity but many false hypotheses are eliminated allowing us to be more liberal in the hypotheses formation and thus including hypotheses that may have been missed otherwise.

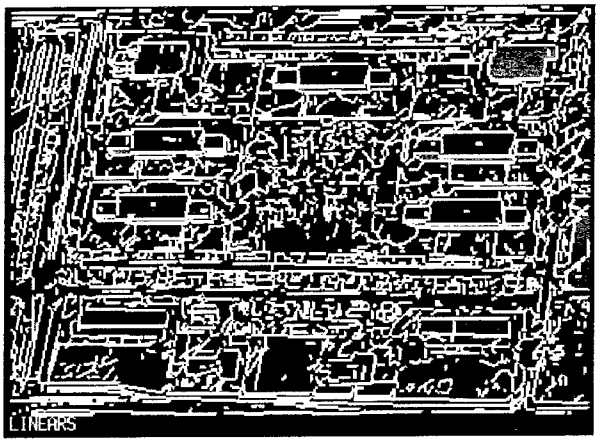


Figure 8. Line segments from PAN image.

Hypothesis Selection

The building detection system applies a series of filters to the hypotheses formed. The remaining hypotheses are then evaluated in the basis of the geometric evidence (underlying line segments that support the hypothesized roof boundaries), in an attempt to select a set of "strong" hypotheses. With HYDICE cues available we skip the initial filtering stages and introduce cue evidence into the

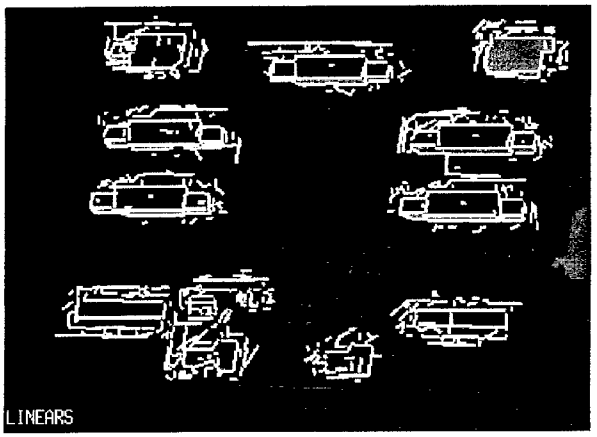


Figure 9. Lines near HYDICE cues.

roof support analysis. The evidence consists of support of a roof hypotheses in terms of the overlap between the roof hypotheses and the HYDICE cue regions. The hypotheses are constructed from matching features in multiple (two in this example) images and are represented by 3-D rectilinear components in 3-D world coordinates. We can therefore project them directly onto the HYDICE cues image to compute roof overlap (See Figure 10). The system requires that the overlap be at least 50% of the projected roof area.

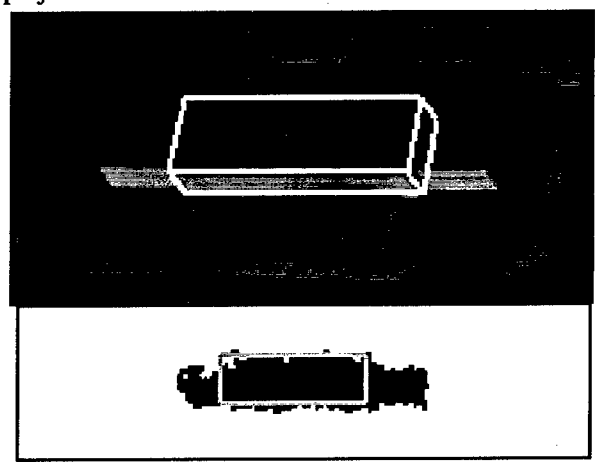


Figure 10. A 3-D hypotheses projected on PAN image (top) and on cue image.

Hypotheses Validation

Just as poor hypotheses can be discarded because they lack HYDICE support, the ones that have a large support see their confidence increase during the verification stage. In this stage, the selected hypotheses are analyzed to verify the presence of shadow evidence and wall evidence. Details of the shadow and wall analysis are given in [Lin et al., 1994]. When no evidence of walls or shadows is found, we require that the HYDICE evidence (overlap) be higher, currently 70%, in order to validate a hypotheses. The 3-D Models constructed with HYDICE

support from the validated hypotheses are shown in Figure 11. For comparison, the model shown in Figure 12 was derived without HYDICE support. Note that false detections are eliminated with HYDICE cueing. Also, the object cue on the lower middle in (Figure 7) is not found to be a building, even with HYDICE support, as the lack of geometric evidence prevented a hypothesis to be formed there. Also, the building components on the top left and on the lower left are not found without HYDICE support but found with it.

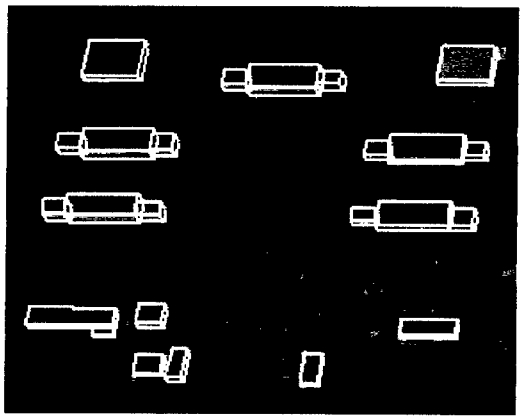


Figure 11. 3-D HYDICE assisted model

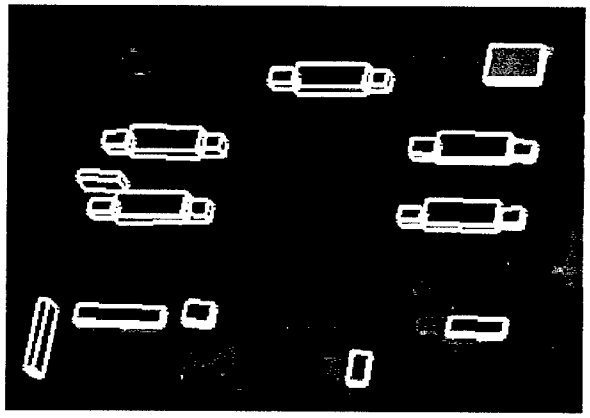


Figure 12. Model from PAN without HYDICE cueing

Once a 3-D model of the buildings is obtained, it is possible to reclassify the roof pixels in the thematic map more accurately by improving the delineation of the roof pixels boundaries and marking missclassified pixels (see Figure 13.)

An evaluation of the quality of results is given next.

5 System Evaluation

Table 2 gives a comparison of the number of features and final result component counts with and without use of HYDICE cues for the Fort Hood example. The two figures given for the line segments and linear structures correspond to the two images that were used, one of which was shown earlier in Figure 4.

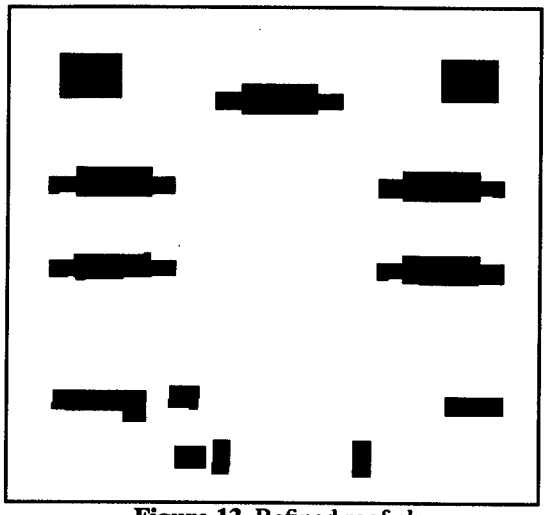


Figure 13. Refined roof class

Table 2: Execution Statistics

Feature	PAN Only	With HYDICE
Line Segments	15938/6976	
Linear Structures	6363/2693	796/652
Hypotheses	3793	636
Selected hypotheses	273	172
Verified hypotheses	115	127
Final hypotheses	20 (2 false)	24 (0 false)

To characterize the increase in performance of the system when HYDICE cues are available we use two basic metrics (see [Nevatia, 1999] for details), detection rate and false alarm rate, as follows:

$$\textit{Detection Rate} = \frac{TP}{(TP+FN)}$$

$$\textit{False Alarm Rate} = \frac{FP}{(TP+FP)}$$

TP, **FP** and **FN** stand for true positives, false positives and false negatives. Note that with these definitions, the detection rate is computed as a fraction of the reference features whereas the false alarm rate is computed as a fraction of the detected features.

In the definitions given above, a feature could be an object, an area element or a volume element. The first level of evaluation is to measure the detection and false alarm rates at the object levels such as for buildings or wings of a complex building. We consider each rectangular part of a rectilinear building as a separate object. A building object will be considered to be detected, if *any* part of it has been detected. Consider the reference model shown in Figure 14. Table 3 shows a summary of detection and

false alarm results for the Ft. Hood example in terms of object parts.

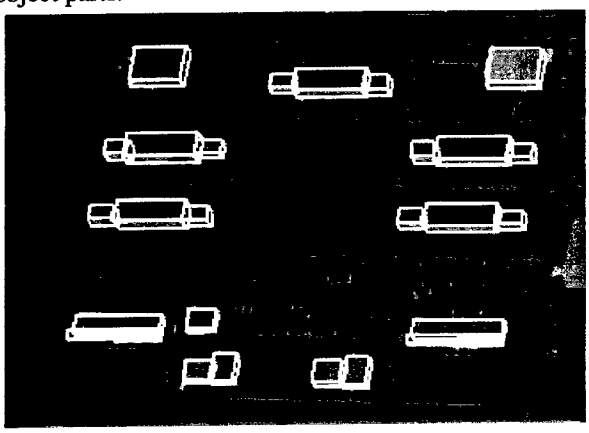


Figure 14. Reference model for evaluation.

Table 3: Component Evaluation

	PAN only	With HYDICE
Reference Model	26	
TP	20	25
FP	2	0
FN	6	1
Detection Rate	0.769	0.961
False Alarm Rate	0.09	0.00

To better reflect the quality of the detected components we also compute the accuracy the overlap between the footprints of the detected and the reference models and in the overlap between the 3-D volume occupied by them.

The area (volume) elements of the reference model that overlap with some area (volume) element of an extracted model can be considered to give the true positive (**TP**) values for the area (volume) elements of the reference model (the remaining elements of the reference models are the false negatives, **FN**). The area (volume) elements of the extracted model that do not overlap with any area (volume) element of the reference model give us the false positives (**FP**) for the area (volume) elements of the extracted model.

One way to combine the results of the above area (or volume) overlap analysis is to consider each area element as an object and count the detection and false alarm rates for all the area elements in the models. Table 4 shows these results for our Ft. Hood example. Ground detection rate is computed for the ground area elements (all elements that are not part of other objects); ground false alarm rate is not shown.

To better characterize the accuracy, we compute the detection rates for the area elements of each reference

Table 4: Ft. Hood Combined Area Evaluation

	PAN Only	with HYDICE
Detection rate	0.7116	0.8453
False Alarm rate	0.1510	0.0768
Ground Detection rate	0.9819	0.9907

building component and the false alarm rates for each extracted building component separately. To visualize the result we compute a cumulative distribution of the detection and false alarm rates. Specifically, we can compute the percentage of building components of the reference model whose area (volume) elements detection rate (**TP**) is at a give value or *higher*. A curve plotting such a distribution is called a **CDR** curve [Nevatia, 1999]; Figure 15a shows the CDR curve for area elements of our Ft. Hood example. Similarly, we can compute the percentage of the building components of the extracted model whose false alarm rate (**FP**) is at a given value or *lower*. A curve plotting such a distribution is called a **CFR** curve; Figure 15b shows the CFR curve for the *area* elements of our Ft. Hood example. We also compute CDR and CFR curves for the *volume* elements for the reference and extracted building components. These are not shown for lack of space. A CDR curve that is consistently higher than another CDR curve indicates consistently better performance (similarly, a CFR curve that is consistently lower is consistently better).

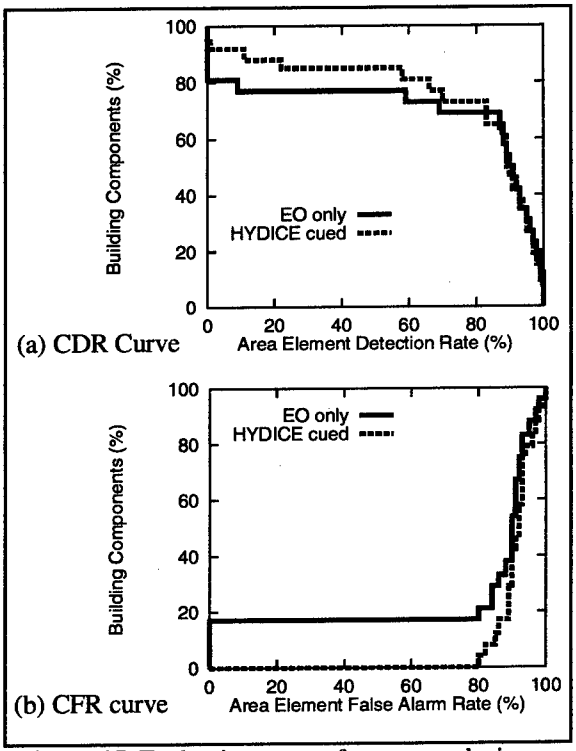


Figure 15. Evaluation curves for *area* analysis.

6 Conclusions

Many challenges remain in terms of data normalization and sub-pixel image registration for successful of data fusion of PAN and HYDICE types of imagery at the sensor level. Hyperspectral data however, provides the capability to discriminate between nearly any set of classes. By introducing an optimal feature design calculation on the 171 bands, we have shown that a good classification of materials can be achieved for production of a thematic map providing effective cues for objects of interest

We have presented a methodology for detection and reconstruction of building structures by using conventional intensity images with cues data derived from HYDICE sensors. Even though the HYDICE data is of a lower resolution and contains some missing elements and artifacts, it has been shown that it can be used to enhance the results of PAN image analysis while substantially reducing the computational complexity. This was accomplished not by combining the information at the sensor level but rather by using analysis of one to guide the analysis of the other. We believe that this paradigm will be suitable for other tasks as well as sensors of different modalities become available for more domains.

References

[Bea & Healey, 1998] T. Bea and G. Healey. "Invariant Subpixel Material Identification in Hyperspectral Imagery," Proceedings DARPA Image Understanding Workshop, Monterey, CA, November, 1998, pp. 809-814.

[Collins et al., 1998] R. Collins, C. Jaynes, Y. Cheng, X. Wang, F. Stolle, A. Hanson, and E. Riseman, "The ASCENDER System: Automatic Site Modeling from Multiple Aerial Images", Computer Vision and Image Understanding Journal, Vol 72, No. 2, November, pp 143-162.

[Ford et al., 1998] S. Ford, C. McGlone, S. Cochran, J. Shuffelt, W. Harvey and D. McKeown, "Analysis of HYDICE Data for Information Fusion in Cartographic Feature Extraction" Proceedings of the International Geoscience and Remote Sensing Symposium, IGARSS'98, Seattle, Washington, July 6-10 1998, Vol. V, pages 2702-2706.

[Grün & Nevatia, 1998] Computer Vision and Image Understanding Journal, Special Issue on Automatic Building Extraction from Aerial Images. A. Grun and R. Nevatia, editors. Academic Press, Vol. 72, No. 2, November.

[Grün et al., 1997] Proceedings of the Ascona Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images II, A. Grün, E. Baltsavias & O. Henricsson, Ed., Brinkhauser Verlag, Switzerland, May.

[Healey, 1999] G. Healey, "Reflectance Estimation and Material Classification for 3D Objects in Aerial Hyperspectral Images," To appear in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, CO, June 1999.

[Jimenez & Landgrebe, 1998] L. Jimenez and D. Landgrebe, "Supervised Classification in High Dimensional Space: Geometrical, Statistical, and Asymptotical Properties of Multivariate Data," IEEE Transactions on System, Man, and Cybernetics, Volume 28, Part C, No. 1, pp. 39-54, February 1998.

[Landgrebe, 1999] D. Landgrebe, Information Extraction Principles and Methods for Multispectral and Hyperspectral Image Data, Chapter 1 of Information Processing for Remote Sensing, edited by C. H. Chen, World Scientific Publishing Co., Inc., River Edge, NJ

[Lee et al., 1999] C. Lee, H. Theiss, J. Bethel, and E. Mikhail, "Rigorous Mathematical Modeling of Airborne Pushbroom Imaging Systems." to appear in the Photogrammetric Engineering and Remote Sensing Journal, 1999.

[Lee & Landgrebe, 1993] C. Lee and D. Landgrebe, "Analyzing High Dimensional Multispectral Data," IEEE Transactions on Geoscience and Remote Sensing, Volume 31, No. 4, pp 792-800, July 1993.

[Lin et al., 1994] C. Lin, A. Huertas and R. Nevatia, "Detection of Buildings Using Perceptual Groupings and Shadows," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, June 1994, pp. 62-69.

[Madhok & Landgrebe] V. MAdhok and D. Landgrebe, "Supplementing Hyperspectral Data with Digital Elevation," To appear in Proceedings of the IEEE International Geoscience and Remote Sensing Symposium," Hamburg, Germany, June 1999.

[Nevatia, 1999] R. Nevatia. "On Evaluation of 3-D Geospatial Modeling Systems," in ISPRS Proceedings of the International Workshop on 3D Geospatial Data Production", Paris, France, April, 1999.

[Noronha & Nevatia, 1997] S. Noronha and R. Nevatia. "Detection and Description of Buildings from Multiple Aerial Images", Proceedings IEEE Conference on Computer Vision and Pattern Recognition, San Juan, PR, June 1997, pp. 588-594.

[Paparoditis et al., 1998] N. Paparoditis, M. Cord, M. Jordan and J.P. Cocquerez, "Building Detection and Reconstruction from Mid- and High-Resolution Aerial Imagery", Computer Vision and Image Understanding Journal, Vol 712, No. 2, November, pp 122-142.

Image Database Indexing using a Combination of Invariant Shape and Color Descriptions *

Ronald Alferez and Yuan-Fang Wang
Computer Science Department
University of California
Santa Barbara, CA, 93106 U.S.A.

Abstract ***Image and video library applications are becoming increasingly popular. The increasing popularity calls for software tools to help the user query and retrieve database images efficiently and effectively. In this paper, we present a technique which combines shape and color descriptors for invariant, within-a-class retrieval of images from digital libraries. We demonstrate the technique on a real database containing airplane images of similar shape and query images that appear different from those in the database because of lighting and perspective. We were able to achieve a very high retrieval rate.***

Keywords: images, video, libraries, features

1 Introduction

Image and video library applications are becoming increasingly popular as witnessed by many national and international research initiatives in these areas, an exploding number of professional meetings devoted to image/video/multi-media, and the emergency of commercial companies and products. The advent of high-speed networks and inexpensive storage devices has enabled the construction of large electronic image and video archives and greatly facilitated their access on the Internet. In line with this, however, is the need for software tools to help the user query and retrieve database images efficiently and effectively.

Querying an image library can be difficult and one of the main difficulties lies in designing powerful features or descriptors to represent and organize images in a library. Many existing image database indexing and retrieval systems are only capable of *between-classes* retrieval (e.g., distinguishing fish from airplanes). However, these systems do not allow the user to retrieve images that are more specific. In other words, they are unable to perform *within-a-class* retrieval (e.g., distinguishing different types of airplanes or different species of fish). This is because the aggregate features adopted by many current systems (such as color histograms and low-ordered moments) capture only the general shape of a class and are not descriptive enough to distinguish objects within a particular class.

The within-a-class retrieval problem is further complicated if query images depicting objects, though belonging to the class of interest, may look different due to non-essential or incidental environment changes, such as rigid-body or articulated motion, shape deformation, and change in illumination and viewpoint. In this paper, we address the problem of *invariant, within-a-class* retrieval of images by using a combination of invariant shape and color descriptors. By analyzing the shape of the object's contour as well as the color and texture characteristics of the enclosed area, information from multiple sources is fused for a more robust description of an object's appearance. this places our technique at an advantage over most current approaches that exploit either geometric information or color information exclusively.

The analysis involves projecting the shape and color information onto basis functions of finite, local support (e.g., splines and wavelets). The projection coefficients, in general, are sensitive to changes induced by rigid motion, shape deformation, and change in illumination and perspective. We derive expressions by massaging these sets of projection coefficients to cancel out the environmental factors to achieve invariance of the descriptors. Based on these features, we have conducted preliminary experiments to recognize different types of airplanes (many of them having very similar shape) under varying illumination and

*Supported in part by a grant from the National Science Foundation, IRI-94-11330

viewing conditions and have achieved good recognition rates. We show that information fusion has helped to improve the accuracy in retrieval and shape discrimination.

2 Technical Description

In this section, we present the theoretical foundation of our image-derived, invariant shape and color features. Invariant features form a compact, intrinsic description of an object, and can be used to design retrieval and indexing algorithms that are potentially more efficient than, say, aspect-based approaches.

The search for invariant features (e.g., algebraic and projective invariants) is a classical problem in mathematics dating back to the 18th century. The need for invariant image descriptors has long been recognized in computer vision. Invariant features can be designed based on many different methods and made invariant to rigid-body motion, affine shape deformation, scene illumination, occlusion, and perspective projection. Invariants can be computed either globally, such is the case in invariants based on moments or Fourier transform coefficients, or based on local properties such as curvature and arc length. See [3, 4, 5] for survey and discussion on the subject of invariants.

As mentioned before, our invariant features are derived from a localized analysis of an object's shape and color. The basic idea is to project an object's exterior contour or interior region onto localized bases such as wavelets and splines. The coefficients are then normalized to eliminate changes induced by non-essential environmental factors such as viewpoint and illumination. We will illustrate the mathematical frameworks using a specific scenario where invariants for curves are sought. The particular basis functions used in the illustration will be the wavelet bases and spline functions. Interested readers are referred to [1] for more details.

Several implementation issues arise in this invariant framework which we will briefly discuss before describing the invariant expressions themselves. [1]

1. How are contours extracted?

Or stated slightly differently, how is the problem of segmentation (separating objects from background) addressed? Segmentation turns out to be an extremely difficult problem and, as fundamental a problem as segmentation is, there is no fail-proof solution. A "*perfect*" segmentation scheme is like the holy grail of low-level computer vision and a panacea to many high level vision problems as well.

We are not in search of this holy-grail, which, we believe, is untenable in the foreseeable future. In an image database application, the problem of object segmentation is simplified because

- Database images can usually be acquired under standard imaging conditions which allow the ingest and catalog operations to be automated or semi-automated. For example, to construct a database of airplane images, many books on civil and military aircrafts are available with standard front, side, and top views taken against a uniform or uncluttered background. (The above is also true for applications in botany and marine biology.) This allows the contours of the objects of interest to be extracted automatically or with the aid of standard tools such as the flood fill mask in Photoshop. Furthermore, the cataloging operations are usually done off-line and done only once. Hence, a semi-automated scheme will suffice.
- On the other hand, query images are usually taken under different lighting and viewing conditions. Objects of interest can be embedded deeply in cluttered background which makes their extraction difficult. However, we can enlist the help of the user to specify the object of interest instead of asking the system to attempt the impossible task of automated segmentation. A query-by-sketch or a "human-in-the-loop" type solution with an easy-to-use graphics interface and segmentation aids such as the flood fill mask is perfectly adequate here and does not impose undue burden on the user. This proved to be feasible in our experiments.

2. How are contours parameterized?

For a contour based description, a common frame of reference is usually needed that allows point correspondences to be established between two contours for comparison. The common frame of reference comprises a common starting point of traversal, a common direction of traversal, and a parameterization scheme that traverses to the corresponding points in the two contours at the same parameter setting. We will first discuss the parameter-

[1] A word on the notational convention: matrices and vectors will be represented by bold-face characters while scalar quantities by plain-face characters. 2D quantities will be in small letters while 3D and higher-dimensional quantities in capital letters. For example, coordinates (bold for vector quantities) of a 2D curve (small letter for 2D quantities) will be denoted by $\mathbf{c}$.

ization issue and then address the issues of point correspondence and traversal direction.

When defining a parameterized curve $\mathbf{c}(t) = [x(t), y(t)]^T$, most prefer the use of the intrinsic arc length parameter because of its simplicity and the fact that it is either invariant or transforms linearly in rigid-body motion and uniform scaling. However, under more general scenarios where shape deformation is allowed (e.g., deformation induced in an oblique view), intrinsic arc length parameter is no longer invariant. Such deformation can stretch and compress different portions of of an object's shape, and a parameterization based on intrinsic arc length will result in wrong point correspondence.

It is well known that many shape deformation and distortion resulting from imaging can be modeled as an affine transform, through which the intrinsic arc length is nonlinearly transformed [2]. An alternative parameterization is thus required. There are at least two possibilities. The first, called *affine arc length*, is defined [2] as: $\tau = \int_a^b \sqrt[3]{\dot{x}\ddot{y} - \ddot{x}\dot{y}}\, dt$ where $\dot{x}, \dot{y}$ are the first and $\ddot{x}, \ddot{y}$ are the second derivatives with respect to any parameter t (possibly the intrinsic arc length), and (a, b) is the path along a segment of the curve.

Another possibility [2] is to use the *enclosed area parameter*: $\sigma = \frac{1}{2}\int_a^b |x\dot{y} - y\dot{x}|\, dt$. One can interpret the enclosed area parameter as the area of the triangular region enclosed by the two line segments from the centroid of an object to two points a and b on the contour. It can be shown that both these parameters transform linearly under a general affine transform [2]. Hence, they can easily be made absolutely invariant by normalizing them with respect to the total affine arc length or the total enclosed area of the whole contour, respectively. We use these parameterizations in our experiments.

3. How are identical traversal direction and starting point guaranteed?

It will be shown that the *invariant signatures* (to be defined later) of two contours are phase-shifted versions of each other when only the starting point of traversal differs. Furthermore, the same contour parameterized in opposite directions produces invariant signatures that are flipped and inverted images of each other. Hence, a match can be chosen that maximizes certain cross-correlation relations between the two signatures.

Allowing an arbitrary change of origin and traversal direction, together with the use of an affine invariant parameterization, imply that **no point correspondence is required in computing our invariants.**

Now we are ready to introduce the invariant expressions themselves. Our invariants framework is very general and considers variation in an object's image induced by rigid-body motion, affine deformation, and changes in parameterization, scene illumination, and viewpoint. Each formulation can be used alone, or in conjunction with others. Due to the page limitation, we can only give a brief discussion of the invariants under rigid-body and affine transform and summarize the invariant expressions under change of illumination and viewpoint. Interested readers are referred to [1] for more details.

Invariants under Rigid-Body Motion and Affine Transform Consider a 2D curve, $\mathbf{c}(t) = [x(t), y(t)]^T$ where t denotes a parameterization which is invariant under affine transform (as described above), and its expansion onto the wavelet basis $\psi_{a,b} = \frac{1}{\sqrt{a}} g(\frac{t-b}{a})$ (where $g(t)$ is the mother wavelet) as $\mathbf{u}_{a,b} = \int \mathbf{c}(t)\psi_{a,b} dt$. If the curve is allowed a general affine transform with the possibility of being traversed from a different starting point and along an opposite direction, then the transformed curve is denoted by: $\mathbf{c}'(t) = \mathbf{m}\mathbf{c}(t') + \mathbf{t} = \mathbf{m}\mathbf{c}(\pm t + t_0) + \mathbf{t}$, where $\mathbf{m}$ is any nonsingular 2×2 matrix, $\mathbf{t}$ represents the translational motion, t_0 represents a change of the origin in traversal, and $\pm$ represents the possibility of traversing the curve either counterclockwise or clockwise:

$$\begin{aligned}
\mathbf{u}'_{a,b} &= \int \mathbf{c}'\psi_{a,b} dt \\
&= \int (\mathbf{m}\mathbf{c}(\pm t + t_0) + \mathbf{t})\psi_{a,b} dt \\
&= \mathbf{m}\int \mathbf{c}(t')\tfrac{1}{\sqrt{a}} g(\tfrac{\mp(t'-t_0)-b}{a}) dt' + \int \mathbf{t}\psi_{a,b} dt \\
&= \mathbf{m}\int \mathbf{c}(t')\tfrac{1}{\sqrt{a}} g(\tfrac{t'-(\pm b+t_0)}{a}) dt' \\
&= \mathbf{m}\int \mathbf{c}(t')\psi(t')_{a,\pm b+t_0} dt' \\
&= \mathbf{m}\mathbf{u}_{a,\pm b+t_0} \cdot
\end{aligned} \tag{1}$$

Note that we use the wavelet property $\int \psi_{a,b} dt = 0$ to simplify the second term in Eq. 1. If $\mathbf{m}$ represents a rotation (or the affine transform is a rigid-body motion of a translation plus a rotation), it is easily seen that an **invariant expression** (this is just one of many possibilities) can be derived using the ratio expression

$$\frac{|\mathbf{u}'_{a,b}|}{|\mathbf{u}'_{c,d}|} = \frac{|\mathbf{m}\mathbf{u}_{a,\pm b+t_0}|}{|\mathbf{m}\mathbf{u}_{c,\pm d+t_0}|} = \frac{|\mathbf{u}_{a,\pm b+t_0}|}{|\mathbf{u}_{c,\pm d+t_0}|}, \tag{2}$$

which is a function of the scale a and the displacement b. If we fix the scale a, by taking the same

Scenarios	Invariant expressions
Rigid-body motion (using spline basis)	$\frac{\lvert \mathbf{u}_{a,b} - \mathbf{u}_{c,d} \rvert}{\lvert \mathbf{u}_{e,f} - \mathbf{u}_{g,h} \rvert}$
Affine transform (using wavelet basis)	$\frac{\begin{vmatrix} \mathbf{u}_{a,b} & \mathbf{u}_{c,d} \end{vmatrix}}{\begin{vmatrix} \mathbf{u}_{e,f} & \mathbf{u}_{g,h} \end{vmatrix}}$
Affine transform (using spline basis)	$\frac{\begin{vmatrix} \mathbf{u}_{a,b} & \mathbf{u}_{c,d} & \mathbf{u}_{e,f} \\ 1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} \mathbf{u}_{g,h} & \mathbf{u}_{i,j} & \mathbf{u}_{k,l} \\ 1 & 1 & 1 \end{vmatrix}}$
Perspective transform (using *rational spline basis* R)	$\int_t (\mathbf{d}(t) - \sum_i \mathbf{p}_i R_{i,k}(t))^2 dt$ where $\mathbf{d}(t)$ is the observed image curve, and $\sum_i \mathbf{p}_i R_{i,k}(t)$ is the database curve in rational spline form.
Change of illumination	$\frac{\lvert [\mathbf{u}_{a_1,b_1}\, \mathbf{u}_{a_2,b_2} \cdots \mathbf{u}_{a_k,b_k}]^T [\mathbf{u}_{a_1,b_1}\, \mathbf{u}_{a_2,b_2} \cdots \mathbf{u}_{a_k,b_k}] \rvert}{\lvert [\mathbf{u}_{c_1,d_1}\, \mathbf{u}_{c_2,d_2} \cdots \mathbf{u}_{c_k,d_k}]^T [\mathbf{u}_{c_1,d_1}\, \mathbf{u}_{c_2,d_2} \cdots \mathbf{u}_{c_k,d_k}] \rvert}$

Table 1: Other invariant measures

number of sample points along each curve, we can construct a function $f_a(x)$ which we call the **invariant signature** of an object as:

$$f_a(x) = \frac{|\mathbf{u}_{a,x}|}{|\mathbf{u}_{a,x+x_0}|} \quad \text{and}$$
$$f'_a(x) = \frac{|\mathbf{u}'_{a,x}|}{|\mathbf{u}'_{a,x+x_0}|} = \frac{|\mathbf{m}\mathbf{u}_{a,\pm x+t_0}|}{|\mathbf{m}\mathbf{u}_{a,\pm(x+x_0)+t_0}|} \qquad (3)$$
$$= \frac{|\mathbf{u}_{a,\pm x+t_0}|}{|\mathbf{u}_{a,\pm(x+x_0)+t_0}|},$$

where x_0 represents a constant value separating the two indices. Then it is easily verified that when the direction of traversal is the same for both contours, $f'_a(x) = \frac{|\mathbf{u}_{a,x+t_0}|}{|\mathbf{u}_{a,x+x_0+t_0}|} = f_a(x + t_0)$. If the directions are opposite, then $f'_a(x) = \frac{|\mathbf{u}_{a,-x+t_0}|}{|\mathbf{u}_{a,-x-x_0+t_0}|} = \frac{1}{f_a(-x-x_0+t_0)}$. As the correlation coefficient of two signals is defined as

$$R_{f(x)g(x)}(\tau) = \frac{\int f(x)g(x+\tau)dx}{\|f\| \cdot \|g\|},$$

we define the **invariant measure** $I_a(f, f')$ (or the similarity measure) between two objects as

$$I_a(f, f') = max_{\tau,\tau'}\{R_{f_a(x)f'_a(x)}(\tau), R_{f_a(x)\frac{1}{f'_a(-x)}}(\tau')\}. \qquad (4)$$

It can be shown [1] that the invariant measure in Eq. 4 attains the maximum of 1 if two objects are identical, but differ in position, orientation, scale, and traversal direction and starting point. Due to the page limit, we will only summarize other invariant expressions in Table 1 without derivation. The entries shown in the table are the *invariant expressions* (similar to Eq. 2). The process of deriving *invariant signatures* (similar to Eq. 3) and *invariant measures* (similar to Eq. 4) are similar and will not be repeated here.

3 Experimental Results

In the following, we will present some preliminary results. The purpose is to provide a proof-of-concept demonstration and to discover research issues that need be addressed for a large-scale implementation and testing. Hence, the database used is of a relatively small size.

The scenario is that of a digital image database comprises a collection of sixteen airplanes in canonical (top) view (Fig. 1). The airplane contours were automatically extracted from the images and invariant shape and color signatures computed off-line. Eleven query images (Fig. 2) were photographed of these airplanes from different viewpoints and under varying illumination. The airplanes in the query images were extracted using a semi-automated process with user assistance. Even though the image database is relatively small, it contains objects of very similar appearance (e.g., models 5 and 6, and models 3, 7, and 14). Furthermore, the query images (Fig. 2) differ greatly from the database images due to large changes in perspective and illumination. This is in contrast with many digital image library retrieval schemes which can perform only between-classes (e.g., airplanes vs. cars) retrievals with small changes in imaging condition.

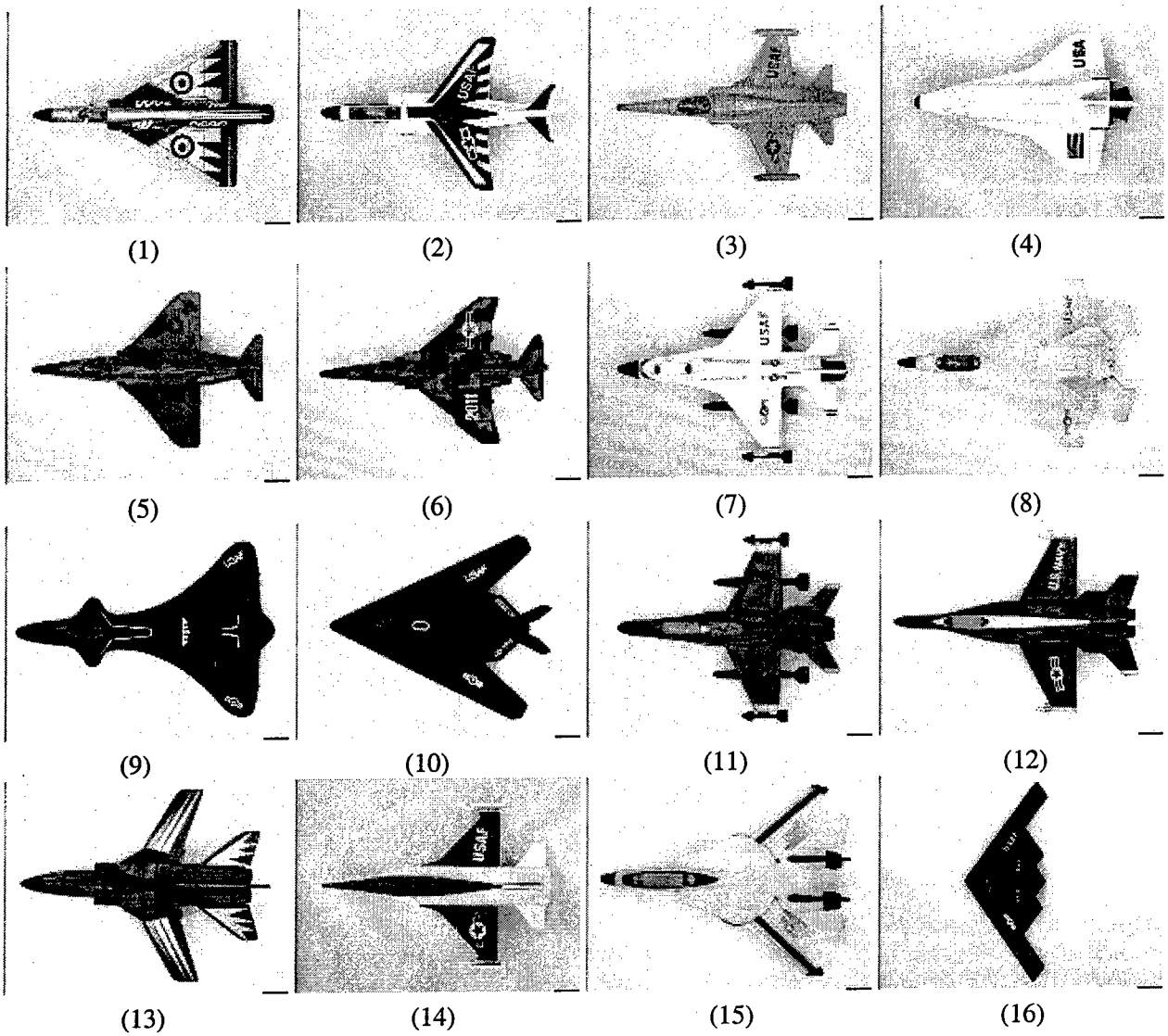


Figure 1: A database of airplane models.

We used a two-stage approach in information fusion. Features invariant to affine deformation and perspective projection were first used to match the silhouette of the query airplane with the silhouettes of those in the database. We then employed the illumination invariants computed on objects' interior to disambiguate among models with similar shape but different colors. **The results show that we were able to achieve 100% accuracy using our invariants formulation for a database comprising very similar models, presented with query images of large perspective shape distortion and change in illumination.**

Table 2 shows the performance of using affine and perspective invariants for shape matching under a large change of viewpoint. For each query image (A through K), the affine and perspective invariant signatures were computed, and compared with the signatures of all models in the database. Correlation coefficients as described in Sec. 2 were used to determine the similarity between each pair of signatures. Each row in Table 2 refers to a query image. Each of the ten columns represents the rank given to each airplane model from the database (shown in parentheses). The columns are ordered from left to right, with the leftmost column being the best match found. Only the top ten matches are shown. The values (not in parentheses) are the correlation coefficients. Entries printed in boldface are the expected (correct) matches.

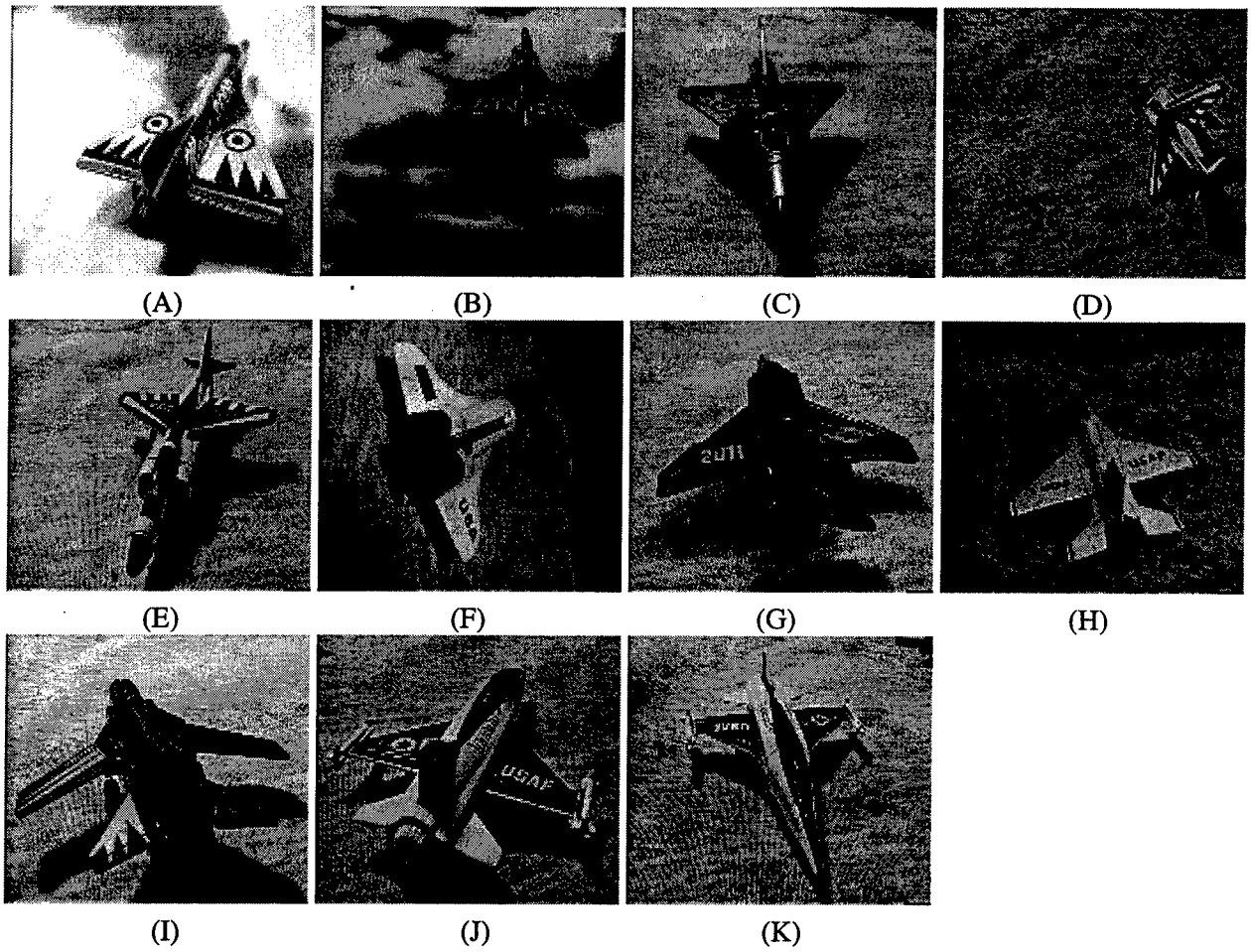


Figure 2: The same airplanes in varying poses and illumination.

As can be seen, all query images were identified correctly. Fig. 3 shows a sample result. The leftmost image is the query image. The top three matches are in the next three columns—the query image (solid) and estimated image (dashed, using perspective invariants) with the corresponding database model are shown.

For this experiment, *all* query images were correctly matched with the models from the database, using affine and perspective invariants. However, the error values of the top two matches for, say, airplane K were very close to each other. This is because the top two matches have similar shapes and both are similar to the query image. The confidence in the selected matches can be strengthened by testing whether the interior regions of the objects are also consistent. Illumination invariants readily applies here.

For illumination invariants, a characteristic curve was uniquely defined on the surface of each airplane model in the database (performed off-line), so that its superimposition over the model emphasizes important (or interesting) color patterns in the image. Our perspective invariants scheme computed the transformation parameters that best match the two given contours. The same parameters were used to transform the characteristic curve defined for each model to its assumed pose in the query image. Hence, the colors defined by the characteristic curve in the model should match the colors defined by the transformed curve in the query image (except for changes due to illumination). Illumination invariant signatures for the query images were then computed, and compared with the signatures stored in the database using Eq. 4.

We show one result of illumination invariants where the (perspective invariant) errors of the 1^{st}

	Rank (using affine and perspective invariants)									
Image	1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}	6^{th}	7^{th}	8^{th}	9^{th}	10^{th}
A	**(1) 0.8792**	(9) 0.7210	(4) 0.6161	(6) 0.4967	(5) 0.4663	(10) 0.4578	(2) 0.4030	(7) 0.3248	(11) 0.2443	(14) 0.2388
B	**(1) 0.9527**	(9) 0.8532	(10) 0.7666	(4) 0.7479	(6) 0.6630	(2) 0.6103	(5) 0.5943	(15) 0.5364	(16) 0.4756	(7) 0.4576
C	**(1) 0.8538**	(4) 0.6806	(2) 0.6521	(9) 0.6016	(6) 0.5623	(5) 0.5353	(10) 0.4446	(14) 0.3359	(7) 0.3095	(11) 0.2386
D	**(2) 0.9283**	(6) 0.9002	(5) 0.8962	(4) 0.8177	(13) 0.8097	(14) 0.7801	(1) 0.7730	(7) 0.7663	(3) 0.7502	(12) 0.7439
E	**(2) 0.9228**	(5) 0.7747	(6) 0.7622	(14) 0.6975	(12) 0.6167	(4) 0.6167	(3) 0.6146	(13) 0.5902	(7) 0.5704	(15) 0.4813
F	**(4) 0.6369**	(1) 0.6002	(9) 0.5810	(6) 0.5291	(10) 0.5205	(14) 0.5056	(5) 0.4486	(11) 0.4283	(2) 0.4036	(7) 0.3946
G	**(6) 0.8254**	(13) 0.7293	(5) 0.7026	(4) 0.6616	(2) 0.6460	(14) 0.6396	(12) 0.6287	(3) 0.6035	(1) 0.5930	(7) 0.5638
H	**(7) 0.8747**	(14) 0.8552	(3) 0.8398	(11) 0.8226	(13) 0.7848	(6) 0.7668	(12) 0.7663	(5) 0.7282	(2) 0.7007	(4) 0.6980
I	**(13) 0.8609**	(6) 0.6890	(3) 0.6563	(14) 0.6468	(12) 0.6343	(5) 0.6107	(7) 0.5916	(2) 0.5849	(15) 0.5775	(1) 0.5516
J	**(14) 0.8815**	(3) 0.8017	(12) 0.7564	(13) 0.7512	(7) 0.7055	(11) 0.6805	(6) 0.6501	(4) 0.6346	(5) 0.5838	(15) 0.5711
K	**(14) 0.8779**	(3) 0.8558	(7) 0.7623	(13) 0.7272	(12) 0.7270	(6) 0.7235	(11) 0.7209	(2) 0.6503	(5) 0.6191	(4) 0.5459

Table 2: Top ten matches between each query image and database models, using affine and perspective invariants. Numbers in parentheses indicate the airplane model selected. The value beneath it is the similarity measure between the selected image and query image. The correct airplane model is in boldface. Each row corresponds to a query image. The columns are arranged left to right, from the best match to worse.

and 2^{nd} best matches differ by a small amount (see Table 2); in this case, query image K. Figs. 4 (a) and (d) show the characteristic curves (the zigzag lines) superimposed over the images of models 14 and 3. The transformed characteristic curves, shown in (b) and (e), is superimposed over the query image K, using parameters estimated from perspective invariants. Finally, (c) and (f) show the illumination invariant signatures. Clearly, the signatures in (c) is much more consistent, which reinforces the results from shape invariants.

4 The Concluding Remarks

We present a technique where shape/color information from interior/contour points is used to describe an imaged object for database retrieval. The technique is superior in that it tolerates changes in appearance induced by incidental environmental factors and is powerful enough for within-a-class retrieval.

References

[1] Ronald Alferez and Y. F. Wang. Geometric and Illumination Invariants for Object Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1999. to appear as a regular paper.

[2] H. W. Guggenheimer. *Differential Geometry*. McGraw-Hill, New York, 1963.

[3] J. Mundy and A. Zisserman (eds.). *Geometric Invariance in Computer Vision*. MIT Press, Cambridge, MA, 1992.

[4] T. H. Reiss. *Recognizing Planar Objects Using Invariant Image Features*. Springer-Verlag, Berlin, 1993.

[5] I. Weiss. Geometric Invariants and Object Recognition. *Int. J. Comput. Vision*, 10(3):207–231, 1993.

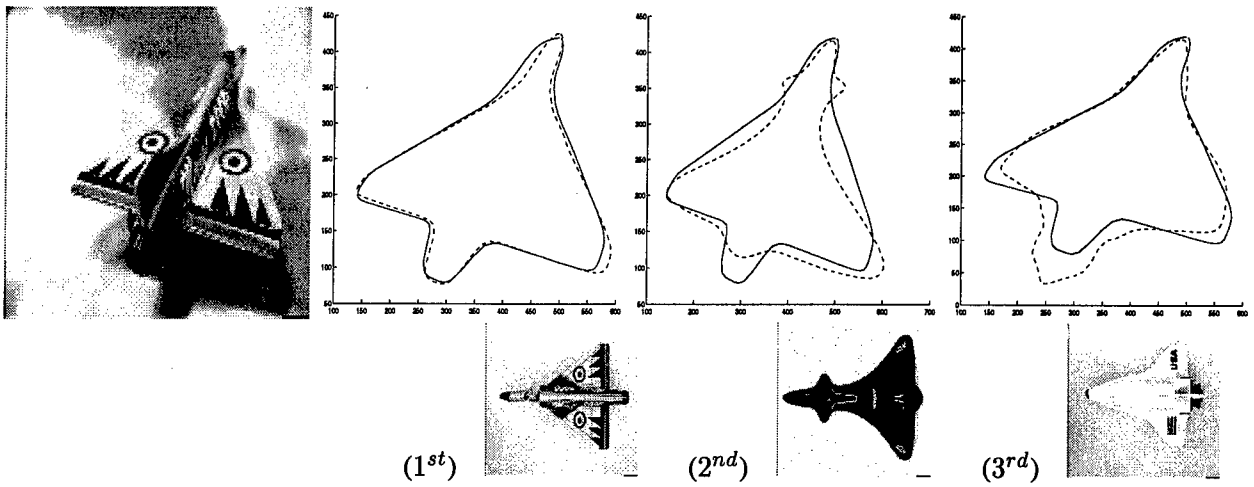


Figure 3: Query image A, with the top three matches from the database, using perspective invariants. (Solid for the query image, dashed for the estimated image using perspective invariants.)

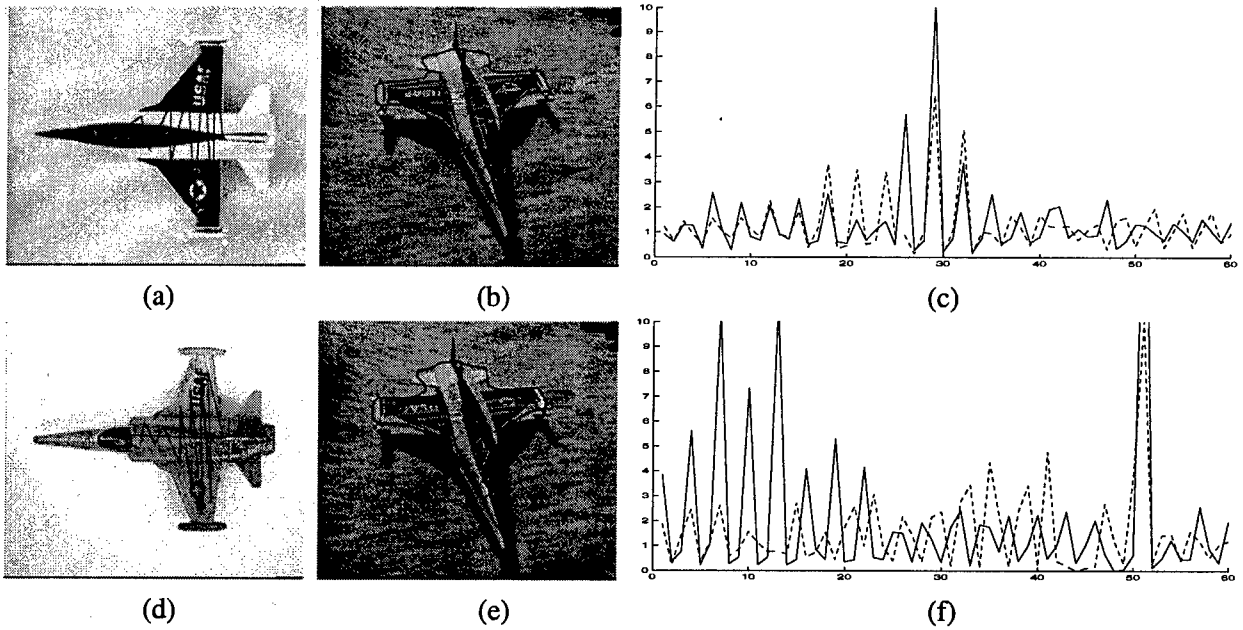


Figure 4: (a),(d) Airplane models with the characteristic curves superimposed, (b),(e) query image with the transformed characteristic curves superimposed, and (c),(f) illumination invariant signatures for query image K (solid) and for models 14 and 3 (dashed).

Fusion of multiple cues for video segmentation

Bikash Sabata and Moises Goldszmidt
SRI International
333 Ravenswood Ave
Menlo Park, CA 94025, U.S.A.
{sabata, moises}@erg.sri.com

Abstract *The segmentation of video into contiguous scenes is becoming an important problem in many applications. Since video data is a rich source of spatio-temporal information, different types of features can be computed in the video data. Each of these features provide a cue for the segmentation of video and is usually sufficient to perform an approximate segmentation of video. However, the features many times provide conflicting evidence for segmentation. Further, since there is strong correlation between the different features, it is not easy to fuse the information from the features to make segmentation decisions. We present a method based on Bayesian Networks that model the dependence between the segmentation decision and the different features. This framework using Bayesian Networks is promising and provides an extensible mechanism for fusion of information.*

Keywords: Video Processing, Bayesian Networks

1 Introduction

The information in video data is being used increasingly for many decision and interpretation tasks. For example, we would like to determine when one scene ended and a new one started so that the relevant segments of the video may be retrieved for display. There is a critical need for efficient management and processing of video data. However, the sheer volume of information in the video data makes it difficult to device algorithms that are efficient and robust. The decisions and interpretations based on video use the feature vectors extracted from the video data. Each of these features provide a cue for the understanding of video and is usually sufficient to make an approximate interpretation about the content of video. However, the features many times provide conflicting evidence. Further, since there is strong correlation between the different features, it is not easy to fuse the information from the features to make interpretations.

This problem of correlated and conflicting evidence from different sources is a common occurrence in multisensor systems and other complex systems. Our goal is to develop a framework to address the problems of information fusion when the features are noisy and highly correlated. The source of these features may arise out of processing of different sensors and/or different filters applied to sensor data such as video. In this paper we address the specific problem of segmenting structured video that arises in broadcast video and movies. However, the techniques for fusion we develop are more general.

We present a method based on Bayesian Networks that model the dependence between the segmentation decision and the different features. The Bayesian network model also explicitly represents the correlations between the different features. These correlations may be known a priori (because of the domain knowledge) or may be learned from the data. We incorporate the prior knowledge into the model and learn the other dependency structures by learning the Bayesian networks from the data.

In section 2 we present a brief overview of the diverse set of techniques used to segment video data. The important lesson here is that

many of techniques that segment video perform reasonably well for restricted classes of videos. In different situations different techniques perform better than others. An important consequence of our framework is that we are able to select the best set of techniques that are sufficient to make reliable and robust decisions for a given class of video data.

In section 3 we present the specific set of features we compute in the video to make the segmentation decision. A common problem we have seen in the past approaches to computing the feature vectors in video is that the feature vectors are computed in individual frames and the temporal dimension is added in as a difference operation. We depart from this approach and present a novel multi-scale filter to compute the color and texture features for each individual frame and their variations along the time axis. In addition, we compute features that are unique to video. These include the spatio-temporal motion tracks of points in video and edge features in the spatio-temporal volumes.

The Bayesian Network based framework for fusion of information from the different features is presented in section 4. In sections 5 and 6 we present the results of our experiments and conclude with some discussions about the current and future work.

2 Background

The majority of the techniques for video segmentation use low-level image features such as pixel differences, differences in the statistical properties of the feature values, histogram comparisons, edge differences, and motion vectors. The key problem is that there are many events in the video that have the same characteristics as scene changes in the low level features. For example, fast camera panning in the scene may have the same color histogram characteristics as a dissolve. Reducing the number of false positive triggers is the main objective of research activities in this area.

A large class of segmentation algorithms compute the boundary between two segments by examining the local pixel values and their statistical properties in frames in a temporal window 2ϵ around the candidate boundary frame b. Zhang, Kankanhalli and Smoliar [1] compute the number of pixels that change value more than a threshold to decide if a boundary has been detected at frame i. Yeo [2] improves the above technique by taking the difference on spatially reduced frames over a symmetric temporal window $[b-\epsilon, b+\epsilon]$ around the candidate frame b. Yeo also detects the gradual change regions by detecting the "plateaus" in the distance measure over temporal windows. Kasturi and Jain [3] segment each frame into regions and compare statistical measures of the pixels in the regions over the frames. To improve the computational efficiency, Taniguchi, Akutsu and Tonomura [4], take temporal samples to process the frames and incrementally increase the sampling where a candidate scene change is detected.

Another approach to boundary detection is to model the transition in terms of the statistics of the pixel values of the frames in the temporal window constituting the transition. Aigrain and Joly [5] and Hampapur, Jain, and Weymouth [6] use such model based methods to capture the different shot transitions.

An alternative class of approaches that use the pixel information more compactly is the histogram of the frames. The histograms could be the intensity histograms or the color histograms. Histograms, of course lose the spatial information of the frames entirely, and therefore are robust with respect to camera motions and reasonable amounts of noise.

Zhang, Kankanhalli and Smoliar [1], and Yeo [2] compare the histograms by using a bin-wise difference of the histograms of the two consecutive frames. Ueda, Miyatake and Yoshizawa [7] use the rate of change of the color histograms to find the shot boundaries. Nagasaka and Tanaka [8] compare many different statistical measures for the histogram distributions. They report that partitioning the frames into 16 regions and using a χ^2 test on color histograms of the regions over the frames performs the best for shot boundary detection.

Their method is robust against zooming and panning but fails to detect gradual changes. Prior statistical properties of the video are incorporated into the decision process by Swanberg, Shu and Jain [9]. They use intensity histogram differences in regions weighted with the likelihood of the region changing in the video.

In addition to the features in the individual frames, the video data also has motion related features. Zabih, Miller and Mai [10] use a method based on edge tracking over frames to determine shot boundaries. The edge distances are measured using the Hausdorff distance measure. Shahraray [11] uses a method similar to the motion vector computation algorithms in most MPEG codecs to compute scene change points. Hsu and Harashima [12] model the scene changes and activities as motion discontinuities. The motion is characterized by considering the sign of the Gaussian and Mean curvature of the spatio-temporal surfaces. Clustering and split-and-merge approaches are then used to segment the video.

In summary the past literature has a large collection of techniques that extract one of the many features from video to detect the segment boundaries. Beyond simple ad-hoc schemes that try to integrate the information from the different features using some kind of weighting procedure, we have not seen any effort to use sophisticated techniques to fuse the evidence from the different cues.

3 Video Features

The lesson learnt from past work and our experiments with the various techniques in the literature is that although the mechanisms don't perform well for all situations, they perform well in a subset of the situations. Our approach is to select a set of these methods and fuse the output of each module to make the final decision about segmentation. In the process, we modified some of these algorithms to make them more robust and efficient. In addition, we propose some new features from video that capture information in video that has not been addressed in the past approaches.

There are four basic types of features we compute and use in the fusion module for making the final segmentation. The first set of features is based on the color distributions in each frame of the video. The second feature set is based on the response of each frame to a set of texture filters. The third feature set scores the frames for a segmentation boundary based on the tracking of significant point features in the video. Finally the fourth feature class computes the likelihood of a segmentation boundary by detecting edges in the spatio-temporal volume that represents the video.

The segmentation is computed by comparing the change across the candidate segment boundary. The change can be measured as a distance $\mathcal{D} = F(S_{b-\epsilon}, S_{b+\epsilon})$ between the video features, $S_{b-\epsilon}$ and $S_{b+\epsilon}$, in the two temporal intervals $[b-\epsilon, b]$ and $[b, b+\epsilon]$ around the boundary b. Every procedure for segmentation that has been discussed in the past literature can be mapped to this formulation. The distance measure is then compared with a threshold value to determine if there exists a boundary at b. The different approaches select different properties $S_{b-\epsilon}$ and $S_{b+\epsilon}$ to represent the intervals and the function that evaluates the distance. In section 4 we will present fusion techniques that either use the $\mathcal{D}$ from each module for making the fused decision or the individual decisions from each module to make the fused decision.

3.1 Color Features

The most common feature used to segment video are the color histograms of each frame. Using the results from Nagasaka and Tanaka [8], we designed a distance measure based on the χ^2 measure between two histograms. We define the histogram of the video frames over the temporal intervals $[b-\epsilon, b]$ and $[b, b+\epsilon]$ around the candidate frame b. Further, we weigh the contributions of the frames to the joint histogram using a Gaussian Mask. This method of computing the color feature vector for video is novel and inspired from the scale space methods in computer vision and the multi-scale filters for edge detection. The

multi-scale Gaussian windowing approach provides a robust mechanism to reliablely estimate the scene boundaries in the presence of noise.

Let $h_t(i)$ be the normalized histogram[1] of the t^{th} frame in the video. The weighted histogram for a temporal window $W = [t_s, t_e]$ is then computed as

$$h_{[t_s,t_e]}(i) = \sum_{t \in [t_s,t_e]} w_t \cdot h_t(i)$$

where w_t is the weight associated with the t^{th} frame in the window. The weights are computed using the one dimensional Gaussian function $G(x) = e^{-\frac{x^2}{2\pi\sigma^2}}$ multiplied with a normalization factor. The Gaussian window size is different for the different scales. The distance measure for the Gaussian windowed histograms for the scale s is given by

$$\mathcal{D}^{color}(s,t) = \sum_i \frac{(h_{[l]}(i) - h_{[r]}(i))^2}{(h_{[l]}(i) + h_{[r]}(i))} \quad (1)$$

where

$$[l] = [t - \epsilon(s),\ t];\ \ [r] = [t,\ t + \epsilon(s)].$$

The set of distance measures for the different scales and colors together form the set of features that capture the relevant color information in video. Figure 1 shows the distance measure at three different scales as a function of time for an example video.

Yeo [2] proposed a flash detector by noticing that flashes produce two closely spaced sharp peaks in the χ^2 distance scores of the video frames. We implemented the above detector to give the candidate frames at which flashes were detected.

3.2 Texture Features

The texture information in each frame can also be used to evaluate the continuity of a segment. The use of texture information for segmentation in video is a novel use of texture filters. We propose a novel distance measure that

[1]Normalization makes $\sum_i h_t(i) = 1$ and $h_t(i)$ can be interpreted as the probability of a pixel taking value i in frame t.

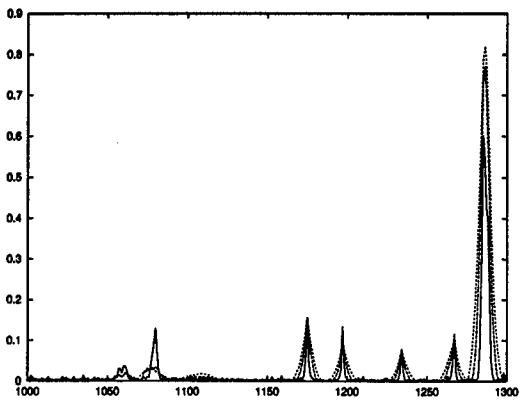


Figure 1: Distance score at each frame for three different scales. At the noise, fast camera motion, and flashes the maxima reduce with increasing scale, while at true boundaries, the maxima remain the same or increase.

uses the texture energy to compute the distance between temporal windows across candidate segment boundaries. The texture energy is computed using the *Gabor filters* proposed in [13]. The Gabor energy method measures the similarity between neighborhoods in an image and Gabor masks. Each Gabor mask consists of Gaussian windowed sinusoidal waveforms with parameters of wavelength λ, orientation θ, phase shift ϕ, and the standard deviation σ. The filter is given by:

$$G(x,y) = e^{-\frac{(x-X)^2+(y-Y)^2}{2\sigma^2}} \times \sin\left(\frac{2\pi(x\cos\theta - y\sin\theta)}{\lambda} + \phi\right)$$

A set of filters is generated by varying the θ and λ. The texture energy for a filter (fixed θ and λ) is calculated as the sum over the phases of the squared convolution values. We implement a total of twelve filters by quantizing the θ into four values and λ into three values.

Next, similar to the case of the color distance measure, we use the texture energy response of each frame to find the difference between the adjacent group of frames. To compute the distance measure at frame i, a Gaussian window at scale s is selected around the frame. The weighted average texture energy is calculated

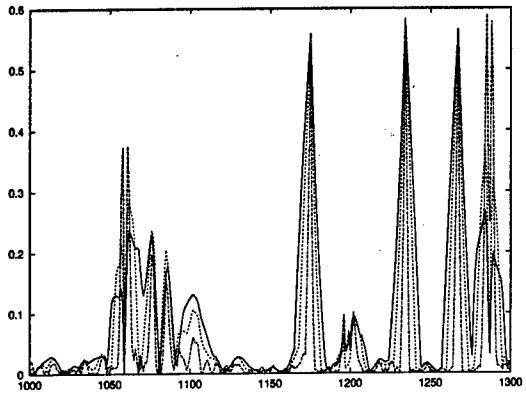


Figure 2: Distance score at each frame for three different scales for one texture filter.

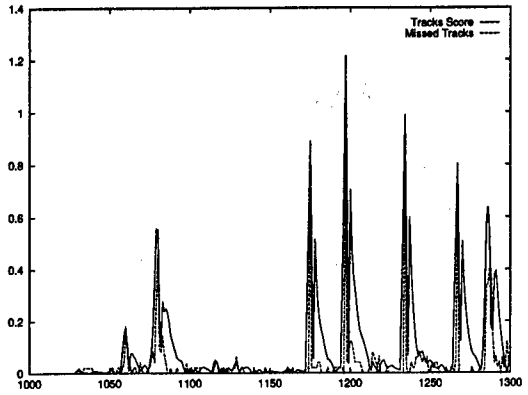


Figure 3: Tracking based distance score and the percentage missed features at each frame.

to the left and right of the frame. The normalized distance between the average texture energy is used as the estimate of the change across the segment boundary.

Figure 2 shows the distance measure computed at the different scales for one of the texture filters on an example video.

3.3 Tracking Based Features

Motion features are important features as they correspond to the real physical phenomenon captured by the video sensor. Using tracks of objects and points in the video for detecting scene changes was used in the past by Zabih, Miller and Mai [10] who tracked edge segments over frames and used a Hausdorff distance measure to evaluate the segment boundaries.

A critical problem in tracking based systems is that of feature selection. We use the method proposed by Shi and Tomasi [14] to detect good point features to track. The features are selected from the frames and tracked across the video. We assign a score to the tracking of the features and use it to evaluate the segment boundaries. The score is computed by weighting the contribution of each feature that is tracked from the last frame to the current frame. The weighting function looks at the history of the feature and assigns a weight that is proportional to the history of the track, however, the incremental increase in the weight for the feature decreases exponentially with the history length. The weight for the i^{th} feature is computed as

$$w_i = 1 - e^{\frac{-p_i}{k}} \quad (2)$$

where p_i is the number of frames in the past through which the i^{th} feature was tracked and k is some constant that determines the sensitivity to the history of the tracks. In addition to evaluating the feature tracks in a frame, we also measure the fraction of features that cannot be tracked in each frame from the past frame. This fraction also gives the evidence about the boundary between video segments.

To compute a distance measure across the segment boundary, we compute the difference between the average track scores in windows on either side of the candidate boundary frame. At frame i, a window of size S is selected around the frame and the average track score is calculated to the left and right of the frame. The difference between the average track scores and the fraction of the missed tracks is used as the distance measure of the tracking module. Figure 3 shows the tracking based distance scores and the fraction of the point features missed in the tracking for each frame.

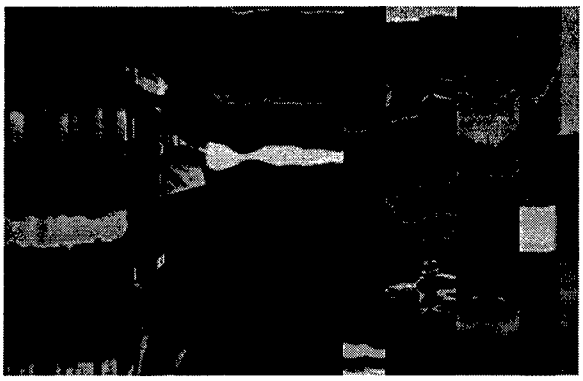


Figure 4: $x-t$ *Section through the spatio-temporal video volume. The* t *axis is along the horizontal direction.*

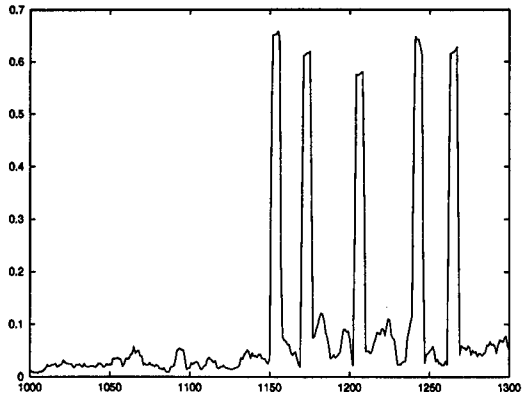


Figure 5: *Distance score based on the edges in spatio-temporal volumes at each frame.*

3.4 Edges In Spatio-Temporal Volumes

Video data is three dimensional data where the temporal dimension is the third dimension. To detect the segmentation boundaries we should study the patterns in the data along the temporal dimensions. This idea was investigated by Otsuji and Tonomura [15] who proposed a projection detection filter to detect cuts in video. They projected the video data along the $x-t$ and $y-t$ planes to generate images that they then use to detect cuts. This construction is based on the work on spatio-temporal surfaces by Baker and Bolles [16].

We use a similar idea and generate sections through the video volume using planes parallel to the $x-t$ and the $y-t$ planes (Figure 4). The edges perpendicular to the t axis in these sections indicate possible video segment boundaries. The fraction of the pixels at any t covered by the horizontal edges is taken as a measure of the segment boundary. Averaging this measure across many sections gives a probability measure of the existence of a segment boundary based on the evidence from edges in spatio-temporal volumes. Figure 5 shows the probability measure evaluated for the different frames in an example video.

4 Bayesian Network Based Fusion

We use capital letters X, Y, Z for variable names, and lower-case letters x, y, z to denote specific values taken by those variables. Sets of variables are denoted by boldface capital letters $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ and assignments of values to the variables in these sets are denoted by boldface lowercase letters $\mathbf{x}, \mathbf{y}, \mathbf{z}$.

A *Bayesian network* over a set of variables $\mathbf{X} = \{X_1, \ldots, X_n\}$ is an annotated directed acyclic graph that encodes a joint probability distribution over $\mathbf{X}$. Formally, a Bayesian network is a pair $B = \langle G, \mathrm{L} \rangle$. The first component, G, is a directed acyclic graph whose vertices correspond to the random variables $X_1, \ldots, X_n$, and whose edges represent direct dependencies between the variables. The second component of the pair, namely L, represents a set of local *conditional probability distributions* (CPDs) $L_1, \ldots, L_n$, where the CPD for X_i maps possible values x_i of X_i and $\mathbf{pa}(i)$ of $\mathbf{pa}(i)$, the set of parents of X_i in G, to the conditional probability (density) of x_i given $\mathbf{pa}(i)$. A Bayesian network B defines a unique joint probability distribution (density) over $\mathbf{X}$ given by the product

$$P_B(X_1, \ldots, X_n) = \prod_{i=1}^{n} L_i(X_i | \mathbf{pa}(i)) \ . \quad (3)$$

When the variables in $\mathbf{X}$ take values from finite discrete sets, we typically represent CPDs as tables that contain parameters $\theta_{x_i|\mathbf{pa}(i)}$ for all possible values of X_i and $\mathbf{pa}(i)$. When the variables are continuous, we can use various parametric and semi-parametric representations for these CPDs.

In this paper, we treat information fusion as a pattern classification problem. We assume that there is one variable A_i for each feature, and a distinguished variable *Outcome* that can take value from the set $\{0, 1, 2\}$ depending on whether the frame is "normal," a "boundary, or a "flash." The objective is given a set of vectors $\mathbf{X} = \{A_1, \ldots, A_n, Outcome\}$, to induce a probability distribution $Pr(A_1, \ldots, A_n, Outcome)$ from this data in the form of a Bayesian network. Given this network the decision on a new scene will be given by:

$$\underset{O}{\text{argmax}} \quad Pr(Outcome = O | a_1, \ldots, a_n), \tag{4}$$

which is the classic definition of a Bayesian classifier [17]. Note that we have translated the fusion problem to that of inducing a probability distribution linking the various features with a decision on the nature of the frame.

There is a recent substantial body of work on inducing Bayesian networks from data (see [18] for example, and references therein). In [19] Friedman et al argue convincingly for using specialized graph structures for classification tasks. As an example, consider a graph structure where the *Outcome* variable is the root, that is, $\mathbf{Pa}(Outcome) = \emptyset$, and each feature has the *Outcome* variable as its unique parent, namely, $\mathbf{pa}(A_i) = \{Outcome\}$ for all $1 \leq i \leq n$. For this type of graph structure, Equation 3 yields $\Pr(A_1, \ldots, A_n, Outcome) = \Pr(Outcome) \cdot \prod_{i=1}^{n} \Pr(A_i | Outcome)$. From the definition of conditional probability, we get $\Pr(Outcome | A_1, \ldots, A_n) = \alpha \cdot \Pr(Outcome) \cdot \prod_{i=1}^{n} \Pr(A_i | Outcome)$, where α is a normalization constant. This is the definition of the *naive Bayesian classifier* commonly found in the literature [17].

The naive Bayesian classifier has been used extensively for classification. It has the attractive properties of being robust and easy to learn—we only need to estimate the CPDs $\Pr(Outcome)$ and $\Pr(A_i \mid Outcome)$ for all attributes. Nonetheless, the naive Bayesian classifier embodies the strong independence assumption that, given the value of *Outcome*, features are independent of each other. Friedman, Geiger and Goldszmidt [19] suggest the removal of these independence assumptions by considering a richer class of networks. They define the TAN Bayesian classifier that learns a network in which each attribute has the class and at most one other attribute as parents. Thus, the dependence among attributes in a TAN network will be represented via a tree structure. Figure 6 shows an example of a TAN network.

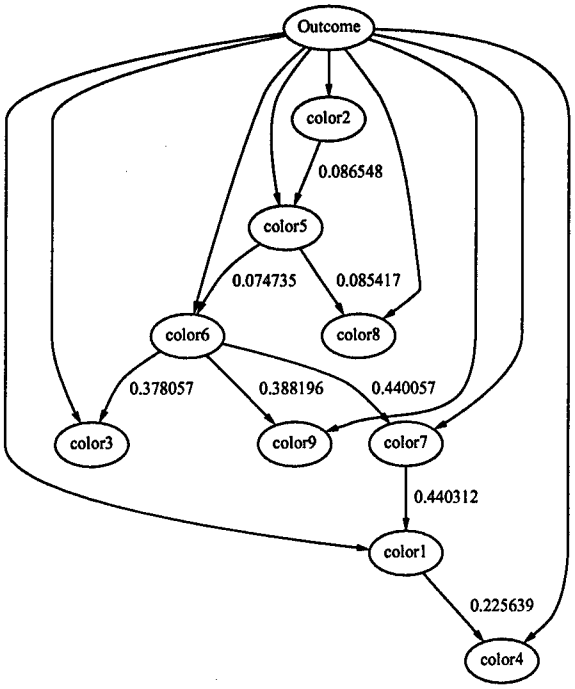


Figure 6: A TAN model learned using only features that take color into account. The numbers on the arcs indicate conditional mutual information between the features.

In a TAN network, an edge from A_i to A_j implies that the influence of A_i on the assessment of *Outcome* also depends on the value of A_j. For example, in Figure 6, the influence of the feature "color1" on *Outcome* depends on the value of "color7," while in the

naive Bayesian classifier the influence of each feature on *Outcome* is independent of other features. These edges affect the classification process in that a value of "color1" that is typically surprising (i.e. $P(color1|Outcome)$ is low) may be unsurprising if the value of its correlated attribute, "color7," is also unlikely (i.e. $P(color1|Outcome, color7)$ is high). In this situation, the naive Bayesian classifier will overpenalize the probability of the class by considering two unlikely observations, while the TAN network of Figure 6 will not do so, and thus will achieve better accuracy.

TAN networks have the attractive property of being learnable in polynomial time [19]. In the next section we show the results of using the TAN classifier for fusing the information provided by the various filters. As a control, we also used the naive Bayes classifier introduced above. As is illustrated in the next section, the lack of correlation modeling between the different features causes a substantial increase in the number of false positives (e.g. classifying normal frame are boundaries).

5 Results

Several experiments were conducted with the Bayesian network model induction and segmentation. However, for the sake of brevity, in this paper we will present only an outline of our experiments and indicate some of the results. This work is still in progress and the next section presents the future directions.

The video segmentation experiments were performed on samples of broadcast news video. The video segments were processed using the different color filters, texture filters, tracking algorithm, and the spatio-temporal edge detectors. We also ran the "flash detector' on all the data. In all there were 51 features: 9 color features, 3 from flash detector output, 36 from the texture filters, 2 from the tracker and 1 from the spatio-temporal edge detector. In video data the fraction of the frames that are segment boundaries and flashes are extremely small. This is because video has 30 frames per second and scene changes do not typically occur more than once in 4-5 seconds. This is a tough problem as only approximately 1% of the data is of type breaks and flashes. We are not interested in accuracy (the percentage of successfully classified frames) as the vast majority of the frames are normal (approximately 99%). Our criteria must be based on the number of *false negatives*, how many segment boundaries or flashes, were missed, and the false positives, how many normal frames were confused by our model for segment boundaries or flashes.

The first Bayesian network model was generated with the data discretized following the method by Fayyad and Irani [20], using the routines in the MLC++ package [21]. We first trained on the whole data set and tested classification on the same dataset. We run the risk of over-fitting the data, but given the nature of the problem (so few instances) we wanted a sanity check. The results where very encouraging. Only 3 of the 36 events we were interested in where missed; namely only 3 false negatives, and 27 were false positives. The same experiment with the naive returned 0 false negatives (i.e., not a single segment boundary or a flash was missed), however the number of false positives jumped to 214 (57 of the normal frames were labeled as boundaries and 157 were labeled as scenes with flash).

We performed a five fold test to check how would the model behave against unseen data. The folds maintained the proportions of the interesting cases in the training data, but naturally it reduced the number of instances. The results show that about 1 in every 6.4 segment boundaries are missed, and that about 1 in every 50.3 normal frames were considered to be boundaries or flashes.

To test whether the model was indeed fusing information we tested the performance of the four feature classes in isolation. The results reveal that indeed fusion took place. The number of false negatives decreased significantly for the filters based on "color," "track," and "flash," and for "texture" even though the number of false negatives for segment boundaries increased by 2 the number of false negatives for flash decreased by the same amount.

It is worth noting that in the case of "texture" the number of false positives decreased by almost 40%.

We attempted to induce a classifier without discretization. We used Gaussians and linear Gaussians as the family of distributions. The results where poor and the model failed to identify the majority of the scenes of interest (25 false negatives). This was a surprising result for us are we are trying to characterize this further at this time. As described in the next section, future work includes the exploration of more sophisticated models, such as those described in [22].

6 Conclusions & Future Work

Fusion of information from multiple sensors or from different computational modules is becoming important in many applications. In particular for multimedia applications (with audio and video) the fusion of cues from the different media channels and from different processing modules is becoming increasingly critical. For example, in the domain of multimedia information processing, applications requiring content based search and retrieval require interpretation of features in the data from all the media sources. In this paper we presented a framework based on Bayesian networks for the fusion of information from multiple sources. This framework is very general and extensible. The preliminary results on our fusion experiments are very encouraging. Currently we are applying this framework to the integration of multiple cues resulting from the processing of audio, video/imagery, speech and text in broadcast news video.

In addition to the fusion framework we also presented novel features to evaluate the content of video. These included the multi-scale color and texture filters and the edges in the spatio-temporal volume of data representing video. The usual approach to feature detection in video is repeated application of the image feature detectors to every frame of the video. Our approach was to design feature detectors specific to video data. This approach we believe is the key to characterizing the structure of video data and extracting features that are relevant to the content of video.

In our experiments with the Bayesian network models we were able to design different networks that performed better on one of the performance metrics at the expense of others. We are currently exploring methods to characterize the different models and quantify the tradeoffs. We are also exploring the use of more sophisticated models such as those in [22] that include mixtures of Gaussians and also a mix of discrete and continuous features. A significant step will be to use models that do not consider the data to be iid but sequences in time.

The Bayesian network model also allows us to evaluate the contribution of the different features towards the final decision using the value of information computations in Bayesian networks. This results in minimal set of features necessary to reliably segment video. In addition, we will generate a decision tree where the output of one feature detector directs the test with the next feature detector. Feature computations tend to be computationally expensive therefore our goal is to provide a decision procedure to determine the redundant computations and the most significant computations.

Acknowledgements
We would like to thank Claire Monteleoni for her help with the experiments.

References

[1] H. Zhang, A. Kankanhalli, and S. Smoliar, "Automatic partitioning of full motion video," *Multimedia Systems*, 1(1):10–28, 1993.

[2] B.-L. Yeo, *Efficient Processing of Compressed Images and Video.* PhD thesis, Princeton University, 1996.

[3] R. Kasturi and R. Jain, "Dynamic vision," in *Computer Vision: Principles* (R. Kasturi and R. Jain, eds.), vol. 1, 1991.

[4] Y. Taniguchi, A. Akutsu, and Y. Tonomura, "Panoramaexcerpts: Extracting

and packing panoramas for video browsing," in *Proc. of the ACM Multimedia Conference*, Seattle, Nov. 1997.

[5] P. Aigrain and P. Joly, "The automatic real-time analysis of file editing and transition effects and its applications," *Computer and Graphics*, vol. 18, pp. 93–103, Jan. 1994.

[6] A. Hampapur, R. Jain, and T. Weymouth, "Digital video segmentation," in *Proc. of ACM Multimedia*, San Francisco, pp. 357–364, Oct 1994.

[7] H. Ueda, T. Miyatake, and S. Yoshizawa, "Impact: An interactive natural motion picture dedicated multimedia authoring system," in *Proceedings of CHI*, New Orleans, pp. 343–350, April 1991.

[8] A. Nagasaka and Y. Tanaka, "Automatic video indexing and full video search for object appearences," in *Visual Database Systems II* (E. Knuth and L. Wegner, eds.), pp. 113–127, Elsevier Science Publishers, 1992.

[9] D. Swanberg, C. Shu, and R. Jain, "Knowledge guided parsing and retrieval in video databases," in *Storage and Retrieval for Image and Video Databases*, vol. 1908 of *Proc of SPIE*, pp. 173–187, Feb. 1993.

[10] R. Zabih, J. Miller, and K. Mai, "A feature-based algorithm for detecting and classifying scene breaks," in *Proc. of ACM Multimedia 95*, (San Francisco), pp. 189–200, Nov. 1995.

[11] B. Shahraray, "Scene change detection and content-based sampling of video sequences," in *Digital Video Compression: Algorithms and Technologies*, vol. Proc. SPIE 2419, pp. 2–13, February 1995.

[12] P. Hsu and H. Harashima, "Detecting scene changes and activities in video databases," in *Proc. of ICASSP 94*, pp. 13–36, April 1994.

[13] I. Fogel and D. Sagi, "Gabor filters as texture discriminator," *Journal of Biological Cybernetics*, vol. 61, pp. 103–113, 1989.

[14] J. Shi and C. Tomasi, "Good features to track," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Seattle, June 1994.

[15] K. Otsuji and Y. Tonomura, "Projection detecting filter for video cut detection," in *Proc. of ACM Multimedia 93*, (Anahiem), pp. 251–257, ACM, August 1993.

[16] H. Baker and R. C. Bolles, "Generalizing epipolar–plane image analysis on the spatiotemporal surface," *International Journal of Computer Vision*, 3:33–49, 1989.

[17] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis.* New York: John Wiley & Sons, 1973.

[18] D. Heckerman, D. Geiger, and D. M. Chickering, "Learning Bayesian networks: The combination of knowledge and statistical data," *Machine Learning*, 20:197–243, 1995.

[19] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," *Machine Learning*, 29:131–163, 1997.

[20] U. Fayyad and K. Irani, "Multi-interval discretization of continuous-valued attributes for classification learning," in *Proc. of Intl. Joint Conf. on AI*, pp. 1022–1027, 1993.

[21] R. Kohavi, G. John, R. Long, D. Manley, and K. Pfleger, "MLC++: A machine learning library in C++," in *Proc. 6^{t}h Intl. Conf. on Tools with Artificial Intelligence*, pp. 740–743, 1994.

[22] N. Friedman, M. Goldszmidt, and T. Lee, "Bayesian network classification with continous attributes: Getting the best of both discretization and parametric fitting," in *Proc. Intl. Conf. on Machine Learning*, 1998.

Application of Low Discrepancy Sequences and Classical Control Strategies for Image Registration

Dinesh Nair
National Instruments
11500 N. Mopac B, Austin, TX 78759 USA
Email: dinesh.nair@natinst.com

Lothar Wenzel
National Instruments
11500 N. Mopac B, Austin, TX 78759 USA
Email: lothar.wenzel@natinst.com

Abstract

Image registration is the process of estimating the affine transformation made up of a rotation, a scale change, and a translation that maps one *n*-dimensional image into a second *n*-dimensional image. In practice, *n* is usually 2 or 3. Image registration usually plays the part of an important and critical first step in applications involving the fusion of information from multiple modalities.

In this paper we introduce two important concepts: (1) non-standard quasi-random sampling of *n*-dimensional images using low discrepancy sequences to select a set of *k* *n*-dimensional points in the image, and (2) the adaptation of a PID control strategy that uses the extracted subset of points to accurately determine the affine transformation that registers one image with another.

I. Introduction

Image registration is the process by which the correspondence between all points in two or more images of the same scene is determined. Image registration is used in image analysis tasks such as motion or change detection, fusion of data from multiple sensing modalities, and image geometric correction. There has been a tremendous increase in the need for good image registration techniques due to the increased use of temporal and multimodal 2-D and 3-D images in medical, remote sensing, and industrial applications. The aim of this paper is to introduce a couple of novel ideas for solving image registration where the images to be registered are an affine transformation, consisting of a spatial shift, rotation and scale change, apart. The registration method introduced here relies on classical control based strategies to determine the affine transformation parameters and the use of a unique sampling technique to perform the registration operation only on a subset of image points, thereby reducing the computation time.

The rest of the paper is organized as follows. We begin by introducing the theory of classical control and explain how this theory can be extended for image processing. Next, we introduce low discrepancy sequences and illustrate its use in image analysis applications. We end this paper by describing the registration framework that uses classical control theory and low discrepancy sequences in a complimentary manner. The objective of this paper is to introduce the ideas and illustrate the potential of the ideas through simple examples.

II. Classical Control

Control is an extremely well developed theory with a wide range of practical applications. Usually, this describes only the most straightforward control problems; a system is given with a specific input – output structure. The system could be a plant, an engine, a biological object, or any other structure

with a clear input – output behavior. The number of input – output channels is low or moderate. The inputs can be interpreted as control actions on the one hand, and noise sources on the other hand. Usually, one has information about statistical properties of the disturbances. The outputs of the given system consist mainly of measurements. These measurements describe internal states, which, in most cases, are not directly measurable. In the simplest case, a so-called set point must be reached and stabilized based on appropriate control actions. Even in highly nonlinear situations, linear models can successfully be used to determine such control strategies.

The general continuous linear model

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t)$$
$$y(t) = Cx(t) + Du(t) \qquad (1)$$

can be described by:

- matrices A, B, C, and D
- vector $x(t)$ of internal states
- vector $u(t)$ of control actions
- vector $y(t)$ of measurements.

Additionally, noise terms can be added if necessary. The most successful and very common control strategy is simply a feedback according to

$$u(t) = -K(t)x(t) \qquad (2)$$

The design of a valid control system consists mainly of the construction of matrices $K(t)$ making the controlled system as efficient as possible. In typical applications the matrix $K(t)$ is constant which simplifies the algorithms considerably.

PID controllers form a widely used subclass of control systems. More complicated systems are controlled by cascaded PID controllers. The term PID stands for a combination of P = proportional, I = integration, and D = differential components. Let $e(t)$ be the difference between the set point and the current measurement. Clearly, the goal is to reach $e(t) = 0$. The PID controller reacts with a control value $u(t)$ according to (many other notations are used)

$$u(t) = Pe(t) + I\int_0^t dse(s) + D\frac{de(t)}{dt} \qquad (3)$$

The quality of a PID controller depends completely on the quality of the choice of parameters (P, I, D). The preferable strategy is to tune (auto-tune) the PID controller using and observing the real behavior of the given system.

2.1 Control and Image Analysis

At a first glance, images (at least beyond a certain size) don't match the requirements of classical control. Each pixel can be interpreted as a separate channel. Though these channels are not completely independent, the overall degrees of freedom are enormous. On the other hand, tasks such as object tracking, the determination of sub-pixel accuracy in the final phase of pattern matching algorithms, or image registration can be formulated in terms of classical control theory. Here, we will closely follow Hoger and Belhumeur's approach. In [1] they describe iterative algorithms capable of tracking objects under very general motion models of the target (see also [2]). A straightforward reformulation and redefinition of the goal leads to a registration algorithm. Three new features are added.

First, Hoger and Bulhumeur's algorithms are tracking mechanisms, whereas our strategy tries to register two given and similar images. Second, the method developed in [1] is a direct consequence of a certain minimization routine. The result can be interpreted as a control strategy with P=1, I=0, and D=0. Third, to reduce the computational load, our algorithm is based on some well-defined parts of given images. More precisely, instead of taking into account all pixels in the image, we restrict the computation to low discrepancy sets. In doing so, there is no significant loss of accuracy but a speed up that is in the order of

pixels used from the images to be registered.

III. Low Discrepancy Sequences

Pseudo-random sequences have been used as a deterministic alternative to random sequences for use in Monte Carlo methods for solving different problems. Recently, it was discovered that there is a relationship between low discrepancy sets and the efficient evaluation of higher-dimensional integrals. Theory suggests that for midsize dimensional problems, algorithms based on low discrepancy sets should out perform all other existing methods by an order of magnitude in terms of the number of samples required to characterize the problem.

Given a function $f(x)$ the problem of calculating the integral

$$I(f) = \int_0^1 f(x)\,dx \qquad (4)$$

in the most efficient manner is not a well posed problem. An approximate strategy could be based on the following procedure:

(A) Construct an infinite sequence $\{x_1, x_2, x_3, \ldots, x_n, \ldots\}$ of real numbers in [0, 1] that does not depend on a specific function f (we do not know anything about f in advance, except some general smoothness properties).

(B) During the n^{th} step of the algorithm calculate $f(x_n)$ and the approximation to the integral in (4) as:

$$I_n(f) = (f(x_1) + \ldots + f(x_n))/n \qquad (5)$$

If a certain criterion is satisfied stop, else repeat step (B). The stopping criterion depends strongly on objectives such as accuracy or speed.

How does this algorithm differ from standard methods such as trapezoidal rule which is based on equally distributed points in [0, 1]? First, there is no relationship between consecutive sets $x_i(n) = i/n$ and $x_i(n)=i/(n+1)$. In other words, if the approximation given in equation (5) fails a given goal, a complete recalculation of numerous f-values is necessary. On the other hand, it is well known that the trapezoidal rule gives a $1/n^2$ rate of convergence for a given continuous function f.

Obviously, the quality of the trapezoidal rule is based on a highly homogeneous set of points. To quantify the homogeneity of a finite set of points, the definition of the so-called discrepancy of a given set was introduced ([3], [4]):

$$D(X) = \sup_R |m(R) - p(R)| \qquad (6)$$

Here, R runs over all rectangles [0, r] with $0 \le r \le 1$, $m(R)$ stands for the length r of the closed interval R, and p(R) is the ratio of the number of points of X in R and the number of all points of X. The definition given in equation (6) can be generalized to the case of d dimensions (d=2, 3, ...), where the term interval must be interpreted as an d dimensional rectangle. The lower the discrepancy the better or more homogeneous the distribution of the set. The discrepancy of an infinite sequence $X = \{x_1, x_2, \ldots, x_n, \ldots\}$, is a new sequence of positive real numbers $D(X_n)$, where X_n stands for the first n elements of X. Other definitions for the discrepancy do exist that avoid the worst-case scenario according to (6).

Clearly, there exist a set of points of given length, that realizes the lowest discrepancy. It is well-known ([5]) that the following inequality holds true for all finite sequences X of length n in the d dimensional unit cube

$$D(X) \le B_d \frac{(\log n)^{(d-1)/2}}{n}. \qquad (7)$$

B_d depends only on d. Except for the trivial case $d=1$, it is not known whether or not the theoretical lower bound is attainable.

Many schemes to build finite sequences X of length n do exist that deliver a slightly worse limit

$$D(X) \leq B_d \frac{(\log n)^d}{n}. \quad (8)$$

There are also infinite sequences X with

$$D(X_n) \leq B_d \frac{(\log n)^d}{n} \quad (9)$$

for all natural numbers n. The latter gave rise to the definition of so-called low discrepancy (infinite) sequences X. The inequality in equation (9) must be valid for all n, where B_d is an appropriately chosen constant. Low discrepancy sequences are also known as quasi-random sequences.

On average, a randomly chosen sequence in [0, 1] has a discrepancy value in the order of $1/\sqrt{n}$ which is far beyond the low discrepancy value in the order of $(\log n)^d/n$.

The relationship between the integration in equation (4), its d-dimensional generalization, and the approximation given by equation (6) for an infinite sequence $X=\{x_1, x_2 \ldots, x_n, \ldots\}$ is given by the Koksma-Hlawka inequality.

$$|I(f) - I_n(f)| \leq V(f) \cdot D \quad (10)$$

where, $V(f)$ is the variation of the function in the sense of Hardy and Krause. Even if a finite $V(f)$ exists (e.g., for smooth functions), the inequality in equation (10) is of no practical use. Very often, the real value of $V(f)$ is unknown and describes only the worst-case. But, at least in principle, a low discrepancy set X should be preferred to all other sequences.

3.1 The Halton Low Discrepancy Sequence

Many of the well-studied low discrepancy sequences in the d-dimensional square can be constructed as combinations of 1-dimensional low-discrepancy sequences. The most popular low discrepancy sequences are based on schemes introduced by Richtmeyer [5], Halton [4], Sobol' [6, 7], Niederreiter [8], and Faure [9]. The book [7] gives a comprehensive introduction into the implementation of low discrepancy sequences (Halton and Sobol'). We will explain the Halton method here in detail. All of our test results are based on Halton and Sobol' sequences.

Halton sequences in 1-d start with the choice of a natural number greater than 1. Though not absolutely necessary, prime numbers $p = 2, 3, 5, \ldots$ are typically chosen. If p is a given prime number and x_n the n^{th} element of the Halton sequence, the following algorithm determines x_n.

(A) write n down in the p-ary system

$$n = n_q \ldots n_0,\ n = n_0 + n_1 \cdot p + \ldots + n_q \cdot p^q$$

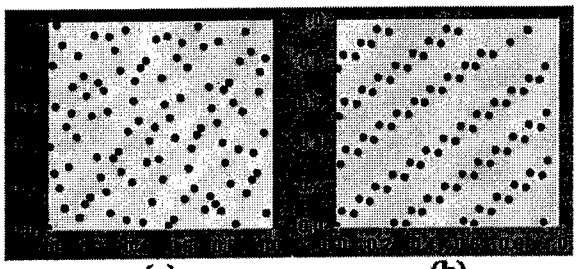
(B) Reverse the order of the digits and add the p-ary point

$$0.n_0 n_1 \ldots n_q$$

(C) It is

$$x_n = n_0 \cdot p^{-1} + n_1 p^{-2} + \ldots + n_q p^{-(q+1)}$$

The n^{th} element of the Halton sequence can be calculated independently of all other elements. As mentioned above, in d dimensions one has to interpret different 1-dimensional Halton sequences as coordinates of points in d dimensions. It is very common to start with the first d prime numbers. Figure 1 shows the first 100 elements of a Halton sequence in the unit square for two different valid choices of starting prime numbers (2, 3) and (13, 17). Obviously, the first couple performs much better at least at the very beginning of the sequence. Because of the relatively low number of pixels in typical image processing applications, homogeneity (low discrepancy value) is always desirable. That is why, all experiments were based on the most straightforward combination (2, 3).



(a) (b)

Figure 1: Distribution of the Halton sampling points on a unit square. (a) With prime numbers (2,3). (b) With prime numbers (13,17).

Halton sequences in 1-D are low-discrepancy sets in the sense of equation (8). More precisely, for all n and for all Halton sequences X that are based on a prime number p it is ([5])

$$D(X) \le B\frac{\log n}{n} \quad \text{with}$$

$$B = \begin{cases} \dfrac{p^2}{4(p+1)\log p} & \text{when p is even} \\ \dfrac{p-1}{4\log p} & \text{when p is odd} \end{cases}$$

A similar result (see [5] again) holds true for Halton sequences in d dimensional unit squares. In a 2-dimensional unit square for the (p,q) Halton sequence with prime numbers p and q the discrepancy is

$$D(X) \le \frac{2}{n} + \frac{(\log n)^2}{n}\left[\left(\frac{p-1}{2\log p} + \frac{p+1}{2\log n}\right)\left(\frac{q-1}{2\log q} + \frac{q+1}{2\log n}\right)\right]$$

3.2 Low Discrepancy sequences and images

Most routines in image processing are based on sampled objects where the resolution is limited by hardware and can not be influenced by the user. On the other hand, low discrepancy sequences are inherently continuous. Given a digital image of size $N^{(1)}$ x $N^{(2)}$, the n^{th} pixel element according to the given low discrepancy sequence $(x_1, x_2, ..., x_n, ...)$ in the unit square has the coordinates $[N^{(1)}.x_n^{(1)}]$ and $[N^{(2)}.x_n^{(2)}]$, where $x_n = (x_n^{(1)}, x_n^{(2)})$ and [] stands for the nearest integer. Because of the homogeneity of low discrepancy sequences, double hits are impossible if n is sufficiently small. The final goal of a combination of low discrepancy sequences and image processing is the significant reduction of information processing necessary for image analysis. Good approximations must be delivered with the aid of a small percentage of all available pixel values.

The field of image processing and image understanding can potentially take advantage of specific properties of low discrepancy sets. To illustrate this, we applied the theory of low discrepancy sequences to some relatively simple image processing and computer vision related operations such as the estimation of gray level image statistics and fast location of objects in a binary image. The results of our experiments are tabulated below. In the first experiment, we estimated the average number of points as a percentage of the mage size needed to estimate the mean of the image with a certain accuracy. Accuracy was defined in terms of the absolute difference from the true value. The image database used in this experiment comprised of images of man-made objects, textures, medical imagery, fractals etc.

Method	1.0 Accuracy	0.5 Accuracy
Halton	0.2	0.7
Random	0.7	3.3
Grid	0.6	2.3

Table 1: % of number of image points required for the estimation of the mean gray-scale value of an image using different sampling techniques.

In the second experiment, the objective was to determine the number of points needed to locate a randomly placed binary rectangle in an image with probability beyond 0.5. The image size was 512x512. The object was assumed to be located if a point on the sequence fell inside the boundaries of the object.

Method	Number of Points
Halton	100
Random	115
Grid	114

Table 2: Number of sample points required for locating an arbitrarily placed object in an 512x512 binary image using different sampling techniques.

Our experiments show that compared to standard methods, the proposed new algorithms require fewer points than regular grid-based sampling and random sampling to accurately characterize images. Hence these algorithms are faster and statistically more robust than conventional sampling techniques.

IV. Image Registration Framework

This section presents the method used for image registration using low discrepancy sequences and a control strategy.

Given two images imageA and imageB, let imageB be a shifted, rotated, and scaled version of imageA. We assume that the four parameters x, y, θ, and s (x-shift, y-shift, θ-rotation, s-scaling factor) are relatively close to 0, 0, 0, and 1, respectively. Given two reasonably sized regions, regionA and regionB, where regionA is part of imageA and regionB is part of imageB respectively, regionA matches regionB with unknown values x, y, θ, and s. The goal is to determine these values accurately.

4.1 The control strategy

The goal is to reduce the distance between regionA and a shifted, rotated, and scaled version of regionB (as part of imageB) with the aid of a step-by-step approach. The original situation is described by an unknown x, y, θ, and s. The control strategy can be divided into two parts ([1]).

Preparation:

(1) Calculate the Prewitt derivatives of regionA and flatten these derivatives. This results in vectors Ix and Iy

(2) Build the matrix

$$M_o = \begin{bmatrix} I_x & I_y & -yI_x + xI_y & xI_x + yI_y \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

Calculate the matrix $\lambda = \left(M_0{}^T M_0\right)^{-1} M_0{}^T$.

The matrix $M_0{}^T M_0$ is 4x4 and non-singular (with exception of some pathological situations).

Control:

(1) Let (x, y, θ, s) be the current estimates.

(2) Let $\rho(\theta,s)$ be the matrix

$$\rho(\theta,s) = \begin{bmatrix} \cos(\theta)/s & -\sin(\theta)/s & 0 & 0 \\ \sin(\theta)/s & \cos(\theta)/s & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{s} \end{bmatrix}$$

(3) Let e be the flattened difference (pixelwise) between regionA and the shifted, rotated, and scaled version of regionB. Then

$$[\Delta x, \Delta y, \Delta\theta, \Delta s] = \rho(\theta,s) * \lambda * e$$

and

$$x_{new} = x + \Delta x$$
$$y_{new} = y + \Delta y$$
$$\theta_{new} = \theta + \Delta\theta$$
$$s_{new} = s + \Delta s$$

(4) Calculate the new $(x_{new}, y_{new}, \theta_{new}, s_{new})$ version of regionB, i.e. shift the old regionB by $(\Delta x, \Delta y)$, rotate it by $\Delta\theta$, and shrink or stretch it by the factor Δs using bilinear interpolation.

(5) Depending on the value of the norm of e stop or go to (3) again.

Figure 2 shows a typical pair of images that were

used for the registration experiments. In this example the image shown in Figure 2(b) is a shifted, rotated, and scaled version of the image in Figure 2(a). Figure 3 depicts the typical response of this registration algorithm. The (x, y, θ, s) data slowly approach the real values which are completely known in this example.

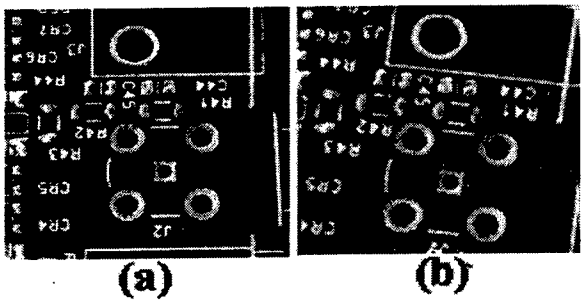


Figure 2: A pair of images used for the registration example. The image in (b) is a shifted, rotated, and scaled version of the image in (a).

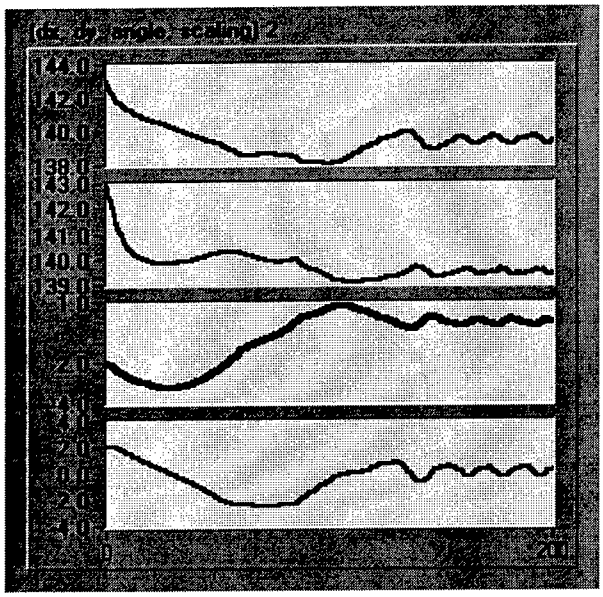


Figure 3: Response time of the control-based registration technique using the entire image.

More efficient strategies:

One of the drawbacks of the described algorithm is its sluggish convergence behavior. Though (3) is a direct consequence of an optimization procedure ([1]), it is usually not the fastest procedure. The performance increases by introducing constant factors kx, ky, $k\theta$, *and* ks. The new estimates are:

$$x_{new} = x + kx * \Delta x$$
$$y_{new} = y + ky * \Delta y$$
$$\theta_{new} = \theta + k\theta * \Delta\theta$$
$$s_{new} = s + ks * \Delta s$$

This is nothing else but four P-controllers acting on the four channels x, y, θ, and s. Figure 4 shows the same situation as before but with k-values: kx=5, ky=5, $k\theta$=5, and ks=5. The performance is significantly better than with the choice of (1, 1, 1, 1).

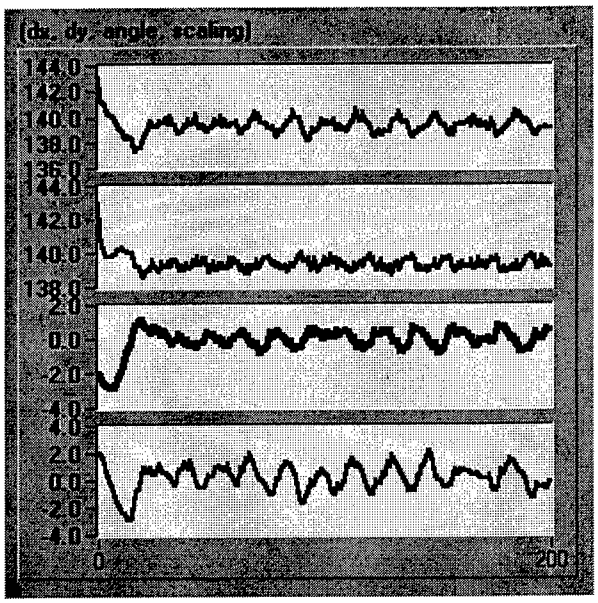


Figure 4: Response of the registration of the images in Figure 2 using a P value of 5 for each of the affine parameters to be estimated.

Autotuning:

The last remark gives rise to the introduction of complete PID-controllers for the four above-mentioned channels. The problem is that there is no set of P-, I-, and D-components covering all situations equally well. In other words, depending on the image content different sets of parameters must be chosen. In the majority of our experiments the set

$kx = 5$, $ky = 5$, $k\theta = 5$, and $ks = 5$ (where s is measured in % times 100) showed an excellent behavior. The I- and D-components were set to 0.

If any of the images being registered is noisy, the D-values should be 0 or very close to this value. In case of noise-free images or filtered versions of the originals a D value different from 0 can result in an even more improved speed of convergence. By introducing an I term, you can reduce the oscillatory behavior seen in figure 4. In general, there is no strict method of determining the optimal sets of P-, I-, and D-parameters. A successful approach, if applicable, can be described as follows. Shift, rotate, and scale regionA artificially and measure the speed of convergence under different conditions, i.e. different sets of parameters. Choose the set with the best behavior.

This method has some obvious similarities to auto-tuning that is used commonly in classical control. Another method is based on adaptive control, i.e. modifications of an initial set of parameters in dependence of the norm of the error e. We prefer the first procedure to the second one because the auto-tuning method produces very similar results for images belonging to the same family.

4.2 Use of Low discrepancy Sequence Points

A careful study of the developed control strategy reveals that the Prewitt derivatives Ix and Iy can easily restricted to specific parts of the image without changing the algorithms. Earlier we demonstrated that deterministic random sequences (low discrepancy sets) out perform other choices based on the same amount of pixels. More precisely, an excellent estimate of the average gray level of a given image can be achieved if only a very small percentage of all pixel values are considered. The well distributed Halton or Sobol' sequences interpreted as pixel positions on an image deliver much better results than randomly chosen pixels or grid like structures. Another advantage of low discrepancy sequences is the ability to add further points without loosing results achieved so far. Both properties make Halton, Sobol' and other low discrepancy sequences superior to comparable choices.

Figure 5 is the result of the new control strategy where only 10% of all points were used. These pixels belong to a Halton sequences. Compared to a full set of pixels, the speed of convergence is as fast as in the original case. Randomly chosen pixels can not guarantee this behavior.

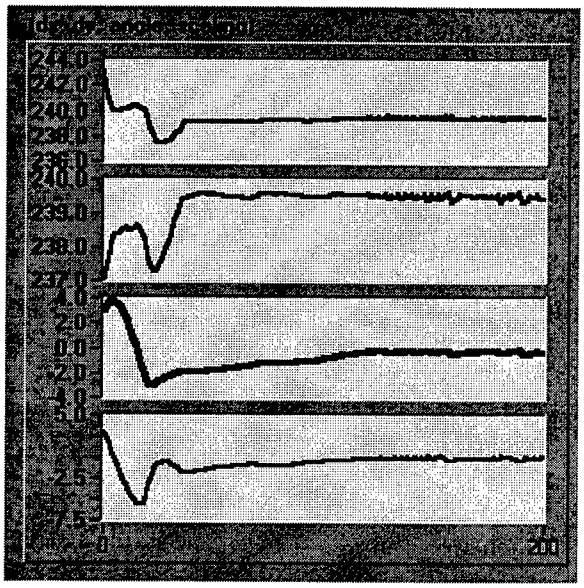


Figure 5: Convergence behavior of the registration algorithm using Halton points.

The described algorithms have their limitations. They do not perform well if the shift is beyond +/-4 pixels (both in x- and y direction), if the rotation is beyond +/-4 degrees, and if the scaling factor is beyond +/-4%. Typical region sizes are in the order of 80x80. If one can not guarantee these conditions an additional step is necessary. Based on pattern matching or

similar techniques, a first estimate must be generated that satisfies the above mentioned parameters.

IV. Summary

In this paper a new image registration technique is presented. The problem of finding the affine transformation parameters between the images to be registered is posed as a classical control problem. An efficient and robust image registration is performed by sampling the images to be registered using low discrepancy sequences and estimating the transformation between two subsets using a strategy based on the PID control.

References

[1] G. D. Hoger, P. N. Belhumeur; "Efficient region tracking with parametric models of geometry and illumination," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No 10, Oct. 1998

[2] B. C. Lucas, T. Kanade; "An iterative image registration technique with an application to stereo vision," Proc. Int'l Joint Conf. Artificial Intelligence, pp. 674 – 679, 1981

[3] J.G. von der Corput; Nederl. Akad. Wetensch. Proc. Ser. B 38 (1935) 813, 1058

[4] J. H. Halton; "On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals," Numer. Math. 2 (1960) 84-90

[5] F. J. Hoglund, R. Keiss; "Multidimensional sampling for simulation and integrals," accepted for publication in Computer Physics Communication

[6] I. M. Sobol'; "On the distribution of points in a cube and the approximative evaluation of integrals," USSR Computational Mathematics and Mathematical Physics, 7, 86-112, 1967

[7] W. Press, S. Teukolsky, W. Vetterling, B. Flannery; Numerical Recipes in C, Cambridge University Press, 1992

[8] H. Niederreiter; Random number generation and quasi-Monte Carlo methods, CBMS-NSF Regional Conference Series in Applied Math., No.63, SIAM, 1992

[9] H. Faure; Discrepance de suites associees a un systeme de numeration (en dimension s), Acta Arithmetica 41, 337-351, 1982

Video Indexing using Color and Textural Features

Dinkar Bhat Jia Wang Katja Sakiewicz N. Nandhakumar

LG Electronics Research Center of America, Inc.
40 Washington Road, Princeton Jn, NJ 08550

Abstract – We propose a color and texture based descriptor for shots in a video, using an appropriate keyframe. The color descriptor is formed through a set of histograms computed in different regions of the keyframe image, at different tessellation levels. The texture descriptor is formed by choosing a set of coefficients in the wavelet transform of the keyframe. Together they form a descriptor for a keyframe, which in turn forms a descriptor for the shot. Experiments on videos in the MPEG-7 test database suggest that the descriptor is reliable for shot retrieval and video indexing.

Keywords: Image Indexing, Color and Texture Features, Multimedia Databases, MPEG-7

1 Introduction

Temporally segmenting video into shots facilitates non-linear video browsing, editing and search ([4],[5]) in digital video management systems [4]. Thus, shots form basic units of video, but their content description for indexing, in terms of color, texture, motion, etc, differs considerably between systems. This lack of standardization results in poor interoperability between systems that manage video. With the onset of the MPEG-7 (a.k.a "Multimedia Content Description Interface") standardization effort, formalizing video structure and its description has taken a very important role [1],[2].

MPEG-7 seeks to standardize a set of descriptors that can be used to describe various types of multimedia information. These descriptors shall be associated with the content itself, to allow fast and efficient searching for material of a user's interest. Audio-visual material that has MPEG-7 data associated with it, can be indexed and searched for. This 'material' may include: still pictures, graphics, 3D models, audio, speech, video, and information about how these elements are combined in a multimedia presentation.

In the context of video, MPEG-7 seeks to define descriptors for shots that would then permit efficient searching and indexing into video, and permit interoperability between video databases. Several requirements have been established for descriptors, including ease of computation, expressability, and comprehensiveness. In this paper, a descriptor for shots is described that we proposed in the MPEG-7 evaluation meeting at Lancaster, U.K earlier this year [9]. The descriptor uses color and textural features of keyframes in shots that have been identified using a shot detection technique like [6].

The color descriptor is composed of a set of histograms computed in different regions of a keyframe image, at different tessellation levels. The texture descriptor is formed by choosing a set of coefficients in the wavelet transform of the keyframe [8]. The color and texture descriptors are combined to form a descriptor for a keyframe. We evaluate the shot retrieval performance of the descriptor using the MPEG-7 test database [3]. Experiments suggest the utility of the descriptor in search and indexing into video.

2 Our Approach

We present a composite descriptor for shots in a video. Each shot Φ in a video consists of a set of consecutive image frames $F = \{f_i; i = 1,..,n\}$. Shot Φ can be compactly described by a set of keyframes:

$$R = \{r_i; i = 1,..,m, r_i \in F, m \le n\}.$$

In this paper, we assume that a shot is represented by a single keyframe, i.e. m=1.

The key frame is described by a combination of color and texture descriptors.

The color descriptor is represented at multiple tessellations of the keyframe r_i. By using multiple tessellations, we obtain varying representations of color distribution - at higher tessellation, the representation is more local. The color at each tessellation level is described using a set of histograms computed at different image regions. Each histogram is specified by the location and size of the image region over which it is computed. The size of image regions over which the histograms are computed is reduced over successive tessellations[1]. Therefore, if a histogram is denoted by $h(\mathbf{p}_j, \mathbf{w}_{j,l})$ where $\mathbf{p}_j$ denotes location of the j^{th} image region, and $\mathbf{w}_{j,l}$ denotes its size, at tessellation level l, then color descriptor H_l at level l is given by:

$$H_l = \bigcup_{j=1}^{N} h(\mathbf{p}_j, \mathbf{w}_{j,l}) \qquad (1)$$

where N denotes the number of image regions over which histograms are computed. The overall color descriptor C_i for keyframe r_i is given by:

$$C_i = \bigcup_{l=1}^{L} H_l \qquad (2)$$

where L is the total number of tessellation levels used.

The texture descriptor consists of a set of coefficients that represent spatial variation of detail in the keyframe image, at different image resolution(s). Detail images of different resolutions are computed using the discrete Haar wavelet transform on the luminance component. Let the detail images at texture resolutions $d_0, d_1, d_2 ..., d_n$ be chosen for representation, and at each texture resolution only q_k highest valued coefficients be preserved, where $k = d_0, d_1, ...d_n$. If the quantized detail image at texture resolution d_i is represented as $\overline{D}_{d_i}$ where $i = 0,1,..,n$, then the texture descriptor T_i is given by:

$$T_i = \{\overline{D}_{d_0}, \overline{D}_{d_1}, ...\overline{D}_{d_n}; q_{d_0}, q_{d_1}, ...q_{d_n}; d_0, d_1, ...d_n\} \quad (3).$$

The overall descriptor Ω_i for keyframe r_i is given by: $\Omega_i = (C_i, T_i)$. Since we assume that the shot is represented by one keyframe, the shot descriptor Ω is identical to its keyframe descriptor.

3 Descriptor Computation

In this section, we discuss how the color and texture descriptors are computed from a keyframe image.

3.1 Computing the Color Descriptor

The procedure to compute the color descriptor consists of two steps: segmentation of a frame image into regions for a fixed number of tessellation levels, and computation of the color histogram of each region. Figure 1 depicts two possible structures, quin-tree and quad-tree, with regular tessellation, i.e. any image region is further sub-divided into equal sized regions to obtain at a higher

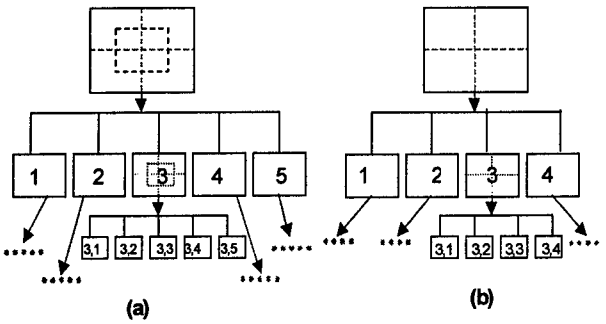


Figure 1: Illustration of Quad-tree and Quin-tree tessellation of a keyframe.

tessellation. Two levels of tessellation are shown. Note that in the case of the quin-tree, subdivided image regions overlap.

[1] In the case of regular tessellations where the image is subdivided into equal-sized regions, the window size is automatically defined by the tessellation level.

3.2 Computing the Texture Descriptor

The texture descriptor T_i is computed using the Haar wavelet transform. Intensity of pixels in a key frame image is projected onto smoothing and detail filters recursively until a multi-resolution representation of the image is obtained. At a coarse resolution, broader details in the image are represented, and at a finer scale, smaller details begin to emerge. Thus, the wavelet representation is a way of describing the texture in the keyframe image.

If the image intensity data in keyframe r_i is represented as $I(x,y)$, then the wavelet transformation of the image can be expressed as:

$$L_d(x,y) = [H_x * [H_y * L_{d-1}(x,y)]_{\downarrow 2,1}]_{\downarrow 1,2}$$
$$D_d^1(x,y) = [H_x * [G_y * L_{d-1}(x,y)]_{\downarrow 2,1}]_{\downarrow 1,2}$$
$$D_d^2(x,y) = [G_x * [H_y * L_{d-1}(x,y)]_{\downarrow 2,1}]_{\downarrow 1,2}$$
$$D_d^3(x,y) = [G_x * [G_y * L_{d-1}(x,y)]_{\downarrow 2,1}]_{\downarrow 1,2}$$

where * denotes convolution, $_{\downarrow 2,1(\downarrow 1,2)}$ denotes sub-sampling along the *x(y)* axis, and $L_0(x,y) = I(x,y)$. H and G denote the Haar low-pass and band-pass filters, composed of separable components $\{H_x, H_y\}, \{G_x, G_y\}$, respectively. The Haar basis defines filters as follows: $H_x = \{1\ 1\}$, $H_y = \{1\ 1\}^T$, $G_x = \{1\ -1\}$, $G_y = \{1\ -1\}^T$. L_d is the output of low-pass filtering, hence it is a low-resolution image at resolution *d*. D_d^1, D_d^2, D_d^3 are outputs[2] of band-pass filtering along specific orientations – horizontal, vertical, and diagonal, respectively. They represent detail at resolution *d*, and hence, they are referred to as detail images. At a resolution *d*, the original image $I(x,y)$ can be reconstructed from the set of images: $\{L_d; D_k^1, D_k^2, D_k^3, k = 1,2..,d\}$.

Only the coarse detail of an image is used as its texture descriptor. The reason is that during retrieval, the broad detail between images are compared without considering finer texture. Therefore, the descriptor is computed in two steps. First, we select appropriate resolutions for representation given by $d_0, d_1, ...d_n$. Second, from detail images $D_{d_i}^1, D_{d_i}^2, D_{d_i}^3$, we select the top (in absolute value) $q_{d_i}^1, q_{d_i}^2, q_{d_i}^3$ coefficients, respectively, where $i = d_0, d_1, ..., d_n$, i.e we quantize the set of coefficients. All other coefficients in the detail images are set to 0. If $\overline{D}_{d_i}^1, \overline{D}_{d_i}^2, \overline{D}_{d_i}^3$ are the resulting quantized detail images and $\overline{D}_{d_i} = \{\overline{D}_{d_i}^1, \overline{D}_{d_i}^2, \overline{D}_{d_i}^3\}$, for $i = d_0, d_1, ..., d_n$, then:

$$T_i = \{\overline{D}_{d_0}, \overline{D}_{d_1}, ..., \overline{D}_{d_n}\}.$$

4 Similarity Measure

To retrieve shots in a database that are similar to a query shot, distance measures are required for measuring similarity between shots. Let Φ_i, Φ_q denote the shots being compared. Φ_i is represented by keyframe r_i, and Φ_q is represented by r_q. The composite descriptor of keyframe r_i is $\Omega_i = (C_i, T_i)$, and that of r_q is $\Omega_q = (C_q, T_q)$. Below, we describe a method to compare the two keyframe descriptors.

4.1 Color Similarity Measure

To determine whether the two keyframes r_i and r_q are similar in terms of color content, the color histogram intersection function [7] is applied to corresponding regions of r_i and r_q as follows:

$$I_{i,q}^j = \bigcap (h_i(\mathbf{p_j}, \mathbf{w_{j,l}}), h_q(\mathbf{p_j}, \mathbf{w_{j,l}}))$$
$$= \sum_{\beta=1}^{\alpha} \min(h_i^{(\beta)}(\mathbf{p_j}, \mathbf{w_{j,l}}), h_q^{(\beta)}(\mathbf{p_j}, \mathbf{w_{j,l}}))$$

[2] The detail images are also referred to as horizontal, vertical and diagonal channels.

where $h^{(\beta)}(.)$ denotes the value of a histogram in bin β, and l is the tessellation level of the color descriptor at which similarity is computed.

To account for different possible contributions of the regions, a weighted similarity measure function normalized by total number of image regions N at level l, is defined as follows:

$$S_c(C_i,C_q)=\frac{\sum_{j=1}^{N}\delta_j\cdot(\frac{I_{i,q}^j}{\|\mathbf{w}_{j,l}\|})}{N} \quad (4)$$

where δ_j is the weight for the jth region, and $\|.\|$ denotes magnitude. δ_j can be varied, for example, to perform home video shot boundary detection, the center region in a Quin-tree division could be emphasized higher than the other four segments since it has a higher chance of capturing the object of interest. Similarity could also be computed at multiple tessellation levels in a hierarchical fashion, starting from the coarsest level.

4.2 Texture Similarity Measure

To match texture descriptors T_i, T_q, we use the following distance criterion [8]. It is assumed that only one texture resolution d_0 is used for representation.

As described earlier, $T_i=\{\overline{D_{d_0}^1},\overline{D_{d_0}^2},\overline{D_{d_0}^3}\}$. Let $T_q=\{\overline{Q_{d_0}^1},\overline{Q_{d_0}^2},\overline{Q_{d_0}^3}\}$ where $\overline{Q}_{d_0}^i, i=1,2,3$ represent quantized detail images of the query (at the same quantization levels as the texture representation for keyframe r_i), at scale d_0. The similarity measure is given by:

$$S_T(T_i,T_q)=\frac{1}{R*Q}\sum_{m=1}^{3}\sum_{(x,y)}w_{(x,y)}^m F(\overline{D}_{d_0}^m(x,y),\overline{Q}_{d_0}^m(x,y)) \quad (5)$$

where $F(.,.)$ is a metric defined as:

$F(a,b) = 1$ if $ab > 0$, and
0, if $ab = 0$.

$Q=q1+q2+q3$, $w_{(x,y)}^m$ is a spatially varying weight function, and R is a normalization constant. In the unweighted case R=1, and $w_{(x,y)}^m=1; m=1,2,3; \forall x,y$. When the two descriptors are identical, $S_T(T_i,T_q)=1.0$. If coefficients in the two descriptors agree, then $S_T(T_i,T_q)=0.0$.

Coefficients in T_i with large absolute value are important, and if corresponding coefficients in T_q do not agree, then the deviation must be penalized significantly. Therefore, if weighting is desired, a logical way would be to set $w_{(x,y)}^m=|D_{d_0}^m(x,y)|$, m=1,2,3. In case multiple resolutions are adopted, $S_T(T_i,T_q)$ is computed individually for each resolution, and the results are summed. The resulting value is then normalized by the number of texture resolutions used.

4.3 Combined Measure

To obtain a combined similarity measure, we weight the two components appropriately. If $S(\Omega_i,\Omega_q)$ is the aggregate measure, then:

$$S(\Omega_i,\Omega_q)=\lambda_1 S_c(C_i,C_q)+\lambda_2 S_T(T_i,T_q) \quad (6)$$

5. Experiments

A prototype system for shot retrieval was developed based on the proposed descriptor, and the similarity scheme described above. For the experiments, approximately 400 shots were collected from three video files belonging to the MPEG-7 test set [3] (Harmony.mpg(**Item V25**), animals.mpg (**Item V14**) and Cm1002.mpg(**Item V19**)). The MPEG-7 test set was created to enable uniform comparison across different descriptors. The size of each keyframe is 160 x 112. The following are the parameters for our experiments:

- The first frame in a shot was used as its keyframe.
- A quin-tree was used, and the number of tessellation levels was set to 2, i.e L=2.
- The number of bins (α) in the color histogram was set to 64.

- One tessellation level ($L=2$) was used for computing color similarity ($S_c(C_i, C_q)$).

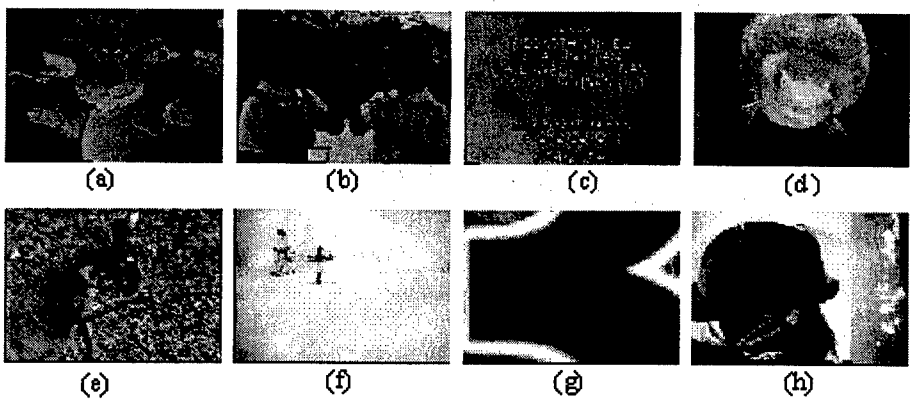


Figure 2: Keyframe images from selected query shots.

- One texture resolution was used for the texture descriptor($d_0 = 3$).
- The quantization level used in the texture descriptor was 60, $q^1_{d_0} = q^2_{d_0} = q^3_{d_0} = 60$.
- The values $\lambda_1 = 0.6, \lambda_2 = 0.4$ were used in equation (6).
- Similarity measures for color and texture are unweighted (in equations (4), (5)).
- When similarity between shots falls below ξ =0.31, they are considered dissimilar.

In our experiments, we used 8 query shots whose keyframes are shown in Figure 2. Precision-recall metrics are used to measure the correctness and completeness of image retrieval. Given a query, let A be the set of top N similar images returned by the search engine, and let I be the ideal set of similar images that were pre-determined by visual inspection. Note that the number of images in A may be less than N, since the images in A must have similarity values $\geq \xi$ with respect to the query. The precision P and recall R of the image retrieval are calculated as follows:

$$P = \frac{|I \cap A|}{|A|},$$

$$R = \frac{|I \cap A|}{|I|}.$$

The average number of images in A for all queries is 15.25. The average number of images in I for all queries is 9.875. Based on the average A and I values, the performance of the retrieval is measured according to the interpretations showed in Table 1.

Interpretation	P	R
high	> 0.55	> 0.75
good	0.30 - 0.55	0.5 - 0.74
low	< 0.3	< 0.5

Table 1: A sample retrieval performance interpretation table. The ranges for precision and recall, for each interpretation, are arbitrary.

Experimental results are reported in Table 2, which shows the average P and R values for N = 40. Apparently, the system does well with both precision and recall. The average precision P value of 0.87 indicates that among the 15.25 retrieved images, 13.26 are similar to the query. The average recall value of R=0.61 indicates that 6.02 of the 9.875 similar images are retrieved.

	a	b	c	d	e	f	g	h	mean
P	0.8	1	0.9	1	0.4	0.9	1	1	0.87
R	1	0.3	0.6	0.6	0.7	1	0.6	0.2	0.61

Table 2: Precision and recall of image retrieval using 9 randomly selected keyframes as query images, with threshold similarity value set at 0.31.

Figure 3 presents two examples of the retrieval results. Figure 4 and Figure 5 show the retrieval effectiveness using precision vs. recall for query shots (a) and (h). The system

successfully retrieved all similar shots from the video database.

5 Discussion

We presented a color and texture descriptor for shots using their keyframes. This descriptor allows for efficient browsing and retrieval of shots from a database, as demonstrated on selected videos in the MPEG-7 test database.

Future research work would proceed along the following directions: a) Development of algorithms to combine descriptors from several keyframes in a shot. b) Selection of suitable parameter values in the descriptor such that MPEG-7 compatible databases can "communicate". c) Experiments with different kinds of video like home video, broadcast video, etc, to evaluate the suitability of the descriptor for diverse classes of video.

MPEG-7 is currently in the process of developing core experiments for descriptors that were deemed important at the evaluation meeting. Selected descriptors would then form part of the experimentation model (popularly known as XM). Design of software components that enable diverse descriptors to cooperate is a key issue that has to be addressed.

6 References

[1] ISO/IEC JTC1/SC29/WG11, "MPEG-7: Context and objectives", 1999.

[2] ISO/IEC JTC1/SC29/WG11, "MPEG-7: Requirements document", 1999.

[3] ISO/IEC JTC1/SC29/WG11, "MPEG-7: Evaluation document", 1999.

[4] J.Y. Chen, C. Taskiran, E. J. Delp, and C. A. Bouman, "ViBE: A new paradigm for video database browsing and search", *Proc. of IEEE Workshop on Content-based Access of Image and Video Libraries*, pp. 96-100, 1998.

[5] F. Arman, A. Hsu, and M.-Y.Chiu, "Feature management for large video databases", *Proc. of SPI/IS&T Conf. On Storage and Retrieval for Image and Video Databases*, vol. 1908, pp 2-12, 1993.

[6] N. V. Patel and I. K. Sethi, "Video shot detection and characterization for video databases", *Pattern Recognition*, 30(4), pp 583-592, 1997.

[7] M. Swain and D. Ballard, "Color indexing", *International Journal of Computer Vision*, 7(1), pp:11-32, 1991.

[8] E. Stollnitz, T. D. Derose, D. H. Salesin, "*Wavelets for computer graphics-Theory and applications*", Morgan Kaufmann Publishers, 1996.

[9] D. N. Bhat, J. Wang, N. Nandhakumar, *MPEG-7, Proposal P68*, Lancaster, UK, February 1999.

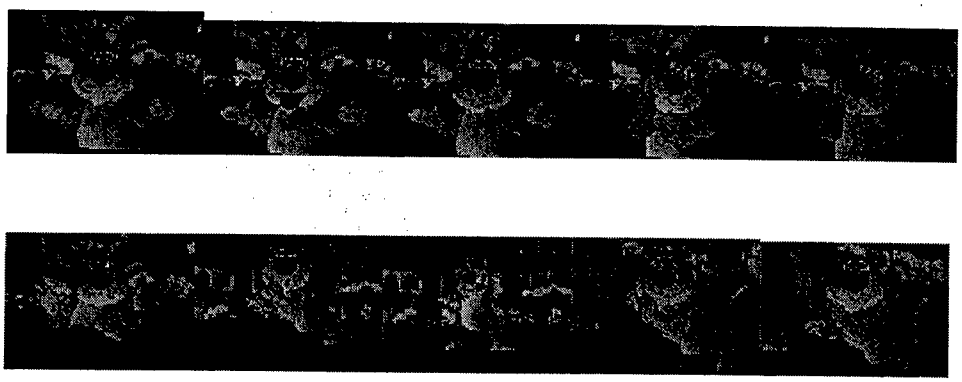
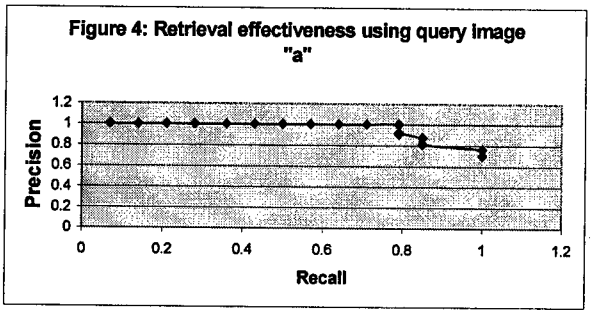
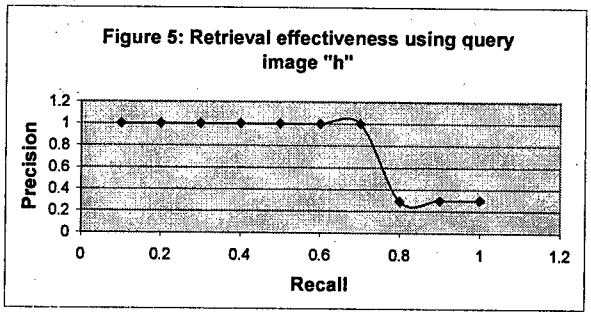


Figure 3(a): Retrieval results using shot number 42 (represented by keyframe 13015) in Harmony.mpg. The query image is shown on the top-left. Retrieved shots are shown in row-major form sorted by similarity value.

Figure 3(b): Retrieval results using shot number 97 (represented by keyframe 8753) in Cm1002.mpg. The query image is shown in a box on the top-left. The retrieved shots are shown in row-major form sorted by similarity value.





Statistical Decision Integration Using Fisher Criterion

S. Shah
Laboratory for Visual Computing
Wayne State University
Dept. of Electrical and Computer Engg.
Detroit, MI, U.S.A.

J. K. Aggarwal
Computer & Vision Res. Ctr.
The University of Texas at Austin
Dept. of Electrical and Computer Engg.
Austin, TX, U.S.A.

Abstract *This paper describes a new methodology for information integration based on the Fisher criterion. A system that uses information from multiple features or sensors can employ redundancy, diversity and complementarity to overcome the shortcomings of single-sensor systems and improve performance. In this paper, a general multifeature/multisensor framework is proposed which does not simply expand the dimensionality of the feature space, but which can discern new features to provide greater discrimination. Using this framework, a more focused methodology is described for localization of objects in complex scenes by learning multiple feature models in images. The methodology is based on a modular structure consisting of multiple classifiers, each of which solves the problem independently based on its input observations. A higher level decision integration is obtained through a supra-Bayesian scheme. Results of the proposed integration scheme are compared to existing combining techniques.*

Keywords: multisensor fusion, object detection, pattern analysis

1 Introduction

The automated interpretation of images to detect and localize objects is a key component of autonomous vision systems. Due to the large amount of data to be processed, the presence of noise in the imagery, the absence of complete information, the ill-posed nature of the problems, and inadequate modeling of the scene and the sensors, such extraction of information is a very complex task. Humans are able to detect and recognize as many as 10,000 distinct objects [2] under varying viewing conditions, while a state-of-the-art object recognition system can recognize relatively few objects. We know very little about the physiological mechanisms with which the human visual system solves and uses solutions to lower-level processes such as depth and shape in the task of object detection and recognition [5].

The process of object localization and recognition involves processing at all levels of machine vision: lower-level vision, as with edge detection and image segmentation; mid-level vision, as with representation and description of pattern shape, and feature extraction; and higher-level vision, as with classification. Since objects are usually characterized by their shape and by the gray-scale representation of the segmented region, detection results directly affect the performance of the system. Past research in machine perception has focused mainly on the use of a single sensing modality, such as a video camera or an infrared camera. A great deal of effort has been devoted to interpreting imagery sensed by each (single) modality separately. However, techniques which use a single modality work only in highly constrained environments and require enormous amounts of computational resources. The use of multiple sensing modalities and the development of "intelligent" algorithms to effectively combine these sensors can overcome the limitations of current approaches to machine vision. In this paper we explore the object detection and localization problem through images obtained from visual and infrared sensors.

This paper presents a framework for sensor fusion along with robust algorithms for the detection process. The framework follows the bottom-up approach and uses Bayesian statistics to account for uncertainties in the process. Multiple features, extracted either from single or multiple sensor images are used to model the object signature. Information from each feature is integrated for focused object

analysis.

The rest of this paper is organized as follows. In section 2, the need for sensor fusion and a modular framework are discussed, providing the basis for the development of the integration framework presented in the following section. In section 3, a Bayesian approach is presented where the problem is formulated as a two-class discrimination case. The theoretical foundations for decision fusion are discussed and the Fisher criterion is considered for determination of the optimal reliability factors. Experimental results obtained from both the visual imagery database and the FLIR imagery database are presented and analyzed in section 4. An example is also presented for detection in registered multisensor data. Finally, section 5 presents the conclusions of this paper.

2 Multisensor Fusion

Multisensor fusion is now widely accepted as being indispensable in vision applications, particularly when any one specific sensor is not guaranteed to provide complete discriminatory information because of the complexity of the scene, poor imaging conditions, or the effect of counter-measures or noise [6]. Multiple sensing modalities are used with great efficacy by several biological perceptual systems. The sensing modalities, as well as the manner in which the sensed signals are fused, are decided by the domain in which the systems function as well as by the application. Fusion of multiple sensor information for reliable analysis is a problem that has been studied in various areas over the years. In certain vision system problems, a complete analysis of a scene is not possible without information from multiple sensors. In such cases, the problem of multisensor fusion is defined as the integration of numerical and spatial sensory data to achieve useful information about an object or a scene that cannot be obtained from single sensor information [13]. This realization of the fundamental limitation of single sensor information has lead to an increasing interest in multisensor systems. Multiple sensor integration techniques studied so far can be broadly categorized in two classes.

1. Model-based approaches
2. Statistical approaches

Model-based techniques try to model the environment in which the system operates and are dependent on the physics of interactions within the scene and the sensor. Similar to physics-based analysis, the heat transfer within the environment is modeled and the radiation received at the thermal sensor is approximated. Similarly, the objects' reflectivity is determined by modeling the light source for visual sensor information. The information can once again be optimally fused. Statistical techniques, on the other hand, can be used to model the uncertainty of the sensor. This information, in turn, provides confidence of information. Bayes' decision criteria can be used to optimally combine information in such a case.

The motivation behind the design of a multisensor system stems from the realization that sensor measurements inherently incorporate varying degrees of uncertainty and are occasionally spurious and incorrect. Further, the spatial and physical limitations of sensor devices often mean that only *partial* information can be provided by a single sensor. Inspired by biological organisms, which are essentially multisensor perception systems, the development of intelligent systems that use multiple sources of information to extract knowledge about the sensed environment seems a natural step forward. The shortcomings of single sensors can be overcome by employing redundancy and diversity [9]. A multisensor fusion framework that integrates information at all processing levels would benefit from the principles of redundancy and diversity while managing the computational complexity. Such a modular framework, which uses Bayesian formulations to detect probable objects and present a coherent system with an associated confidence and error estimate at each level of the system, is proposed in figure 1. The lowest level of the system

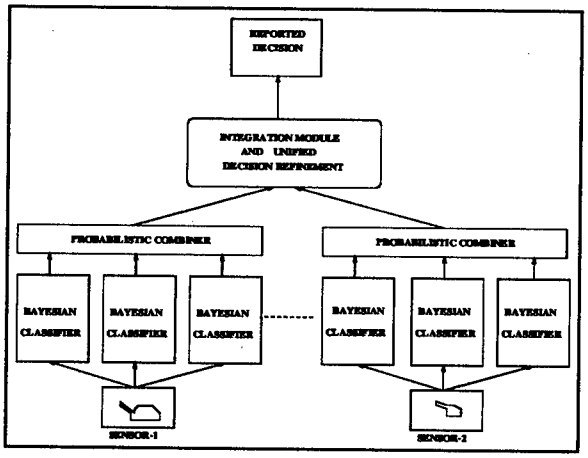


Figure 1: System module overview for information integration to achieve better detection/segmentation.

performs multiple feature extraction from images of each sensing modality. Multiple features or multiple sensors are treated the same in this framework. Several expert modules are used to characterize the distribution of these features, which aid in discriminating probable objects from the background. Detecting objects in both infrared and visual images poses a challenging problem due to enhanced clutter, the high degree of variations observed in natural environments, and faint object signatures. Individual modules are trained to identify varying object signatures. By incorporating different features, an object signature can be successfully identified to get an initial estimate of probable regions. A probabilistic combining stage ensures the fusion of multiple classifier outputs to maximize the object signature. In the event of registered multisensor data, the sensor fusion module incorporates a region refinement to obtain the final object segmentation by integrating object signature detected by individual sensor modules. In the absence of registered data, a region refinement stage can be implemented which analyses the detected region independent of information from other possible sources. To show the efficacy of the proposed framework, examples are provided for multisensor integration using range intensity and infrared images for the problem of object detection.

3 Information Integration

As improvements in sensors have been realized through the development of imaging technologies, the optical limits of sensor resolution have been reached. The second generation FLIR is a prime example of this. More advanced sensors will provide only small improvements over present capabilities [10]. In order to improve performance, we must exploit the use of available multiple sensors, as well as integration of spatial and temporal information. Data from more than one sensor can be fused by several techniques: information/data, pixel, feature, and decision-level fusion. Data fusion refers to the incorporation of object data from several sources, e.g., imaging sensor, object information, GPS, digital maps, etc. Pixel fusion involves the overlay of pixels from disparate sources to form an image. Feature fusion correlates information from two or more sources prior to making a decision. Decision fusion is a voting scheme in which each information source is polled as to the presence of objects.

In most data fusion systems, the information extracted from images or sensors is represented as measures of belief in an event. The information can be either numerical or symbolic. Its representation as numerical values leads to a quantification of the characteristics that have to be taken into account in a fusion process. Indeed, one of the main tasks of data fusion is to combine information issues from several sources to obtain a better decision than can be had from one source only, by reducing imprecision and uncertainty and increasing completeness [3]. In any object detection system, the events to which degrees of beliefs are assigned are related to the absence or presence of objects of interest. The degrees of beliefs are modeled in different ways, depending on the chosen mathematical framework, e.g., membership degrees to a fuzzy set in fuzzy set theory, necessity functions in possibility theory, mass, belief, or plausibility functions in Dempster-Shafer evidence theory, or probabilities in data fusion methods based on probability and Bayesian theory. When several pieces of information have to be combined, these degrees are combined in the form $F(x_1, x_2, \ldots, x_n)$, where x_i denotes the representation of information issues from source i. The question is: what information combination operator F should be chosen? It should be emphasized that the problem of data integration is very complex and that there are many issues begging explanation. These include the effect of individual expert error distributions on the choice of fusion strategy, explicit differentiation between decision ambiguity, competence and confidence, and the relationship between dimensionality reduction and multiple expert fusion, with its implicit dimensionality expansion.

3.1 Theoretical Framework for Fusion

The following discussion concentrates on the problem of statistical data integration, where a decision is to be made based on a group of experts. In the general m-class case, we consider that we have n experts each representing the given input/source, Z, by a distinct vector. Let x_i be the measurement vector, such as a feature extracted from the image, used by the i^{th} classifier. Each class ω_k is modeled by the probability density function $p(x_i|\omega_k)$ and its prior probability is denoted by $P(\omega_k)$. The models are assumed to be mutually exclusive as each model is associated with distinct measurement features/source. According to Bayesian theory, the input Z should be assigned to class ω_j provided the posterior probability of the interpretation is maxi-

mum, i.e.

$$\begin{aligned} & \textit{assign} \quad Z \rightarrow \omega_j \quad \textit{if} \qquad (1) \\ P(\omega_j|x_1,\ldots,x_n) \quad & = \quad \max_k P(\omega_k|x_1,\ldots,x_n) \end{aligned}$$

Rewriting the posterior probability $P(\omega_k|x_1,\ldots,x_n)$ using Bayes' theorem, we have

$$P(\omega_k|x_1,\ldots,x_n) = \frac{p(x_1,\ldots,x_n|\omega_k)P(\omega_k)}{p(x_1,\ldots,x_n)} \quad (2)$$

where $p(x_1,\ldots,x_n)$ is the unconditional joint probability density function. This can be expressed in terms of the conditional distribution as

$$p(x_1,\ldots,x_n) = \sum_{j=1}^{m} p(x_1,\ldots,x_n|\omega_j)P(\omega_j). \quad (3)$$

Assuming that the features/sources used by the classifiers are conditionally statistically independent, we can write the joint distribution as

$$p(x_1,\ldots,x_n|\omega_k) = \prod_{i=1}^{n} p(x_i|\omega_k). \quad (4)$$

Substituting from 4 and 3 into 2, we get

$$P(\omega_k|x_1,\ldots,x_n) = \frac{P(\omega_k)\prod_{i=1}^{n} p(x_i|\omega_k)}{\sum_{j=1}^{n} P(\omega_j)\prod_{i=1}^{n} p(x_i|\omega_j)} \quad (5)$$

Thus the decision rule can be given as:

$$\begin{aligned} & \textit{assign} \quad Z \rightarrow \omega_j \quad \textit{if} \qquad (6) \\ P(\omega_j)\prod_{i=1}^{n} p(x_i|\omega_j) \quad & = \quad \max_{k=1}^{m} P(\omega_k)\prod_{i=1}^{n} p(x_i|\omega_k) \end{aligned}$$

Under the assumption that the posterior probabilities computed by the respective classifiers will not deviate dramatically from the prior probabilities, the more commonly used sum rule is derived in [8]. Some of the other decision rules used are also discussed. They include the **max rule**, **min rule**, **median rule**, and the **majority vote rule**.

In deriving the decision rule above, it is conceded that the conditional independence assumption may be deemed unrealistic in many situations. However, for applications where the feature extractors are distinct, this assumption will hold. Further, this assumption will provide an adequate and workable approximation of reality, which may be more complex. Finally, most routinely used combining schemes are based on this assumption. For the sum rule, the assumption that the posterior class probabilities do not deviate greatly from the priors is unrealistic in most applications, and would introduce gross approximation errors. On the other hand, the product formulation has the drawback that a single recognition engine can inhibit the overall fusion by outputting a close to zero probability. For the special case of normally distributed assessments of the individual classifier outputs about the true class, the product formulation provides the final estimate by minimizing the variance over all inputs.

To avoid the zero probability problem, a measure of reliability for individual experts/classifiers can be considered in which the corresponding classifier contributes minimally to the final decision. In such a case, the modified product formulation is given by

$$p(x_1,\ldots,x_n|\omega_k) = \prod_{i=1}^{n} p(x_i|\omega_k)^{w_i}. \quad (7)$$

The introduction of weight factors, w_i, clearly reflects the expertise of individual classifiers, but it is not clear as to how they should be chosen. Rewriting equation 5, we have

$$P(\omega_k|x_1,\ldots,x_n) = \frac{P(\omega_k)\prod_{i=1}^{n} p(x_i|\omega_k)^{w_i}}{\sum_{j=1}^{n} P(\omega_j)\prod_{i=1}^{n} p(x_i|\omega_j)^{w_i}} \quad (8)$$

For a two-class discrimination, we can write the combined probability as

$$\begin{aligned} P(\omega_k|x_1,\ldots,x_n) \quad & = \quad \frac{P(\omega_k)[\prod_{i=1}^{n}\{\frac{p(\omega_k|x_i)}{P(\omega_k)}\}^{w_i}]}{A+B} \\ A \quad & = \quad P(\omega_j)[\prod_{i=1}^{n}\{\frac{p(\omega_k|x_i)}{P(\omega_k)}\}^{w_i}] \\ B \quad & = \quad P(\omega_j^c)[\prod_{i=1}^{n}\{\frac{p(\omega_k^c|x_i)}{P(\omega_k^c)}\}^{w_i^c}] \quad (9) \end{aligned}$$

A similar formulation has been shown in [1, 4] and is termed the logarithm opinion pool. The interpretation of this combined estimate is known to be unimodal and less dispersed than the linear combination of individual assessments. As p^w is a monotonic function of p, we can simplify equation 9 to the logarithmic form as (ignoring the normalizing denominator)

$$\log P(\omega_k|x_1,\ldots,x_n) = \quad (10)$$
$$\log P(\omega_k) + \sum_{i=1}^{n} w_i \log\{\frac{p(\omega_k|x_i)}{P(\omega_k)}\}$$

Considering the odds formulation where $O_G = \frac{P(\omega_k|x_1,\ldots,x_n)}{1-P(\omega_k|x_1,\ldots,x_n)}$, $O_i = \frac{p(\omega_k|x_i)}{1-p(\omega_k|x_i)}$ and $O_{\omega_k} = \frac{P(\omega_k)}{1-P(\omega_k)}$, we can rewrite the group estimate as

$$\log(\frac{O_G(\omega_k)}{O_{\omega_k}}) = \sum_{i=1}^{n} w_i \log(\frac{O_i}{O_{\omega_k}}) \quad (11)$$

3.2 Reliability Factors

To determine the weight factors for optimal discrimination, we consider the Fisher criterion. The reliability factors are computed based on the measure of within-class and between-class information, as introduced by Fisher [7]. The classifier outputs for a given class should be chosen such that they are clustered closely as compared to the outputs given for any other class. Once again, considering equation 11 or 11, let

$$\phi_i = \log\{\frac{p(\omega_k|x_i)}{P(\omega_k)}\}. \tag{12}$$

The group estimate can then be written as

$$y = \sum_{i=1}^{n} w_i \phi_i \tag{13}$$

where the decision is based on the maximum y_i for all classes. In the general case, let $\Omega = \{\omega_1, \ldots, \omega_k\}$ be the k classes into which the input Z is to be classified. Given the training patterns and the corresponding outputs, let X_ω be the set of input belonging to class ω and let N_ω be the number of patterns. The mean of class ω can then be defined as

$$\begin{aligned} \mu_\omega &= \frac{1}{N_\omega} \sum_{x \in X_\omega} y = \frac{1}{N_\omega} \sum_{x \in X_\omega} W^T \phi \\ &= W^T (\frac{1}{N_\omega} \sum_{x \in X_\omega} \phi) = W^T \hat{\phi_\omega} \end{aligned} \tag{14}$$

where $\hat{\phi_\omega}$ is a $n \times 1$ column vector given by:

$$\hat{\phi_\omega} = \frac{1}{N_\omega} \sum_{x \in X_\omega} \phi. \tag{15}$$

Similarly, the mean over all classes can be defined by

$$\begin{aligned} \mu &= \frac{1}{N} \sum_{x \in X} y = \frac{1}{N} \sum_{x \in X} W^T \phi \\ &= W^T (\frac{1}{N} \sum_{x \in X} \phi) = W^T \hat{\phi} \end{aligned} \tag{16}$$

where $\hat{\phi}$ is again the $n \times 1$ column vector given by

$$\hat{\phi} = \frac{1}{N} \sum_{x \in X} \phi = \frac{1}{N} \sum_{\omega \in \Omega} N_\omega \hat{\phi_\omega}. \tag{17}$$

The between-class variance is then defined as

$$S_B(W) = \sum_{\omega \in \Omega} N_\omega (\mu - \mu_\omega)^2 \tag{18}$$

Using equation 17,

$$\sum_{\omega \in \Omega} N_\omega (W^T \hat{\phi_\omega}) = N(W^T \hat{\phi}) \tag{19}$$

the between-class variance can be simplified as

$$\begin{aligned} S_B(W) &= \sum_{\omega \in \Omega} N_\omega (W^T \hat{\phi_\omega})^2 \\ &- N(W^T \hat{\phi})^2 \end{aligned} \tag{20}$$

The within-class variance is given as

$$S_W(W) = \sum_{\omega \in \Omega} \sum_{x \in X_\omega} (\mu_\omega - y)^2 \tag{21}$$

$$= \sum_{\omega \in \Omega} \sum_{x \in X_\omega} (W^T \hat{\phi_\omega} - W^T \phi)^2 \tag{22}$$

Simplifying, we get

$$\begin{aligned} S_W(W) &= \sum_{x \in X} (W^T \phi)^2 \\ &- \sum_{\omega \in \Omega} N_\omega (W^T \hat{\phi_\omega})^2 \end{aligned} \tag{23}$$

In trying to minimize the within-class variance and maximize the between-class variance, the selection criterion can be given as

$$J(W) = S_B(W) - \alpha[S_W(W) - 2W^T I] \tag{24}$$

where I is an $n \times 1$ column vector of ones. α weights the within-class variance with respect to the between-class variance and acts as the regularization parameter. Differentiation with respect to the weight factors and setting the derivative to zero, we get

$$\frac{\partial J}{\partial W} = \frac{\partial S_B(W)}{\partial W} - \alpha \frac{\partial S_W(W)}{\partial W} - 2\alpha I \tag{25}$$

Using the identity $\frac{\partial (W^T I)}{\partial W} = I$, we have

$$\frac{\partial S_B(W)}{\partial W} = 2 \sum_{\omega \in \Omega} N_\omega (W^T \hat{\phi_\omega}) \hat{\phi_\omega} - 2N(W^T \hat{\phi}) \hat{\phi} \tag{26}$$

and

$$\frac{\partial S_W(W)}{\partial W} = 2 \sum_{x \in X} (W^T \phi) \phi - 2 \sum_{\omega \in \Omega} N_\omega (W^T \hat{\phi_\omega}) \hat{\phi_\omega} \tag{27}$$

Simplifying equation 25

$$\frac{1}{2} \frac{\partial J}{\partial W} = \alpha I \tag{28}$$

Thus the optimal solution for the reliability factors **W** is given by

$$\mathbf{W} = \alpha((1+\alpha)\sum_{\omega\in\Omega} N_\omega(\hat{\phi_\omega}^T\hat{\phi_\omega}) - N(\hat{\phi}^T\hat{\phi}) - \alpha\sum_{x\in X}(\phi^T\phi))^{-1}I \quad (29)$$

4 Experimental Results

The features used for robust detection of objects in complex scenes included color, regularity, local deviation, and homogeneity in images [12]. The data was divided into two distinct sets, one for validating the system and the other for testing the system. The **validation set** for both the Colorado(visual) and Comanche(FLIR) data consisted of 40 images under varying background conditions. Some of the images had high clutter and very small objects. A separate set of 5 images was used to compute the image statistics and estimate the object signature distribution parameters using the modified EM algorithm [11]. The **testing set** for the Colorado data consisted of 65 images of similar objects, once again in varying environmental settings and ambient conditions. For the Comanche data, a set of 144 images were used from three distinct sites in varying environmental conditions. The **validation set** was used to validate the classifier design and determine the confidence/reliability of individual classifiers.

The detected regions were individually analyzed by performing a connected component analysis on the output of pixel classification and regions consisting of less than 100 pixels were removed. A region growing procedure with compactness and edge linearity constraints [11] was used to isolate final regions. An example of a typical visual image and the detected object is shown in figure 2. Figure 3 shows

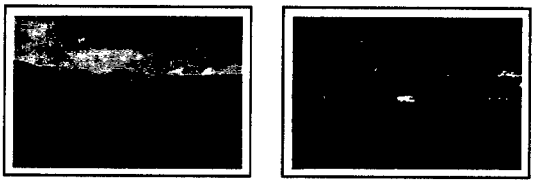


Figure 2: Typical visual input and detected object.

an example of a typical FLIR input image and the result of detection and segmentation.

Based on the computed reliability factors, the classification rate of 96.59% and a false alarm of

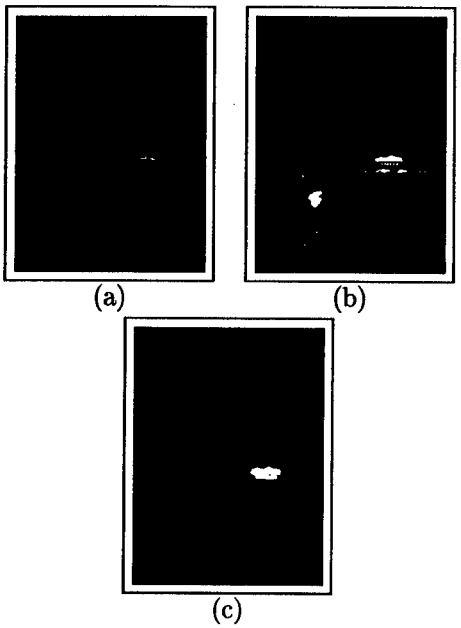


Figure 3: Preprocessing steps applied to typical FLIR image (a), the results obtained after initial detection (b), and the result of final segmentation (c).

4.3% was obtained with a threshold of 0.7 in the visual dataset. For images from the Comanche FLIR dataset, the best results were obtained with a total of 143 regions detected, an object classification rate of 99.3% and a false alarm of 1.9%.

In another experiment, the lower level of individual classifiers was merged into one classifier and the decision integration module was removed. Thus, the features were concatenated to give just a single feature vector. The same experiments were then repeated. This was done to verify the advantage of the proposed methodology. The overall detection rate in this case dropped to 81.5% for the visual data and 83.2% for the FLIR data with a threshold of 0.7.

To further evaluate the developed methodology, we compare the results with those obtained by using some of the existing classification and combining schemes. As variants for the integration scheme, we consider the *Sum Rule* [8] which gives

$$P(\omega_k|X) = \sum_i^n P_i(\omega_k|x_i) \quad (30)$$

A modification of that is the weighted sum rule, which is simply

$$P(\omega_k|X) = \sum_i^n w_i P_i(\omega_k|x_i) \quad (31)$$

The weights in these experiments were determined by cross validation on the validation set for each classifier. The third variant would be just to consider a majority vote, which is

$$P(\omega_k|X) = \max_i^n P_i(\omega_k|x_i) \qquad (32)$$

Standard classification algorithms were also considered. A Nearest Neighbor(NN) algorithm (1-NN and 3-NN) was used to estimate the final binary decision. The NN classifier assigns a test pattern x to the same class ω_i as the training pattern $x_i^p \in \chi_i$ nearest to x in the feature space. χ_i is the set of all training patterns belonging to the class ω_i. Given in the discriminant function form, the classification can be given as:

$$\delta_{NN}(x) = \alpha_i \ni g_i(x) \geq g_j(x) \ \forall j \neq i \qquad (33)$$

where α_i is the assigned label and $g_i(x)$ is the distance measure and can be written as

$$g_i(x) = -\|x - x_i^p\| \qquad (34)$$

and

$$\|x - x_i^p\| \leq \|x - x_j^p\| \ \forall x_i^p \in \chi_i, i \neq j \qquad (35)$$

The decision tree (C4.5) and one level rule induction (1-R) algorithms were also used. In all of these algorithms, the training patterns were the posterior estimates obtained from individual classifiers. Tables 1 and 2 present the average results over 20 runs for the same visual and FLIR images respectively, as used in earlier experiments.

Algorithm	3000 Data Pts.
Our Algorithm	96.59%
Sum Rule	83.6%
Wtd. Sum Rule	91.5%
Maj. Vote	76.8%
1-NN	88.2%
3-NN	89.2%
C4.5	85.1%
1-R	75.9%

Table 1: Comparison of combining schemes for visual data.

In addition, to demonstrate the generalizability of the developed fusion framework, an example of multisensor detection is considered. The images used in these examples were obtained by a FLIR and range sensor. Three of the features introduced earlier, excluding color, were used on the FLIR and range intensity images to model the objects of interest and the background. A validation set consisting of non-registered images was used to determine the weight factor for the integration of features for each sensor. The final sensor integration was performed by considering equal weight contributions. In the presense of registered dataset, the weight factor contributions for sensor integration can easily be computed as discussed in section 3.1. For the two examples shown here, the registration was performed manually to maximize pixel overlap between the two sensor images. Figure 4(a) and (b) show the input images, (c) and (d) show the detection after multifeature integration, and (e) shows the result after sensor integration.

Algorithm	3000 Data Pts.
Our Algorithm	99.3%
Sum Rule	87.5%
Wtd. Sum Rule	93.0%
Maj. Vote	77.7%
1-NN	90.9%
3-NN	92.3%
C4.5	84.0%
1-R	84.7%

Table 2: Comparison of combining schemes for FLIR data.

5 Summary

In this paper, we have presented a methodology for object region localization/detection. Multiple image statistics are independently computed based on generic measures in visual computation. We introduce a modular computational structure consisting of multiple classifiers, each of which attempts to solve the global problem based on its input observations. A higher level decision integrator oversees and collects evidence from each of the individual modules and combines it to provide a final decision while considering the redundancy and diversity of individual classifiers. A Bayesian realization of the methodology is presented. Each classifier module models the object signature probability density function based on the computed image statistics and the final integration is achieved in a supra-Bayesian scheme. How the object models benefit the object detection process is demonstrated by the

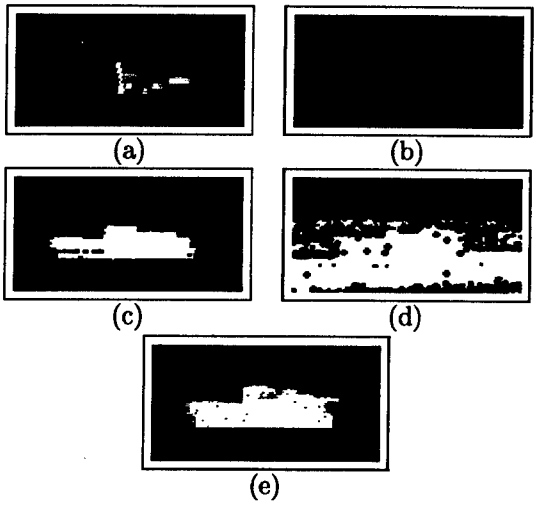


Figure 4: Registered multisensor (a) FLIR and (b) Range intensity images, (c) and (d) detected object after feature integration, and (e) final detection after sensor integration.

effectiveness of the computed image features and the better understanding of each feature's validity with respect to contextual parameters that the use of object models provides. The results presented here were obtained on images from the Fort Carson Colorado dataset and the Comanche FLIR dataset. The results are compared to results obtained by using some of the existing pattern recognition techniques. It is seen that considerable improvement is obtained through the use of the methodology presented in this chapter. This methodology can be extended to multiple sensors if the input data is registered. An example of multisensor detection is presented which shows the efficacy of the proposed framework. The framework allows for the use of maximal information, thus the performance would improve as more discriminatory information is added through extra classifiers.

References

[1] J. A. Benediktsson and P. H. Swain. Consensus theoretic classification methods. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(4):688–704, 1992.

[2] I. Biederman. Human image understanding: Recent research and a theory. *Computer Vision, Graphics and Image Processing*, 32:29–73, 1985.

[3] I. Bloch. Information combination operators for data fusion: A comparative review with classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 26(1):52–67, 1996.

[4] R. F. Bordley. A multiplicative formula for aggregating probability assessments. *Management Science*, 28(10):1137–1148, 1982.

[5] T. Caelli, M. Johnston, and T. Robinson. 3D object recognition: Inspiration and lessons from biological vision. In A. K. Jain and P. J. Flynn, editors, *Three-Dimensional Object Recognition Systems*, pages 1–16. Elsevier Science Publishers, 1993.

[6] C. C. Chu and J. K. Aggarwal. Multisensor fusion for computer vision. In J. K. Aggarwal, editor, *The Issues, Analysis and Interpretation of Multisensor Images*, pages 37–62. Springer Verlag, 1993.

[7] R. A. Fisher. The use of multiple measurements in taxonomical problems. *Annals of Eugenics*, 2:179–188, 1936.

[8] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):226–239, 1998.

[9] J. Manyika and H. F. Durrant-Whyte. *Data Fusion and Sensor Management: A decentralized information-theoretic approach.* Ellis Horwood, 1994.

[10] J. A. Ratches, C. P. Walters, R. G. Buser, and G. D. Guenther. Aided and automatic target recognition based upon sensory input from image forming systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(9):1004–1019, 1997.

[11] Shishir Shah and J. K. Aggarwal. A Bayesian segmentation framework for textured visual images. In *Proc. of Computer Vision and Pattern Recognition*, pages 1014–1020, 1997.

[12] Shishir Shah and J. K. Aggarwal. Multiple feature integration for robust object localization. In *Proc. Computer Vision and Pattern Recognition*, Santa Barbara, CA, 1998.

[13] Shishir Shah, J. Eledath, J. Ghosh, and J. K. Aggarwal. Multisensor fusion for scene classification: An experiment in human form detection. In *Proceedings of International Conference on Image Processing*, pages 199–202, 1997.

Session WB4
Fusion for Target Tracking III
Chairs: X. Rong Li University of New Orleans, USA
T. Kirubarajan, University of Connecticut, CT, USA

Track Fusion Algorithms in Decentralized Tracking Systems With Feedback in a Fighter Aircraft Application

Mathias Karlsson, Anders Malmberg, Thomas Jensen and Leif Axelsson
SAAB AB, Gripen
Decision and Mission Support
SE-581 88 Linköping, Sweden
firstname.surname@saab.se

Abstract

In a decentralized tracking system, sensor tracks are fused into central tracks, which will benefit from all the advantages of the sensor tracks. The performance of the system could be improved if information from the central node is fed back to the sensor nodes. However, the problem of cross-correlation between the sensor tracks arises and must be handled.

In this paper, three different track fusion techniques, for sensor level and central level fusion, are investigated. The methods to handle the cross-correlations are; calculating decorrelated state estimates, using Covariance Intersection (CI) and enlarging the covariance matrix to compensate for cross-correlation. The fusion algorithms have been tested in a fighter aircraft application, consisting of one radar sensor and one angle-only measuring Infrared Search and Track (IRST) sensor with local Kalman Filters in a hierarchical tracking system with feedback.

The conclusion of this paper is that, in order to get a functioning decentralized tracking system with feedback, the cross-correlation has to be taken into account. The decorrelated state estimate method and the Covariance Intersection method manages to do this with about the same performance, whereas the third method fails to handle the cross-correlation.

Also, the effect of different measurement and feedback rates are investigated.

Keywords: Track fusion, decentralized tracking, hierarchical tracking system, track feedback, decorrelation, Covariance Intersection

1. Introduction

In air combat, information advantage over the opponent is vital for the success of the operation. For that reason, modern fighter aircraft have extensive sensor suites to track other objects. The multitude of sensors makes it impossible for the pilot to use each individual sensor efficiently without some sort of data processing, i.e. sensor data fusion, and a sensor manager [1]. Since system modularity and high computational performance are needed in a fighter aircraft application, a decentralized tracking approach is preferable.

In order to form a unified picture of the vicinity, all sensor information is fused into a central track. The fused track will then benefit from all advantages of the sensor estimates.

The tracking performance of each individual sensor could be increased if information is shared between the sensors via track feedback. The problem that arises is that tracks from different sensors may be based on common data. If the resulting cross-correlations are not properly handled, the statistical properties of the tracks could be ruined.

2. The system architecture

A hierarchical decentralized tracking approach (see figure 1) is based on that data (measurements) are preprocessed in the sensor nodes, using e.g. Kalman Filters. The sensor nodes will then provide a central node with the preprocessed data, which are invoked in a central fusion process. The fusion process in the central node includes all sensor data and therefore that node has the best possible conception of the vicinity.

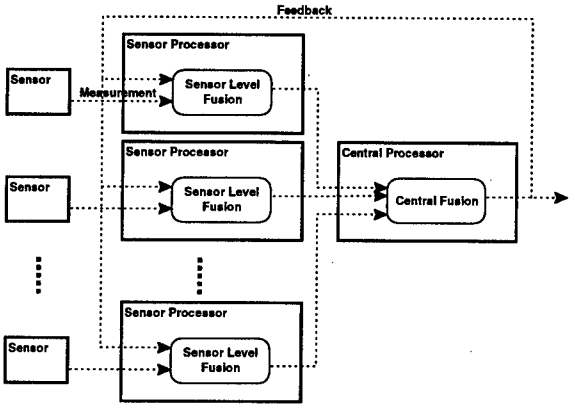


Figure 1. Architecture of the decentralized tracking system with feedback.

If there is no track feedback to the sensor trackers, the only available data in the sensor nodes are the measurements that the sensor has made itself. This could be a problem if the data (measurement) rate is low and the sensor track accuracy thereby is poor.

Therefore, it would be preferred that the sensor trackers also had information from the other sensor trackers. If so, the sensor track would not degenerate too much between measurement updates, even though

measurements are made very seldom. This could make tracking in the sensor nodes easier and more reliable.

In a hierarchical system, information sharing is accomplished by feeding the central tracks back to the sensor trackers. Of course, this means that there have to be fusion processes in the sensor nodes as well, invoking the fed back tracks into the sensor tracks. Also, since data are transferred in both directions the problem of using the same information several times has to be taken care of in the fusion algorithms in the central as well as the sensor nodes. However, it is possible to deal with this problem in a hierachical network. Investigations of other distributed architectures can be found in [2] and [3].

3. Fusion Algorithms

To achieve the desired track accuracy and statistical properties of the track, a fusion algorithm has to be employed both in the sensor nodes and the central node of a decentralized tracking system with track feedback.

Since data are distributed in the system, the problem of double counting information arises. If this problem is not properly handled, the statistical properties and the quality of the track could be ruined.

Three different fusion algorithms, denoted Filter A, B and C, are described in the following sections. The filters can be used in any of the nodes in the system. A fusion node is shown in figure 2. The remote tracks to be fused are denoted (x_i, P_i) and the fused tracks, in the node where the fusion takes place, are denoted (x_{tot}, P_{tot}).

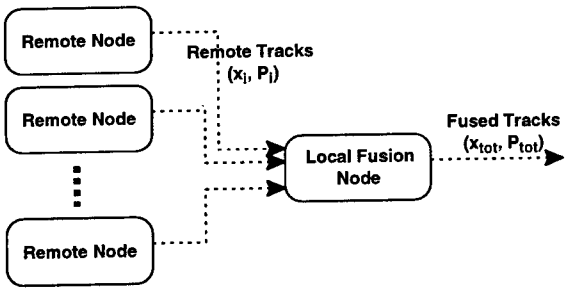


Figure 2. A fusion node in the decentralized tracking system.

3.1 Filter A

The filter fuses the estimates in the fusion node with the estimates reported from other (remote) nodes. Since the estimates reported from one node at different times are created by partly the same data, they are highly correlated. To deal with this problem, the old information (i.e. the inverse of the predicted covariance matrix times the predicted state estimate) is subtracted from the reported new information [4]. In this way decorrelated state estimates are formed out of the remote tracks. The decorrelated tracks are then merged into fused tracks. The update equations are

$$P_{tot}^{-1}(k|k) = P_{tot}^{-1}(k|k-1) + \sum_i P_i^{-1}(k|k) - P_i^{-1}(k|k-1) \qquad \text{Eq. 1}$$

$$\hat{x}_{tot}(k|k) = P_{tot}(k|k) \cdot \left\{ P_{tot}^{-1}(k|k-1) \cdot \hat{x}_{tot}(k|k-1) + \sum_i \left\{ P_i^{-1}(k|k) \cdot \hat{x}_i(k|k) - P_i^{-1}(k|k-1) \cdot \hat{x}_i(k|k-1) \right\} \right\} \qquad \text{Eq. 2}$$

A system without feedback, using this fusion algorithm in the central node, is a decentralized Kalman Filter, which can be derived from a centralized Kalman Filter. It is optimal in the Kalman sense if each sensor has independent measurement noise and the kinematic models are linear.

The term $\hat{x}_i(k|k-1)$ has to be predicted in the fusion node. It is important to use the same kinematic model in both the fusion node and in the remote nodes. Otherwise, it may lead to numerical problems in the filter. Another issue is that different remote trackers could use different kinematic models and therefore cause problems in the fusion node.

3.2 Filter B

This fusion filter is a simplified version of Filter A. It is based on the assumption that the tracks to be fused are independent stochastic variables. This is an unrealistic assumption in a system with feedback. However, if no track feedback is used and the tracks origin from totally different information sources or are sufficiently separated in time, the assumption of independent tracks is a realistic approximation. The expressions for updating the covariance matrix and state vector are

$$P_{tot}^{-1}(k|k) = \sum_i P_i^{-1}(k|k) \qquad \text{Eq. 3}$$

$$\hat{x}_{tot}(k|k) = P_{tot}(k|k) \cdot \sum_i P_i^{-1}(k|k) \cdot \hat{x}_i(k|k) \qquad \text{Eq. 4}$$

To compensate for correlations between object tracks and thus maintain consistency, the resulting covariance matrix, P_{tot}, could be enlarged by multiplying it by a constant. In the simulations presented in this paper, the multiplier 1.3 is used.

3.3 Filter C

This fusion filter is, unlike the other filters, not dependent on whether the tracks to be fused are independent or not. The resulting covariance matrix is a convex combination of the information matrices of the tracks to be fused. The convex combination guarantees that the true covariance matrix, regardless of what the correlation is, lies within the resulting covariance matrix. In existing literature, this method of fusing sensor tracks are referred to as *Covariance Intersection* (CI) [5].

The expressions for updating the covariance matrix and state vector are as follows:

$$P_{tot}^{-1}(k|k) = \sum_i \omega_i \cdot P_i^{-1}(k|k) \qquad \textit{Eq. 5}$$

$$\sum_i \omega_i = 1 \qquad \textit{Eq. 6}$$

$$\hat{x}_{tot}(k|k) = P_{tot}(k|k) \cdot \sum_i \omega_i \cdot P_i^{-1}(k|k) \cdot \hat{x}_i(k|k) \qquad \textit{Eq. 7}$$

The weights, ω_i, are chosen by minimizing some norm of the resulting covariance matrix, P_{tot}. Depending on the norm that is chosen to be minimized, this could be more or less time consuming for the system. Consequently, it is very important to pick a norm that yields sufficient tracking quality, while not being too time consuming to minimize. Two examples of norms that could be used are the trace or determinant of P_{tot}. In the simulations in this paper, the determinant norm is used. For a more thorough investigation of Covariance Intersection, see [5].

4. Evaluation

In this section, prerequisites of the simulations are formulated, parameters of the implemented system are described and some *Measures of Performance* (MOP) are stated.

4.1 Prerequisites

The simulations were made in MatLab5 on a 166 MHz Pentium PC. In the simulation model, no association was conducted. All associations were assumed to be correct.

In the simulations, one radar and one IRST sensor were modeled. Plots and data are the result of 20 runs Monte Carlo simulations.

4.2 Sensor Tracking Filters

The radar node was modeled by an *Extended Kalman Filter* (EKF) in Cartesian Coordinates. Standard deviations for the measurement model were; σ_d = 20 m (distance), σ_b = 10 mrad (bearing) and σ_e = 30 mrad (elevation). The process noise in the filter had a standard deviation of 30 m/s^2 in all directions of a Cartesian Coordinate system. The process noise models object accelerations up to 3g. The prediction interval was 1/3 s and the measurement interval was 5 s if nothing else is stated.

The IRST node was modeled by an EKF in *Modified Spherical Coordinates* (MSC) [6]. Standard deviations for the measurement model were; σ_b = 0.5 mrad (bearing) and σ_e = 0.5 mrad (elevation). The process noise in the filter had a standard deviation of 50 m/s^2 in all directions of a Cartesian Coordinate system. The prediction interval was 1/9 s and the measurement interval was 2 s.

Both the radar and IRST tracking Kalman Filters are supported by track feedback every 2 s in the simulations with feedback, where nothing else is stated.

As a comparison, a central Kalman Filter was modeled. All measurements were reported directly to it and invoked into the filter. The filter was an EKF in Cartesian Coordinates with the same parameters as the radar filter.

4.3 Simulation Scenario

In all simulations, the scenario in figure 3 was used. The own aircraft starts at the bottom of the figure and flies more or less to the north. It is heading for the enemy aircraft, which is flying westward. Both aircraft trajectories start in the point highlighted by a +. The duration of the scenario is 300 s.The aircraft in the scenario are maneuvering between 0 and 3g which makes their maneuvering quite realistic. In figure 4, the accelerations of the target and the own aircraft is plotted to give a survey of the maneuvers in the scenario.

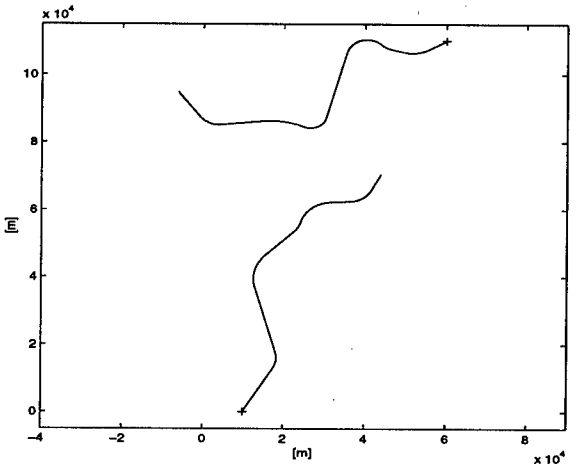


Figure 3. Scenario used in the evaluation of the fusion algorithms. The own aircraft flies northward and the enemy flies westward.

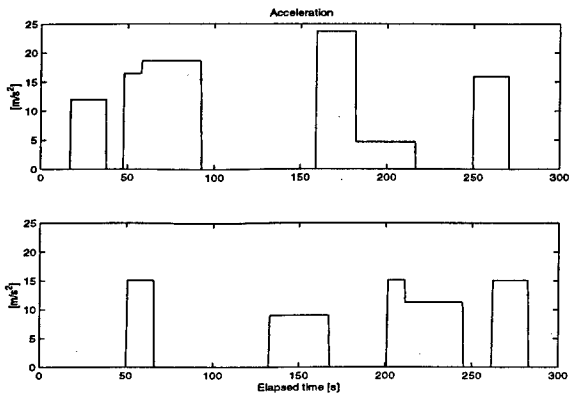


Figure 4. Accelerations of the target (above) and the own aircraft (below) in the scenario.

4.4 Measures of Performance

In this paper, two *Measures of Performance* (MOP) are considered; track deviation and track uncertainty consistency. Both the deviation in position and velocity are used. The deviations are defined as the differences between the estimated and the true values;

$$\delta_i = \left| \hat{x}_i - \bar{x}_i \right| \qquad \textit{Eq. 8}$$

$$\Delta v_i = \left| \hat{v}_i - \bar{v}_i \right| \qquad \textit{Eq. 9}$$

A way to decide if a tracking algorithm is performing well, is to check if the error of the estimated state vector is statistically corresponding to the covariance matrix. Consistency may be checked using standard hypothesis testing techniques. The measure used to test consistency is the $\chi^2(6)$ distributed variable, $\varepsilon(k)$, the *normalized state error* squared;

$$\tilde{x}(k|k) = x(k) - \hat{x}(k|k) \qquad \textit{Eq. 10}$$

$$\varepsilon(k) = \tilde{x}^T(k|k) P^{-1}(k|k) \tilde{x}(k|k) \qquad \textit{Eq. 11}$$

The statistical 95% interval that $\varepsilon(k)$ should stay within are further on presented as two lines in the consistency plots. For a thorough investigation on consistency and the theory behind it, see [7].

5. Results

In this section, simulation results are presented and comparisons between the different algorithms are made. The central fusion is performed in MSC. However, it is shown in 5.2 that Cartesian fusion is possible.

In all simulations in this paper, the measurement frequencies are 0.2 Hz for the RR and 0.5 Hz for the IRST. The feedback rate is 0.5 Hz if no other value is stated.

It is important to make sure that the quality of the central track is not degraded as a result of track feedback to the sensor tracking filters. Consequently, an exploration of the effects of track feedback on the central track has to be made. Different parameter settings such as measurement interval and feedback interval must be studied.

5.1 Comparison Simulations

In figure 5, the resulting track quality of a single radar with measurement frequency 0.2 Hz is presented.

When the single radar is used, the consistency is not very good and the position and velocity deviations are far too large to be acceptable.

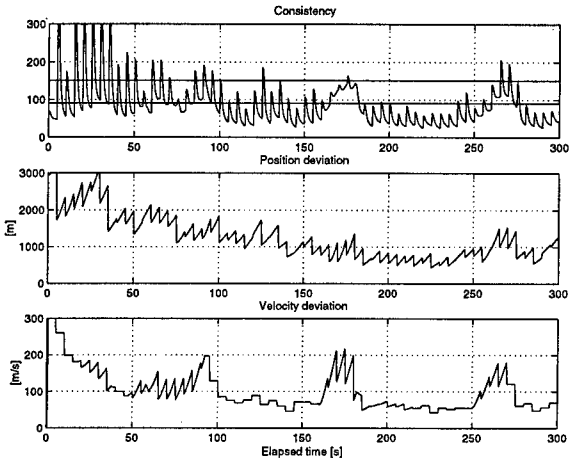


Figure 5. Consistency and track deviations for a single radar, using no multisensor features.

In figure 6, the performance of a central *Extended Kalman Filter* (EKF) that is fusing both the radar and the IRST measurements is shown. These simulations are used as a comparison to the filter configurations tested in the following sections.

It is interesting to observe that the covariances are overestimated in the central EKF (the consistency measure ε lies below the statistical interval). This should be kept in mind when studying the results of Filter A, B and C further on in the paper. It can also be seen that the errors in the position and velocity estimates are considerably lower than for the single radar case. The periods with lower tracking quality (in the periods 50-90 s and 160-180 s) are caused by the fact that the target is maneuvering (compare to figure 4) and it is consequently more difficult to track the target during these periods.

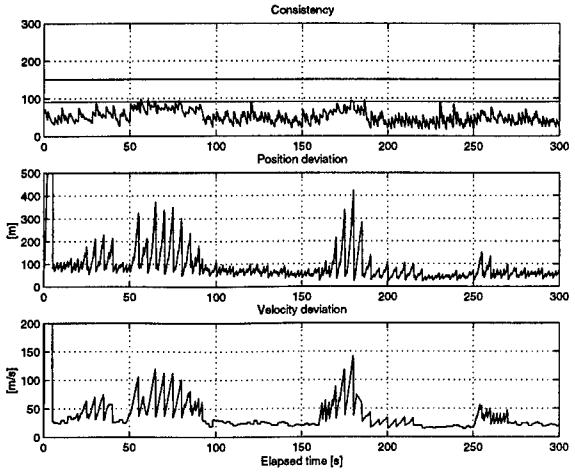


Figure 6. Consistency and track deviations for a centralized EKF.

As another comparison, the performance of a decentralized system without feedback, using Filter A in the central node, is shown in figure 7. It shows that this filter has about the same kinematic deviations as the central Kalman Filter. However, the consistency is not quite as good as for the central Kalman Filter.

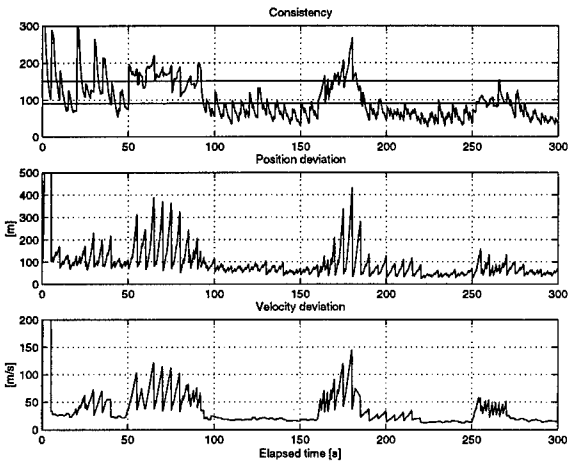


Figure 7. Consistency and track deviations for a decentralized tracking system using Filter A without feedback.

5.2 Filter A/Filter A with Feedback

In figure 8, the results of simulations of a system with Filter A in both the sensor nodes and the central node are shown. The MOP:s show that the filter is consistent and has position and velocity deviations comparable to the deviations of the decentralized filter without feedback in figure 7. The consistency is slightly better in the beginning of the scenario than for the filter without feedback.

Compared to the central Kalman Filter in figure 6, a feedback system using Filter A results in the same track deviations. The difference is that the consistency is not quite as good, especially when the target is maneuvering.

The conclusion is that Filter A gives a fully functioning decentralized system with feedback.

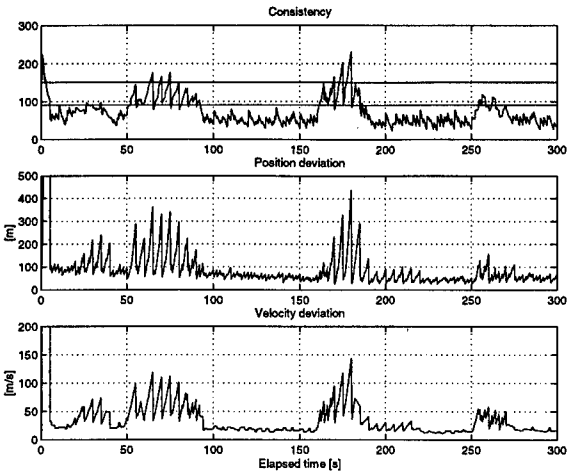


Figure 8. Results of simulations with Filter A in both the sensor nodes and the central node.

In figure 9, the performance of the decentralized tracking system when Cartesian fusion is used in the central node is exhibited. This will serve as the example that the tracking system works in both MSC and Cartesian Coordinates. The result of the Cartesian fusion does not significantly differ from the MSC fusion. Therefore, further results in this paper will be based on simulations where MSC is used in the central fusion.

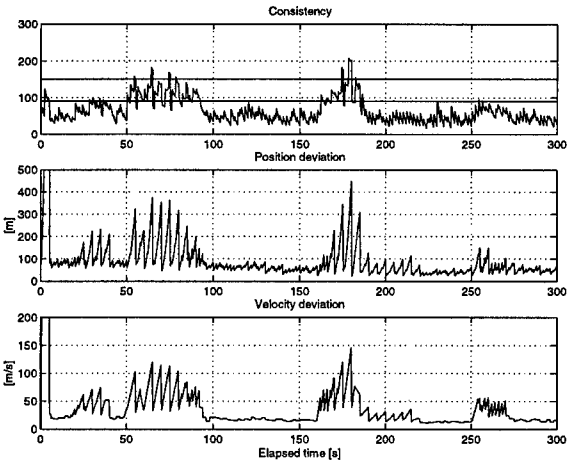


Figure 9. Results of simulations with Filter A in both the sensor nodes and the central node and Cartesian fusion in the central node.

5.3 Filter B/Filter B with Feedback

In figure 10, the results of simulations of a system with Filter B in both the sensor nodes and the central node are shown. In the consistency plot, it can be seen that the filters do not manage to handle the cross-correlation problem and maintain the consistency.

The deviations are not disastrous, but a filter with these properties is totally malfunctioning due to the underestimation of the uncertainty.

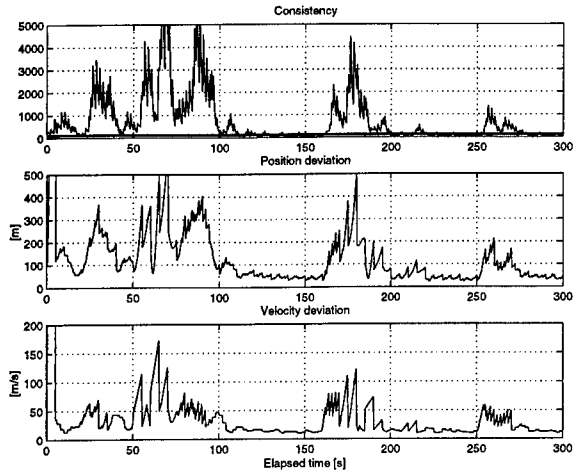


Figure 10. Results of simulations with Filter B in both the sensor nodes and the central node.

5.4 Filter C/Filter C with Feedback

In figure 11, the results of simulations of a system with Filter C in both the sensor nodes and the central node are shown. The filter handles the cross-correlation problem and is fully functional. It manages to maintain the same level of consistency as Filter A in figure 8. The track deviations are also comparable.

The system has slightly better consistency than the system without feedback in figure 7, but the track deviations are about the same.

The conclusion is that Filter C also gives a fully functioning decentralized system with feedback with performance similar to Filter A.

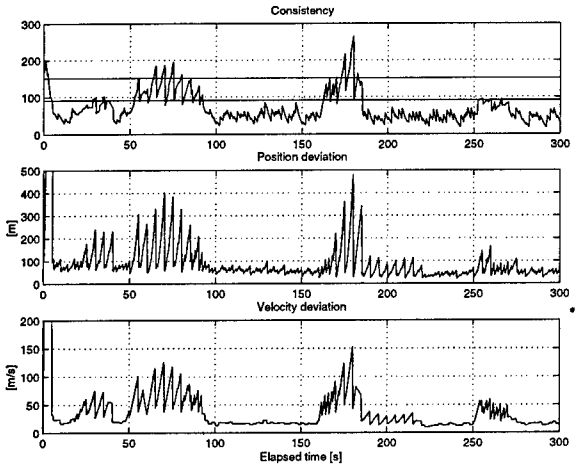


Figure 11. Result of simulations with Filter C in both the sensor nodes and the central node.

5.5 Other System Configurations with Feedback

Other combinations of the fusion filters than those that are shown in 5.2 to 5.4 have also been tested. The conclusion is that all combinations of Filter A and Filter C result in systems with performance similar to the systems in 5.2 and 5.4. All combinations involving Filter B fails to produce consistent track estimates in accordance with the system in 5.3.

5.6 The Effect of Different Measurement Rates

In all simulations in 5.2 to 5.4, the radar measures with 0.2 Hz and the IRST measures with 0.5 Hz. The results of simulations of the system with Filter A in both the sensor nodes and central node with different measurement frequencies are shown in figure 12 and figure 13.

In figure 12, the results of simulations with the radar measuring with 0.1 Hz and the IRST measuring with 1 Hz are shown. Compared to figure 8, we can see that the tracking quality is worse due to the lower radar measurement frequency. The position deviation increases considerably in the intervals between the radar measurements. The quality is much worse when the target is maneuvering and the angle-only measuring IRST has problems with estimating the distance to the target. However, the consistency is not ruined, so the decreased radar measurement rate does not ruin the functionality of the decentralized tracking system.

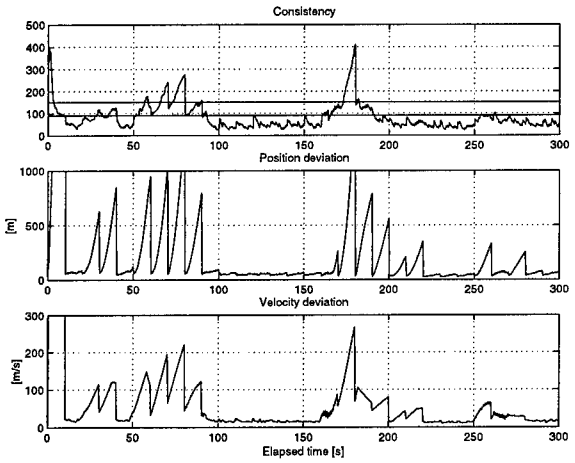


Figure 12. Consistency and track deviations when Filter A is used and the measurement intervals are 10 s for radar and 1 s for IRST.

In figure 13, the results of simulations with both sensors measuring with 1 Hz are presented. Compared to figure 8 and figure 12, we can see that the tracking quality is much better. Since the radar is measuring the distance frequently, the position deviation does not increase that much between the measurements. There are no problems tracking the target when it is

maneuvering and the consistency is very good at all times.

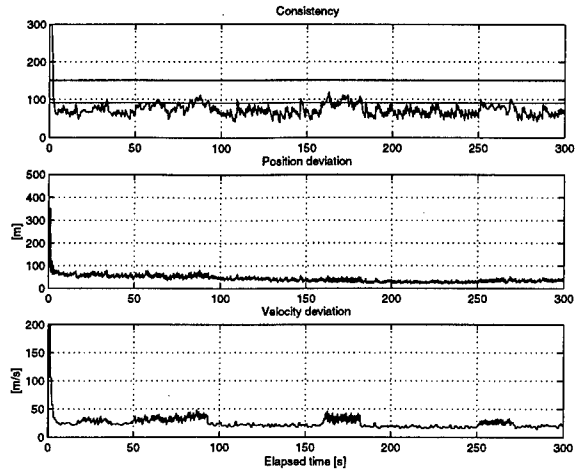


Figure 13. Consistency and track deviations when Filter A is used and the measurement intervals are 1 s for radar and 1 s for IRST.

5.7 The Effect of Different Track Feedback Rates

In all simulations in 5.2 to 5.4, the track feedback rate for both radar and IRST is 0.5 Hz. The results of simulations of the system with Filter A in both the sensor nodes and central node with feedback frequency 0.2 Hz are shown in figure 14.

The performance shown in figure 14 should be compared to the results in figure 8. There are no noticeable effect on the central track of changing the feedback interval from 0.5 Hz to 0.2 Hz. The reason for this is that the information in the central node is not decreased as a result of a decreased feedback frequency. This is also affirmed by the fact that a system without feedback has about the same performance, see figure 7.

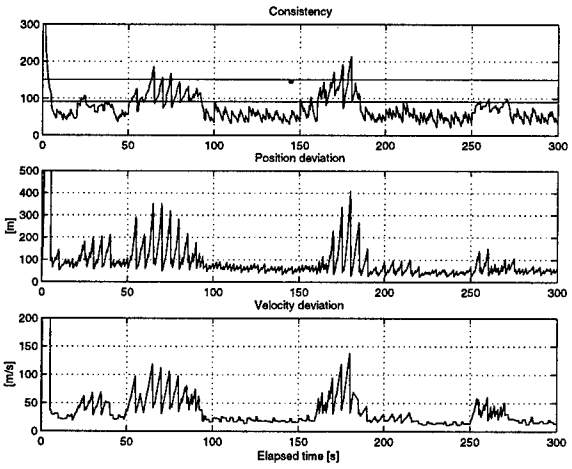


Figure 14. Consistency and track deviations when Filter A is used and the feedback intervals are 5 s for both radar and IRST.

5.8 Benefits of Feedback in the Sensor Tracking Filters

Direct improvements of track accuracy in the central node are not to be expected as a result of track feedback, since the information in the system is not increased, rather distributed. The real improvements can be found in the sensor nodes, which will have access to more information as a result of the track feedback. In figure 15, the track quality achieved in the radar node with track feedback is exhibited.

There are almost no differences between the track quality in figure 15 and the track quality of the central fusion in figure 8. The reason is that when track feedback is used, all nodes in the tracking system have access to the same amount of information. Something that effects the sensor track quality is the feedback rate. If the feedback rate is decreased, so is the sensor track quality due to that the sensor node has access to less information.

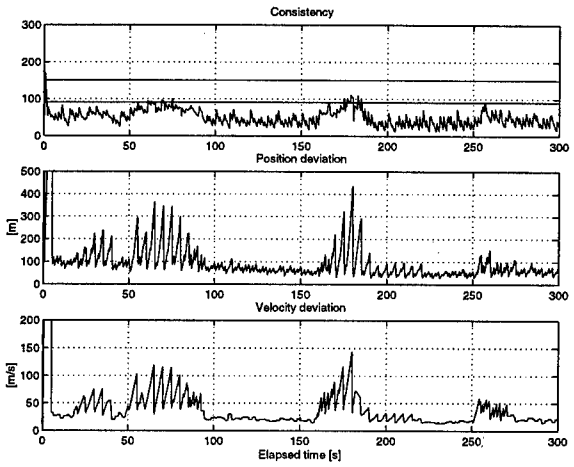


Figure 15. Consistency and track deviations in the radar node when track feedback is used.

In figure 5, the track quality for a single radar without feedback is shown. Compared to when track feedback is used (see figure 15), the result is disastrous concerning consistency and track quality. The maneuvering target is extremely hard to track if one radar without track feedback is used.

This shows that the benefit of track feedback is considerable in the sensor tracking nodes. Since the feedback results in sensor tracks of higher quality, the ability to associate the right measurement to the right track is improved. Thereby, the total tracking performance of the system increases.

This section investigates the improved tracking quality in the radar, but in the IRST sensor there may be even more enhancements. The IRST sensor does not measure any distance to the objects and therefore has poor track accuracy in the distance direction. By using feedback, the IRST gets the same range conception as the rest of the system and thus the

association of new measurements can be greatly simplified.

6. Conclusions

The main conclusion in this paper is that, in a decentralized tracking system with feedback, the cross-correlation between the tracks must be considered. Any combination of Filter A and C (estimate decorrelation and Covariance Intersection) proves to solve this problem. However, Filter B does not handle the cross-correlation and can therefore not be used in a system with feedback. The resulting track quality and uncertainty consistency for Filter A and C are approximately the same. Also, the results are about the same in Cartesian Coordinates and MSC.

The performance of a decentralized system with feedback is comparable to a centralized system, with respect to the quality of the fused tracks. The main benefit from feedback is improved tracking quality in the sensors, which for instance leads to enhanced measurement-to-track association performance. This is especially apparent when the measurement frequencies are low.

Measurement rates can be altered in the tracking system without ruining the consistency. Of course, the track quality is degraded if the measurement rate is decreased as a result of the decline in information in the system, but the system is still functioning.

Feedback rates can be altered in the tracking system without ruining the functionality, considering consistency and track quality. The only effect is that the sensor tracks have access to less information and thus have lower quality.

7. References

[1] Jensen, T., Karlsson, M., Malmberg, A. and Norberg, P., *A Fuzzy Logical Approach to Fighter Aircraft Sensor Management in a Complex Multisensor Environment,* Proceedings of Advances in Systems, Signals, Control and Computers, Vol. 2, pp. 190-195, 1998.

[2] Chong, C. Y., *Distributed Architectures for Data Fusion*, Proceedings of the 1998 International Conference on Multisource-Multisensor Information Fusion (FUSION'98), pp. 84-91, 1998.

[3] Grime, S. and Durrant-Whyte, H.F., *Data Fusion in Decentralized Sensor Networks*, Control Eng. Practice, Vol.2, No. 5, pp. 849-863, 1994.

[4] Drummond, O. E., *Feedback in Track Fusion Without Process Noise*, Signals and Data Processing of Small Targets, SPIE vol. 2561, pp 369-383, 1995.

[5] Uhlmann J. K., Julier S. J. and Csorba M., *Nondivergent Simultaneous Map Building and Localization using Covariance Intersection*, Proceedings of the SPIE Aerosense Conference, vol. 3087, 1997.

[6] Stallard, D. V., *An Angle-Only Tracking Filter in Modified Spherical Coordinates*, Proceedings of the AIAA Guidance Navigation Control Conference, pp.542-550, 1987.

[7] Bar-Shalom, Y. and Fortmann, T. E., *Tracking and Data Association*, Academic Press, 1991.

IPDAF in Distributed Sensor Networks for tracking an occulted ground-target in clutter

Jean Dezert
ONERA (DTIM), 29 Av. de la Division Leclerc
92320 Châtillon, France
E-mail: dezert@onera.fr

Abstract - ***An improved version of the Integrated Probabilistic Data Association Filter (IPDAF) and the IJPDAF based on a new concept of probability of target perceivability has been recently introduced for tracking one or several targets by a single sensor. IPDAF and IJPDAF algorithms allow to perform online track initiation, maintenance, confirmation and termination as well using an appropriate target perceivability probability decision logic. This paper deals with the development of a DSN (Distributed Sensor Networks) version of the new IPDAF algorithm. Simulation results of this new DSN/IPDAF algorithm for tracking a single occasionally occulted ground-target in a cluttered urban environment is presented for a simple 2D scenario.***

Keywords: Distributed Estimation, Multisensor Target Tracking, IPDAF, DSN, perceivability.

1 Introduction

A distributed sensor network (DSN) is a set of sensors connected by a communication network to a set of local processing nodes. These nodes process measurements and communicate among themselves in order to track the target. An important problem in distributed tracking is how to decide whether local tracks delivered at the local processing level represent the same target. We assume here that this track-to-track association problem has been solved (see [6] for discussion). In previous works done by K.C. Chang and al. during last decade [9, 7, 8, 10, 11, 21], the DSN sensor target tracking problem has been solved on the basis of classical PDAF and/or JPDAF algorithms (also coupled with Interacting Multiple Model (IMM) approach for maneuvering target tracking). It has already been shown that performances obtained with distributed estimation algorithms are very close to the optimal performance obtained by a centralized estimation algorithm. Moreover it is well known that DSN has many advantages over a centralized system in terms of reliability, extended coverage, better use of information and so forth. These Distributed PDAF/JPDAF algorithms have however been developed with an implicit strong assumption that the targets are *always perceivable* by the sensors. A target is said to be perceivable if it is present in the environment and not hidden/occulted in the field of view of the sensor. Of course in many real situations and like the one described in this paper, this is not always the case. To remove this total perceivability assumption, new versions of the Integrated Probabilistic Data Association Filter (IPDAF) and IJPDAF for a single sensor/tracker have been developed recently in [14, 13] which includes a more rigourous concept of target perceivability [15, 18] into its formalism than privious works of Colegrove [12] and Musicki [22]. Hereafter we extent this new IPDAF for DNS in order to extend their application fields to more realistic situations.

2 Problem formulation

We consider an s-node distributed sensor network as in [8] where each node processes the local measurements from its own sensor based on a local IPDAF and sends the local estimates to the fusion processor periodically. The fusion processor then sends back the processed results after each communication time. The dynamic of the target in track is modeled as

$$\mathbf{x}(k+1) = \mathbf{F}(k)\mathbf{x}(k) + \mathbf{v}(k) \qquad (1)$$

where $\mathbf{x}(k)$ is the state vector and $\mathbf{v}(k)$ is the process noise assumed to be zero-mean and Gaussian

with a known covariance matrix $\mathbf{Q}(k)$. The target detection probability P_d^i for each sensor i is assumed to be known. The equation measurement for the target relative to sensor i is

$$\mathbf{z}^i(k) = \mathbf{H}^i(k)\mathbf{x}(k) + \mathbf{w}^i(k) \qquad (2)$$

where $\mathbf{H}^i(k)$ is a known observation matrix and $\mathbf{w}^i(k)$ is the corresponding measurement noise assumed to be zero-mean, Gaussian with a given covariance $\mathbf{R}^i(k)$. Furthermore noise sequences $\{\mathbf{v}(k)\}$ and $\{\mathbf{w}^i(k)\}$ $(k = 1, 2, \ldots)$ are assumed to be mutually independent and independent of initial state vector $\mathbf{x}(0)$.

The classical gating technique [4] with a given probability P_g^i $(i = 1, \ldots, s)$ is used for the selection of measurements. For each sensor $i = 1, \ldots, s$, the set of the m_k^i validated measurement at time k and the cumulative set of measurements are denoted

$$\mathbf{Z}^i(k) = \{\mathbf{z}_{j_i}^i(k)\}_{j_i=1}^{m_k^i} \quad \text{and} \quad \mathbf{Z}^{i,k} = \{\mathbf{Z}^i(l)\}_{l=1}^{k}$$

The distributed estimation problem we have to solve is the reconstruction of the global conditional pdf $p(\mathbf{x}(k)|\mathbf{Z}^{1,k}, \ldots, \mathbf{Z}^{s,k})$ from the local ones $p(\mathbf{x}(k)|\mathbf{Z}^{1,k}), \ldots, p(\mathbf{x}(k)|\mathbf{Z}^{s,k})$. Under linear models and Gaussian noise assumptions, this problem reduces to evaluate $\hat{\mathbf{x}}(k|k) = E[\mathbf{x}(k)|\mathbf{Z}^{1,k}, \ldots, \mathbf{Z}^{s,k}]$ from local estimates with its covariance $\mathbf{P}(k|k)$.

3 The Local IPDAF

At a given node associated with a sensor s, the local tracking is assumed to be done with the new IPDAF. This tracking filter is an extension of the classical PDAF which integrates the concept of target perceivability.

At any time k, the target state of perceivability with respect to a given sensor s and its complement is represented by the exhaustive and exclusive events

$$\begin{aligned} O_k^s &\triangleq \{\text{target is perceivable from } s\} \\ \bar{O}_k^s &\triangleq \{\text{target is unperceivable from } s\} \end{aligned}$$

When there are m_k^s validated measurements at time k, the intersection of these events with the classical data association events involved in the PDAF formalism [4]

$$\begin{aligned} \theta_i^s(k) &\triangleq \{\mathbf{z}_{j_s}^s(k) \text{ comes from target}\} \\ \theta_0^s(k) &\triangleq \{\text{none of } \mathbf{z}_{j_s}^s(k) \text{ comes from target}\} \end{aligned}$$

defines a new set of integrated association events

$$\begin{aligned} \mathcal{E}_{-j_s}^s(k) &\triangleq \bar{O}_k^s \cap \theta_{j_s}^s(k) \qquad j_s = 1, \ldots, m_k^s \\ \mathcal{E}_{\bar{0}}^s(k) &\triangleq \bar{O}_k^s \cap \theta_0^s(k) \\ \mathcal{E}_0^s(k) &\triangleq O_k^s \cap \theta_0^s(k) \\ \mathcal{E}_{j_s}^s(k) &\triangleq O_k^s \cap \theta_{j_s}^s(k) \qquad j_s = 1, \ldots, m_k^s \end{aligned}$$

Since any target measurement cannot arise without target perceivability, events $\mathcal{E}_{-j_s}^s(k), j_s = 1, \ldots, m_k^s$ are impossible and we have $P\{\mathcal{E}_{-i}(k)|.\} = P\{\emptyset|.\} = 0$. Only events $\mathcal{E}_{\bar{0}}^s(k)$, $\mathcal{E}_0^s(k)$ and $\mathcal{E}_{j_s}^s(k)$ $(j_s = 1, \ldots, m_k^s)$ may have a non null probability to occur. The development of a new PDAF (called IPDAF) based on these integrated association events yields the following updating equations (see [14, 15] for complete derivation) which are valid for $m_k^s \geq 0$:

$$\hat{\mathbf{x}}^s(k|k) = \sum_{j_s=\bar{0},0}^{m_k^s} \beta_{j_s}^s(k)\hat{\mathbf{x}}_{j_s}^s(k|k) \qquad (3)$$

$$\begin{aligned} \mathbf{P}^s(k|k) = &\Big[\sum_{j_s=\bar{0},0}^{m_k^s} \beta_{j_s}^s(k)\mathbf{P}_{j_s}^s(k|k)\Big] \\ &- \hat{\mathbf{x}}^s(k|k)\hat{\mathbf{x}}^s(k|k)' + \sum_{j_s=\bar{0},0}^{m_k^s} \beta_{j_s}^s(k)\hat{\mathbf{x}}_{j_s}^s(k|k)\hat{\mathbf{x}}_{j_s}^s(k|k)' \end{aligned} \qquad (4)$$

where the conditional estimates and their covariances are

$$\hat{\mathbf{x}}_{\bar{0}}^s(k|k) = \hat{\mathbf{x}}^s(k|k-1) \qquad (5)$$

$$\hat{\mathbf{x}}_0^s(k|k) = \hat{\mathbf{x}}^s(k|k-1) \qquad (6)$$

$$\hat{\mathbf{x}}_{j_s}^s(k|k) = \hat{\mathbf{x}}^s(k|k-1) + \mathbf{K}^s(k)\tilde{\mathbf{z}}_{j_s}^s(k) \qquad (7)$$

$$\mathbf{P}_{\bar{0}}^s(k|k) = \mathbf{P}^s(k|k-1) \qquad (8)$$

$$\mathbf{P}_0^s(k|k) = [\mathbf{I} + q_0^s\mathbf{K}^s(k)\mathbf{H}^s(k)]\mathbf{P}^s(k|k-1) \qquad (9)$$

$$\mathbf{P}_{j_s}^s(k|k) = [\mathbf{I} - \mathbf{K}^s(k)\mathbf{H}^s(k)]\mathbf{P}^s(k|k-1) \qquad (10)$$

with the following computations [14, 15, 16]

$$\begin{aligned} q_0^s &\triangleq \frac{P_d^s(P_g^s - P_{gg}^s)}{1 - P_d^s P_g^s} \\ P_g^s &\triangleq P\{\chi_{n_{z^s}}^2 \leq \gamma\} \\ P_{gg}^s &\triangleq P\{\chi_{n_{z^s}+2}^2 \leq \gamma\} \\ \mathbf{S}^s(k) &= \mathbf{H}^s(k)\mathbf{P}^s(k|k-1)\mathbf{H}^s(k) + \mathbf{R}^s(k) \\ \mathbf{K}^s(k) &= \mathbf{P}^s(k|k-1)\mathbf{H}^s(k)'[\mathbf{S}^s(k)]^{-1} \\ \hat{\mathbf{z}}^s(k|k-1) &= \mathbf{H}^s(k)\hat{\mathbf{x}}^s(k|k-1) \end{aligned}$$

$$\tilde{\mathbf{z}}^s_{j_s}(k) = \mathbf{z}^s_{j_s}(k) - \hat{\mathbf{z}}^s(k|k-1)$$

$$\tilde{\mathbf{z}}^s(k) = \sum_{j_s=1}^{m^s_k} \beta^s_{j_s}(k)\tilde{\mathbf{z}}^s_{j_s}(k)$$

The integrated a posteriori data association probabilities $\beta_{j_s}(k) \triangleq P\{\mathcal{E}_{j_s}(k)|\mathbf{Z}^s(k), m^s_k, \mathbf{Z}^{k-1,s}\}$ $(j_s = \bar{0}, 0, \ldots, m^s_k)$ taking into account the target perceivability are given by

- when $m^s_k = 0$,

$$\beta^s_{\bar{0}}(k) = 1 - P^{O^s}_{k|k-1,0} \quad (11)$$

$$\beta^s_0(k) = P^{O^s}_{k|k-1,0} \quad (12)$$

- when $m^s_k > 0$,

$$\beta^s_{j_s}(k) = \frac{1}{c^s} e_{j_s}(k) P^{O^s}_{k|k-1,m^s_k} \quad (13)$$

$$\beta^s_0(k) = \frac{1}{c^s} b^s_0(k) P^{O^s}_{k|k-1,m^s_k} \quad (14)$$

$$\beta^s_{\bar{0}}(k) = \frac{1}{c^s} b^s_{\bar{0}}(k)(1 - P^{O^s}_{k|k-1,m^s_k}) \quad (15)$$

where c^s is a normalization constant and

$$e_{j_s}(k) \triangleq \frac{1}{P^s_g}\mathcal{N}[\tilde{\mathbf{z}}^s_{j_s}(k); 0; \mathbf{S}^s(k)]$$

$$b^s_0(k) \triangleq \frac{m^s_k}{V^s_k}\frac{1 - P^s_d P^s_g}{P^s_d P^s_g}\xi^s_k$$

$$b^s_{\bar{0}}(k) \triangleq \frac{m^s_k}{V^s_k}\frac{1}{P^s_d P^s_g}[P^s_d P^s_g + (1 - P^s_d P^s_g)\xi^s_k]$$

V^s_k is the volume of the measurement validation gate for sensor s [4, 5] and ξ^s_k, $\mu_F(.)$ are defined as

$$\xi^s_k \triangleq \frac{\mu_F(m^s_k)}{\mu_F(m^s_k - 1)}$$

$\mu_F(.) \triangleq$ pmf of number of false alarms in V^s_k

If a Poisson model with clutter density λ^s for μ_F is assumed, the predicted and updated conditional and unconditional target perceivability probabilities ($P^{O^s}_{k|k-1} \triangleq P\{O^s_k|\mathbf{Z}^{k-1,s}\}$ and $P^{O^s}_{k|k} \triangleq P\{O^s_k|\mathbf{Z}^{k,s}\}$) can be expressed as [14, 15, 16]

$$P^{O^s}_{k|k-1,m^s_k} = \frac{(1 - \epsilon^s_k)P^{O^s}_{k|k-1}}{1 - \epsilon^s_k P^{O^s}_{k|k-1}} \quad (16)$$

with

$$\epsilon^s_k \triangleq \begin{cases} P^s_d P^s_g & m^s_k = 0 \\ P^s_d P^s_g(1 - \frac{m^s_k}{\lambda^s V^s_k}) & m^s_k \neq 0 \end{cases} \quad (17)$$

and

$$P^{O^s}_{k|k-1} = \pi^s_{11} P^{O^s}_{k-1|k-1} + \pi^s_{21}(1 - P^{O^s}_{k-1|k-1}) \quad (18)$$

$$P^{O^s}_{k|k} = \frac{(1 - \phi^s_k)P^{O^s}_{k|k-1}}{1 - \phi^s_k P^{O^s}_{k|k-1}} \quad (19)$$

$$\phi^s_k \triangleq \begin{cases} P^s_d P^s_g & m^s_k = 0 \\ P^s_d P^s_g(1 - \frac{1}{\lambda^s}\sum_{j_s=1}^{m^s_k} e_{j_s}) & m^s_k \neq 0 \end{cases} \quad (20)$$

Hence $P^{O^s}_{k|k-1}$ and $P^{O^s}_{k|k}$ can be computed on-line recursively as soon as the *design parameters* $\pi^s_{11} \triangleq P\{O^s_k|O^s_{k-1}\}$, $\pi^s_{21} \triangleq P\{O^s_k|\bar{O}^s_{k-1}\}$ and $P^{O^s}_{1|0}$ have been set. In practice, the clutter density λ^s is usually unknown. To implement the IPDAF, we have to replace λ^s by its estimation based on the Bayesian (conditional mean) estimation, the maximum likelihood method or the least squares method recently developed in [15, 19]. Theoretical investigations on design of IPDAF trackers for perceivability probability enhancement can be found in [17].

Finally with some elementary algebra $\mathbf{P}^s(k|k)$ given by (4) can take the following forms depending on m^s_k

- when $m^s_k = 0$, $\mathbf{P}^s(k|k) =$

$$[\mathbf{I} + q^s_0 P^{O^s}_{k|k-1,0}\mathbf{K}^s(k)\mathbf{H}^s(k)]\mathbf{P}^s(k|k-1)$$

- when $m^s_k > 0$, $\mathbf{P}^s(k|k) =$

$$\beta^s_{\bar{0}}(k)\mathbf{P}^s(k|k-1) + \beta^s_0(k)[\mathbf{I} + q^s_0\mathbf{K}^s(k)\mathbf{H}^s(k)]\mathbf{P}^s(k|k-1) + (1 - \beta^s_{\bar{0}}(k) - \beta^s_0(k))\mathbf{P}^{c,s}(k|k) + \tilde{\mathbf{P}}^s(k)$$

with

$$\mathbf{P}^{c,s}(k|k) = [\mathbf{I} - \mathbf{K}^s(k)\mathbf{H}^s(k)]\mathbf{P}^s(k|k-1)$$

$$\tilde{\mathbf{P}}^s(k) = \mathbf{K}^s(k)[\sum_{j_s=1}^{m^s_k} \beta^s_{j_s}(k)\tilde{\mathbf{z}}^s_{j_s}(k)\tilde{\mathbf{z}}^s_{j_s}(k)' - \tilde{\mathbf{z}}^s(k)\tilde{\mathbf{z}}^s(k)']\mathbf{K}^s(k)'$$

The local state prediction is done according to classical prediction equations, i.e.

$$\hat{\mathbf{x}}^s(k+1|k) = \mathbf{F}(k)\hat{\mathbf{x}}^s(k|k)$$

$$\mathbf{P}^s(k+1|k) = \mathbf{F}(k)\mathbf{P}^s(k|k)\mathbf{F}'(k) + \mathbf{Q}(k)$$

4 The Distributed IPDAF

Given the local statistics delivered by s local IPDAF of a s-node sensor network[1], we are now looking for the solution of the distributed estimation problem in order to retrieve the optimal global target state estimate and its covariance which are given by [2]

$$\hat{\mathbf{x}}(k|k) = E[\mathbf{x}(k)|\mathbf{Z}^{1,k},\dots,\mathbf{Z}^{s,k}] = \sum_{j_1=\bar{0},0}^{m_k^1} \cdots \sum_{j_s=\bar{0},0}^{m_k^s} \beta_{j_1,j_s}(k)\hat{\mathbf{x}}_{j_1,j_s}(k|k) \quad (21)$$

with

$$\beta_{j_1,j_s}(k) \triangleq P(\mathcal{E}_{j_1}^1(k),\dots,\mathcal{E}_{j_s}^s(k)|\mathbf{Z}^{1,k},\dots,\mathbf{Z}^{s,k})$$
$$\hat{\mathbf{x}}_{j_1,j_s}(k|k) \triangleq E[\mathbf{x}(k)|\mathbf{Z}^{1,k},\mathcal{E}_{j_1}^1(k),\dots,\mathbf{Z}^{s,k},\mathcal{E}_{j_s}^s(k)]$$

and

$$\mathbf{P}(k|k) = \sum_{j_1=\bar{0},0}^{m_k^1}\sum_{j_s=\bar{0},0}^{m_k^s} \beta_{j_1,j_s}(k)\mathbf{P}_{j_1,j_s}(k|k) + \sum_{j_1=\bar{0},0}^{m_k^1}\sum_{j_s=\bar{0},0}^{m_k^s} \beta_{j_1,j_s}(k)\big[\hat{\mathbf{x}}_{j_1,j_s}(k|k)\hat{\mathbf{x}}_{j_1,j_s}(k|k)' - \hat{\mathbf{x}}(k|k)\hat{\mathbf{x}}(k|k)'\big]$$

These previous equations are always valid whatever the values of $m_k^1,\dots,m_k^s$ are. If there is no validated measurement for a given node at a given time, the corresponding summation must be only computed from $\bar{0}$ up to 0.

If we assume the measurement errors from sensors independent, the joint conditional estimates with their covariances can be obtained from the optimal distributed fusion equations of Chong [9, 8, 2, 6].

$$\hat{\mathbf{x}}_{j_1,j_s}(k|k) = \mathbf{P}_{j_1,j_s}(k|k)\Big[\mathbf{P}(k|k-1)^{-1}\hat{\mathbf{x}}(k|k-1) + \sum_{i=1}^{s}\mathbf{P}_{j_i}^i(k|k)^{-1}\hat{\mathbf{x}}_{j_i}^i(k|k) - \sum_{i=1}^{s}\mathbf{P}^i(k|k-1)^{-1}\hat{\mathbf{x}}^i(k|k-1)\Big] \quad (22)$$

$$\mathbf{P}_{j_1,j_s}{}^{-1}(k|k) = \mathbf{P}(k|k-1)^{-1} + \sum_{i=1}^{s}\mathbf{P}_{j_i}^i(k|k)^{-1} - \sum_{i=1}^{s}\mathbf{P}^i(k|k-1))^{-1} \quad (23)$$

When all nodes communicate every scan the global and local prior estimates are the same (i.e. $\hat{\mathbf{x}}^i(k|k-1) = \hat{\mathbf{x}}(k|k-1)$ and $\mathbf{P}^i(k|k-1) = \mathbf{P}(k|k-1)$) and then eqs. (22) and (23) will reduce to

$$\hat{\mathbf{x}}_{j_1,j_s}(k|k) = \mathbf{P}_{j_1,j_s}(k|k)\Big[[\sum_{i=1}^{s}\mathbf{P}_{j_i}^i(k|k)^{-1}\hat{\mathbf{x}}_{j_i}^i(k|k)] - (s-1)\mathbf{P}(k|k-1)^{-1}\hat{\mathbf{x}}(k|k-1)\Big] \quad (24)$$

$$\mathbf{P}_{j_1,j_s}{}^{-1}(k|k) = [\sum_{i=1}^{s}\mathbf{P}_{j_i}^i(k|k)^{-1}] - (s-1)\mathbf{P}(k|k-1)^{-1} \quad (25)$$

The derivation of $\beta_{j_1,j_s}(k)$ is quite complicated and will not be detailed here. We refer the reader to [8] for a complete derivation. Assuming the independence between sensor measurements and between events $\mathcal{E}_{j_1}^1(k),\dots,\mathcal{E}_{j_s}^s(k)$ given the target state, then the final expression for $\beta_{j_1,j_s}(k)$ is

$$\beta_{j_1,j_s}(k) = \frac{1}{c}\gamma(\mathcal{E}_{j_1}^1(k),\dots,\mathcal{E}_{j_s}^s(k))\prod_{i=1}^{s}\beta_{j_i}^i(k) \quad (26)$$

where c is a normalization constant such that

$$\sum_{j_1=\bar{0},0}^{m_k^1}\cdots\sum_{j_s=\bar{0},0}^{m_k^s}\beta_{j_1,j_s}(k) = 1$$

and where the correlation factor $\gamma(\mathcal{E}_{j_1}^1(k),\dots,\mathcal{E}_{j_s}^s(k))$ is given by

$$\int p(\mathbf{x}(k)|\mathbf{Z}^{1,k-1},\mathbf{Z}^{s,k-1})\frac{\prod_{i=1}^{s}p(\mathbf{x}(k)|\mathcal{E}_{j_i}^i(k),\mathbf{Z}^{i,k})}{\prod_{i=1}^{s}p(\mathbf{x}(k)|\mathbf{Z}^{i,k-1})}d\mathbf{x}$$

Using the gaussian distribution approximation and moment matching method, it can be shown

[1] s represents now the total number of sensors in the DSN instead of typical sensor index as in previous section

[2] due to space limitation, notation j_1, j_s must actually be read $j_1,\dots,j_s$ and sometimes $\mathbf{Z}^{1,k},\mathbf{Z}^{s,k}$ as $\mathbf{Z}^{1,k},\dots,\mathbf{Z}^{s,k}$

that $\gamma(\mathcal{E}^1_{j_1}(k),\ldots,\mathcal{E}^s_{j_s}(k))$ can be approximated by

$$\sqrt{\frac{|\mathbf{P}_{j_1,j_s}(k|k)|\prod_{i=1}^{s}|\mathbf{P}^i(k|k-1)|}{|\mathbf{P}(k|k-1)|\prod_{i=1}^{s}|\mathbf{P}^i_{j_i}(k|k)|}}e^{-\frac{1}{2}D_{j_1,j_s}}$$

with

$$\begin{aligned}D_{j_1,j_s} \triangleq &\Big[\sum_{i=1}^{s}\hat{\mathbf{x}}^i_{j_i}(k|k)'\mathbf{P}^i_{j_i}(k|k)^{-1}\hat{\mathbf{x}}^i_{j_i}(k|k)\\ &-\hat{\mathbf{x}}^i(k|k-1)'\mathbf{P}^i(k|k-1)^{-1}\hat{\mathbf{x}}^i(k|k-1)\Big]\\ &+\hat{\mathbf{x}}(k|k-1)'\mathbf{P}(k|k-1)^{-1}\hat{\mathbf{x}}(k|k-1)\\ &-\hat{\mathbf{x}}_{j_1,j_s}(k|k)'\mathbf{P}_{j_1,j_s}(k|k)^{-1}\hat{\mathbf{x}}_{j_1,j_s}(k|k)\end{aligned}$$

When all nodes communicate every scan, $\gamma(\mathcal{E}^1_{j_1}(k),\ldots,\mathcal{E}^s_{j_s}(k))$ will reduce to

$$\sqrt{\frac{|\mathbf{P}_{j_1,j_s}(k|k)||\mathbf{P}(k|k-1)|^{s-1}}{\prod_{i=1}^{s}|\mathbf{P}^i_{j_i}(k|k)|}}e^{-\frac{1}{2}D_{j_1,j_s}}$$

with

$$\begin{aligned}D_{j_1,j_s} \triangleq &\Big[\sum_{i=1}^{s}\hat{\mathbf{x}}^i_{j_i}(k|k)'\mathbf{P}^i_{j_i}(k|k)^{-1}\hat{\mathbf{x}}^i_{j_i}(k|k)\Big]\\ &-(s-1)\hat{\mathbf{x}}(k|k-1)'\mathbf{P}(k|k-1)^{-1}\hat{\mathbf{x}}(k|k-1)\\ &-\hat{\mathbf{x}}_{j_1,j_s}(k|k)'\mathbf{P}_{j_1,j_s}(k|k)^{-1}\hat{\mathbf{x}}_{j_1,j_s}(k|k)\end{aligned}$$

where $\hat{\mathbf{x}}_{j_1,j_s}(k|k)$ and $\mathbf{P}_{j_1,j_s}{}^{-1}(k|k)$ are obtained from (24) and (25) respectively.

5 Simulation results

A two-dimensional single ground-target tracking problem is considered here. The target is assumed to move on a road in a town with (nearly) constant velocity of 36 km/h during 110 s from crossroad A towards the crossroad C as on figure 1. Only three buildings $B1$, $B2$ and $B3$ have been simulated in our scenario. The target dynamic model (i.e. piecewise constant white acceleration model) with discretization over time interval of length $T = 1s$ is [5]

$$\mathbf{x}(k+1) = \mathbf{F}\mathbf{x}(k) + \mathbf{G}\mathbf{v}(k)$$

where $\mathbf{x}(k) = [x\ \dot{x}\ y\ \dot{y}]'$ is the target state vector at time k and $\mathbf{F}$ and $\mathbf{G}$ are given by

$$\mathbf{F} = \begin{bmatrix}1 & T & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & T\\ 0 & 0 & 0 & 1\end{bmatrix} \qquad \mathbf{G} = \begin{bmatrix}T^2/2 & 0\\ T & 0\\ 0 & T^2/2\\ 0 & T\end{bmatrix}$$

The process noise $\mathbf{v}(k)$ representing the acceleration during one period is a zero-mean Gaussian white noise having covariance $\mathbf{Q}_v = diag(q_v, q_v)$ with $q_v = (0.001m/s^2)^2$. The magnitude of the process noise has been chosen very low in order to force the target to move on the segment $[A;C]$ (middle of the road). The true initial target state is assumed to be $\mathbf{x}(0) = [-800\text{ m}\ 10\text{ m/s}\ -450\text{ m}\ 0\text{ m/s}]'$.

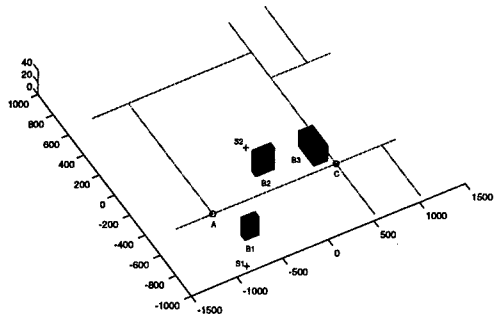


Figure 1: Urban environment scenario

We have considered a 2-nodes DSN with full communication at every scan. The sensor $S1$ is located at position $(-850\text{ m}, -950\text{ m})$ and $S2$ at $(-100\text{ m}, -50\text{ m})$. It is assumed that only position measurements are available, i.e.

$$\mathbf{z}^i(k) = \mathbf{H}\mathbf{x}(k) + \mathbf{w}^i(k) \qquad i = 1,2$$

with

$$\mathbf{H} = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}$$

Figure 2 shows the line of sight between sensors and the true target position for a given realization of the process noise. On average for our scenario, the target is occulted by building $B1$ for sensor $S1$ during period [25s; 72s] and by $B2$ for sensor $S2$ during period [50s; 92s]. Thus during period [50s; 72s] the target is occulted for both sensors.

Both sensors have same measurement precision. The standard deviation of measurement errors are 5 meters on x and y coordinates. The detection probabilities for both sensors are equal to 0.7 and the false alarm rates are both equal to $\lambda = 0.0003FA/m^{-2}$. The initial state estimate for both sensors is estimated using the so-called two-point differencing technique (TPD) [5, 6] (see also [20] for recent advances).

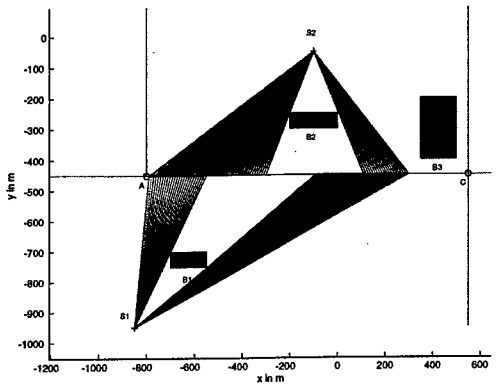


Figure 2: Perceivability scenario (top view)

At each scan, each node will process its own set of sensor measurements first using local IPDAF, then will send its local processed results to the fusion node. After receiving the information from local nodes, the fusion node will use the the distributed fusion algorithm presented at the end of section 4 to construct the global estimate and will send the results back to each local node at every sampling time. Both local IPDAF use the same set of design parameters ($P_g = 0.99$, $\pi_{11} = 0.988$, $\pi_{21} = 0.05$ and $P^O_{1|0} = 0.5$) and the true value λ for clutter density.

Simulations were carried out with 50 Monte Carlo runs. The results of successful runs for decentralized trackers (without fusion) are plotted on figures 3 and 4. A successful run is defined when the estimated target position is within 30 m of the true target position for the last three scans [7]. Figures 5 and 6 show the averaged performances of the successful runs for the decentralized case. We can observe from figure 5 and figure 6 that the target perceivability probabilities estimated by the local IPDAF fit well their true values even when the perceivability mode is switching. Obviously in nominal mode (for $k > 20s$), the rms position errors increase with time when the target becomes unperceivable by the sensors. The maximum of rms errors are obtained for k around 72 s and 92 s. These instants correspond to the end of the unperceivability period for each sensor. For the decentralized case, out of 50 runs, sensor 1 alone and sensor 2 alone only track the occulted target successfully in 29 and 41 runs, respectively.

Figures 7 and 8 show the results obtained with the distributed IPDAF (distributed communication scheme at every scan). According to the results plotted on the figures, the distributed IPDAF performs better than the single sensor configurations. In nominal mode, the maximum rms position error is now obtained for $k = 72$ s which corresponds to the end of the period where the target is unperceivable by both sensors simultaneously which makes sense with the theory. In such case, the DIPDAF sucessfully tracks the target in 48 out of 50 runs. Note also that the quality of estimation using both sensors in terms of mean square error and in terms of target perceivability estimation is significantly better than with the decentralized scheme. In our simulations the averaged number of false alarms per gate was around 0.5. The simulations shows the usefulness and the improvement of DIPDAF with respect to decentralized schemes for tracking an occulted ground-target in an urban cluttered environment.

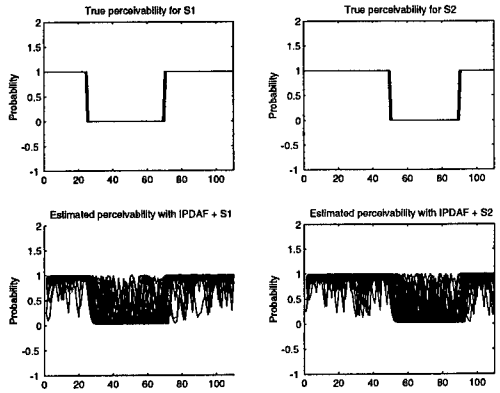


Figure 3: Estimated perceivability probabilities (decentralized communication case)

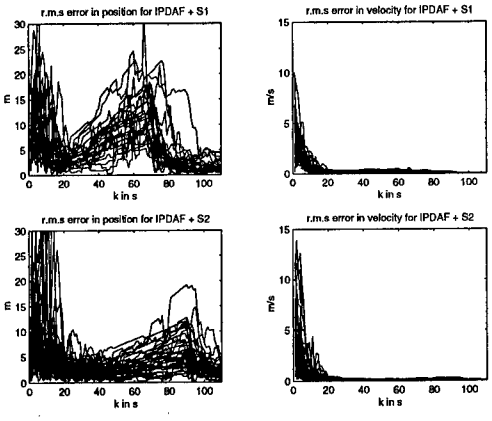


Figure 4: R.M.S. errors for successful runs (decentralized communication case)

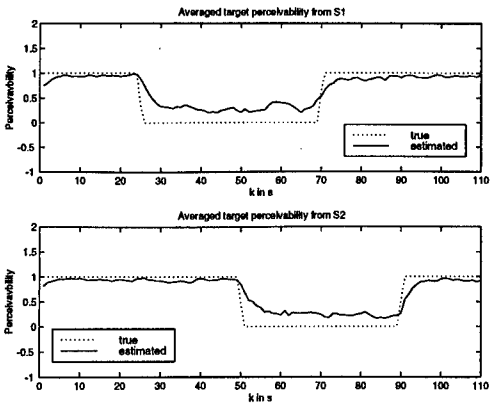


Figure 5: Averaged perceivability probabilities (decentralized communication case)

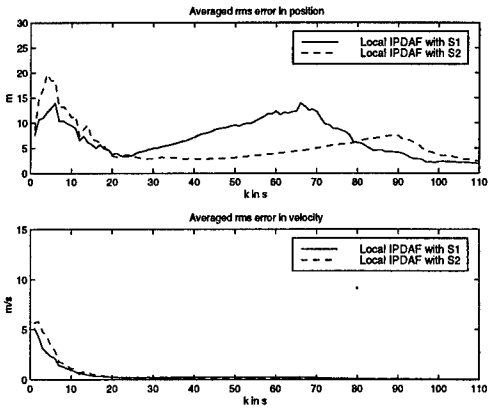


Figure 6: Averaged R.M.S. errors for successful runs (decentralized communication case)

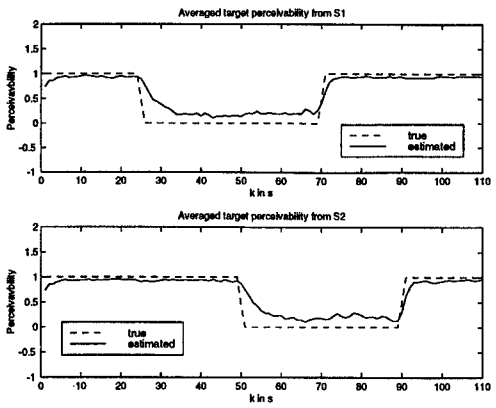


Figure 7: Averaged perceivability probabilities (distributed communication case)

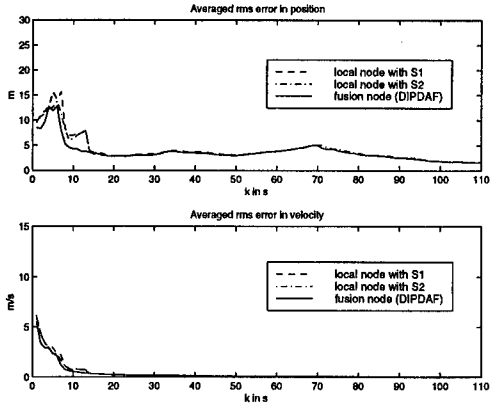


Figure 8: Averaged R.M.S. errors for successful runs (distributed communication case)

6 Conclusion

From a new formulation of IPDAF based on a recent method for target perceivability probability estimation and by following the theoretical approach of Chang and al. [8, 11, 21], a distributed version of IPDAF (called DIPDAF) has been proposed here (with implicit assumption of lossless communication of sufficient statistics). This algorithm takes into account the information fusion in a distributed sensor network. This new DIPDAF is fully coherent and intuitively appealing with the Distributed PDAF formulation [2] as soon as the target perceivability probabilities for each sensor becomes unitary. This filter has been successfully implemented for tracking a ground-target occasionnally occulted in a cluttered urban environment on a simple 2-nodes 2D scenario. Extension of this new tracker for tracking maneuvering target with or without different local observation models could also be developed by taking into account methodology described in previous works [7] and [1, 3]. Another extension of this algorithm for multi-target tracking based on the IJPDAF developed in [13] is under investigations.

References

[1] Alouani A.T.,"Linear Distributed Estimation", *Proceedings of the ACC*, pp. 412-417, (1987).

[2] Alouani A.T.,"Distributed Tracking with Probabilistic Data Association", *Proceedings of the 26th IEEE CDC*, Los Angeles, pp. 1612-1613, (1987).

[3] Alouani A.T.,"Distributed Estimators for Multi-Sensor Systems", *Advanced Information Processing in Automatic Control*, pp. 273-277, (1990).

[4] Bar-Shalom Y., Fortmann T.E., "Tracking and Data Association", *Academic Press*, (1988).

[5] Bar-Shalom Y., Li X.R., "Estimation and Tracking: Principles,Techniques, and Software", *Artech House*, (1993).

[6] Bar-Shalom Y., Li X.R., "Multitarget-Multisensor Tracking: Principles and Techniques", *YBS Publishing*, (1995).

[7] Chang K.C., Bar-Shalom Y., "Distributed Adaptive Estimation with Probabilistic Data Association", *Automatica*, **vol. 25**, no 3, pp. 359-369, (1984).

[8] Chang K.C., Bar-Shalom Y., "Joint Probabilistic Data Association in Distributed Sensor Networks", *IEEE Trans. AC*, **vol. 31**, no 10, pp. 889-897, (1986).

[9] Chong C.Y., "Hierarchical Estimation", *Proceedings of MIT/ONR C3 Workshop*, Monterey, CA, (1979).

[10] Chong C.Y., Mori S., Chang K.C., "Information Fusion in Distributed Sensor Networks", *Proceedings of American Control Conference*, pp. 830–835, (1985).

[11] Chong C.Y., Mori S., Chang K.C.,"Distributed Multitarget Multisensor Tracking", *in Multitarget-Multisensor Tracking : Advanced Applications*, Artech House (Y. Bar-Shalom Editor), (1990).

[12] Colegrove S.B., Ayliffe J.K., "An extension of Probabilistic Data Association to Include Track Initiation and Termination", *20th IREE International Convention,Melbourne*, pp. 853-856, (1985).

[13] Dezert J., Li N., Li X.R.,"Theoretical Development of an Integrated JPDAF for Multitarget Tracking in Clutter", *Proc. Workshop ISIS-GDR/NUWC, ENST, Paris*, (1998).

[14] Dezert J., Li N., Li X.R.,"A New Formulation of IPDAF for Tracking in Clutter", *To appear in Proc. of European Control Conference (ECC99)*, Karlsruhe, Germany, (Sept.1999).

[15] Li X.R., Li N., "Intelligent PDAF : Refinement of IPDAF for Tracking in Clutter", *Proc. 29th Southeastern Symp. on Syst. Theory*, pp. 133-137, (1997).

[16] Li X.R., "Tracking in Clutter with Strongest Neighbor Measurements - Part 1 : Theoretical Analysis", *IEEE Trans. AC*, **vol. 43**, pp. 1560-1578, (1998).

[17] Li N., Li X.R., "Theoretical Design of Trackers for Tracking Probability Enhancement", *Proc. of SPIE Conf., Orlando*, **vol. 3373**, (1998).

[18] Li N., Li X.R., "Target Perceivability : An Integrated Approach to Tracker Analysis and Design", *Proc. of Fusion 98 Intern. Conf., Las Vegas*, pp. 174-181, (1998).

[19] Li X.R., Li N., Zhu Y. , "Integrated Real-Time Estimation of Clutter Density for Tracking", *submitted to IEEE Trans. on Signal Processing.*, (1999).

[20] Li X.R., He C.,"Optimal Initialization of Linear Recursive Filters", *submitted for journals*, (1999).

[21] Luh P.B., Bar-Shalom Y., Chang K.C., "Centralized and Distributed Algorithms for Multitarget-Multisensor Tracking Systems", *Control and Dynamic Systems : Advances in Theory and Applications*, Academic Press (C.T. Leondes Ed.), **vol. 31**,pp. 33-74, (1989).

[22] Musicki D., Evans R.J., Stankovic S., "Integrated Probabilistic Data Association", *IEEE Trans. AC*, **vol. 39**, pp. 1237-1241, (1994).

An Efficient Method for Uniformly Generating Poisson-Distributed Number of Measurements In a Validation Gate

T.-J. Ho and M. Farooq

Department of Electrical and Computer Engineering
Royal Military College of Canada
Kingston, Ontario K7K 5L0
Canada

ho-t@rmc.ca and farooq@rmc.ca

Abstract: ***In this paper, we consider the problem of clutter scattering in a gate under the assumptions: 1)the number of false validated measurements that are generated can be described by a Poisson model; and 2)false validated measurements are generated uniformly. Also, we present an efficient method for generating the measurements that satisfactorily meet the stated assumptions. The method is based upon a proposition related to the principle axis theorem which, in turn, is utilized as a guideline for the generation of a random vector with a uniform distribution in*** $\mathbb{R}^n$. ***This method not only produces a desirable result within definitely finite steps, but also reduces computational loads quite often required by an existing technique cited in [7]. The efficiency and the feasibility of the proposed method is demonstrated by an example in target tracking.***

Keywords: clutter; target tracking; validation gate; Poisson-distributed number; Monte Carlo simulations

1 Introduction

When a target is tracked in a realistic environment, measurements are usually affected by uncertainties (e.g. the targets of interest, clutter, and false alarms). To analyze the performance of a tracking scheme, the error covariance matrix of a tracking filter usually serves as a performance measure. However, the error covariance matrices for the probability data association filter (PDAF) and the joint probability data association filter (JPDAF) incorporate random terms (see [1]). As a result, to characterize the statistics of tracking errors, it is essential to perform Monte Carlo simulations. To do these simulations, one must generate false measurements in a validation gate to characterize the clutter. The development of algorithms for target tracking [6] (for a single taget) and [5] (for multiple targets) is based on assumptions that 1) the number of false validated measurements can be described by a suitable Poisson model; 2) the false validated measurements are uniformly distributed in the gate and are independent from scan to scan. Following this objective, the primary aim of this paper is to present a method which can efficiently and satisfactorily generate false validated measurements in accordance to the above assumptions.

The method proposed in this paper follows the same assumptions as in [7]. However, the method in [7], when incorporated in evaluating the performance of a PDAF or a JPDAF, can be computationally expensive because the method requires an indefinite number of iterations to arrive at the desired results. In view

of this drawback, our method provides an efficient alternative with acceptable accuracy. This proposed method consists of two stages. First, we use the Poisson-distributed random generator in [3] to generate a number of false validated measurements in a gate. Secondly, each of the innovation corresponding to a false validated measurement is generated as follows. We transform a positive semi-definite matrix to a diagonal form with eigenvalues as diagonal entries. The criterion to check that a measurement falls within a gate can be expressed as a quadratic inequality of n variables with its eigenvalues as coefficients; thus, we can generate a random vector in n steps. Subsequently, this vector is transformed back to obtain a clutter. Hence, once the number of clutter points in a gate is determined, say m, we can achieve the objective in $m \times (n+2)$ steps. Our method can be easily implemented in a Matlab environment.

The paper is organized as follows. Section 2 describes the background of the problem and gives a brief coverage of some notations used throughout the paper. In Section 3, the existing method in [7] to generate the false validated measurements in a gate is summarized. The drawbacks of the method are also discussed in this section, and then the development to yield a computationally more efficient method with satisfactory performance is presented. Section 4 presents the Monte Carlo simulation results for a target-tracking example using a standard PDAF including either the method of [7] or the proposed method. The simulation results characterize the statistics of actual rms errors as well as error covariances in the example. The approximate computational efforts of these methods are compared quantitatively. Finally, conclusions are presented in Section 5.

2 Preliminary

In target tracking, a gate is formed around a predicted target position. Gating is used for eliminating measurements which are unlikely to belong to a target. Thus, a measurement falling within the gate is more likely from the target of interest, and hence, is referred to as a validated measurement.

We begin by reviewing quantities to be used in defining a gate. At time instant k, let the predicted state vector be $\hat{x}(k|k-1)$. In addition, the measurement is $z(k) = Hx(k) + \nu(k)$ where H is the measurement matrix and $\nu(k)$ is a zero-mean white Gaussian measurement noise with a covariance R. The residual vector is defined by taking the difference between the measurement and predicted measurement,i.e., $\mu = z(k) - H\hat{x}(k|k-1)$. The residual covariance matrix $S = HPH' + R$ is positive definite where P is the one step prediction covariance matrix, and the superscript $'$ denotes the transposition. The time index k will be omitted for notational convenience. Thus the n-dimensional "g-sigma" ellipsoid gate is

$$\mathbb{R}_g^n = \{\mu : \mu' S^{-1} \mu < g^2\} \qquad (1)$$

where g is a positive number and determines the size of the validation gate.

When the performance of a filter is evaluated, two commonly used assumptions on how clutter scatters in a gate are as follows.

1. The number of false measurements in the gate can be described by a Poisson model.
2. From one scan to another or within a scan, the detection of false measurements is uniformly distributed in the gate and the measurements are independent.

The first assumption yields a mathematical model for generating the number of false validated measurements at each time instant. Generally, we do not have the exact information regarding how validated measurements scatter in a gate. It is reasonable to assume that each validated measurement is likely to be a target of interest. For this reason, the second assumption reveals that the false validated measurements scatter uniformly in a gate.

Fro brevity, the time index k is omitted from some of the notations. In addition, we list the following notations for clarity.

V: the volume of the n-dimensional "g-sigma" ellipsoid $\mathbb{R}_g^n$;

$\lambda_{min}(S^{-1})$: the smallest eigenvalue of the matrix S^{-1};

$\lambda_{max}(S^{-1})$: the largest eigenvalue of the matrix S^{-1};

α: the spatial density in a Poisson model;

$U(a,b)$: a set of random points uniformly distributed in the interval (a,b).

3 Main Results

In this section, we discuss two methods of generating the validated false measurements in a gate. These include method A in [7], and the proposed method. The main drawbacks of method A are presented in Section 3.1. Then, the proposed method is presented in Section 3.2 to overcome the shortcomings of method in Section 3.1. The proposed method is computationally efficient and yields desired results.

3.1 Method A [7]

This method consists of two stages. To generate the Poisson-distributed number of false validated measurements, an algorithm is adopted from [3] as follows.

Stage I.
Poisson Random Generator (PRG)

Set $Num = -1$, $m = exp(\alpha V)$.
Repeat the following until $m < 1$
    Generate θ uniform on $(0,1)$
    Set $Num = Num + 1$, $m = m\theta$
Output Num

The following procedure distributes the false validated measurements uniformly in a gate.

Stage II.
1. l=1;
2.Repeat until $l > Num$
3. Generate random numbers
4. $\mu_i \in U(-\gamma/\lambda_{min}(S^{-1}), \gamma/\lambda_{min}(S^{-1}))$,
5. where $1 \leq i \leq n$;
6.Loop: Let $\mu = [\mu_1 \mu_2 ... \mu_n]$.
7. If $\mu'\mu \leq \gamma/\lambda_{max}(S^{-1})$,
8. then
9. l=l+1;
10. else
11. if $\mu' S^{-1} \mu \leq \gamma$,
12. then
13. l=l+1;
14. else go to Loop
15. end
16. end.

Remark 1: Although the work in [7] did not indicate the procedure to generate the number of false validated measurements by using a Poisson model; however, PRG can be employed in general to satisfy assumption 1.

Remark 2: It is worthwhile to note that the verification criterion in line 7 of Stage II needs to be satisfied to generate a random vector. As a result, the completion of Stage II may require an indefinite number of iterations.

3.2 Method B

Before we present the proposed method, the following lemma and proposition serve as a starting point for the development of the method. The lemma is adopted from [4]; it also can be found in many other textbooks on linear algebra.

Lemma 1. *If A is a $n \times n$ symmetric matrix, then there exists an orthogonal matrix L such that the matrix $L^{-1}AL$ is a diagonal matrix whose entries on the main diagonal are the eigenvalues of the matrix A.*

Proposition 1. *Let A be a $n \times n$ symmetric matrix. A vector $\eta = Lx$ with $\eta' A \eta < \gamma$ is uniform in $\mathbb{R}^n$ where L is an orthogonal matrix*

such that $L' = L^{-1}$ *and the matrix*

$$L^{-1}AL = \begin{bmatrix} \lambda_1 & 0 & 0 & ... & 0 \\ 0 & \lambda_2 & 0 & ... & 0 \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 0 & 0 & ... & 0 & \lambda_n \end{bmatrix}$$

where $\lambda_1 \leq \lambda_2 ... \leq \lambda_n$ *and* λ_i, $1 \leq i \leq n$, *is an eigenvalue of* A, *and the vector* $x = [x_1 ... x_n]'$ *is uniform in* $\mathbb{R}^n$ *where*

$$x_1 \in U(-(\gamma/\lambda_1)^{1/2}, (\gamma/\lambda_1)^{1/2})$$

and

$$x_i \in U(-(\tau_i/\lambda_i)^{1/2}, (\tau_i/\lambda_i)^{1/2}),$$

with $2 \leq i \leq n$, $\tau_i = (\gamma - \lambda_1 x_1^2 - ... - \lambda_{i-1} x_{i-1}^2)$.

Proof: With lemma 1, $\eta' A \eta < \gamma$ can be rewritten as the following quadratic form

$$\lambda_1 x_1^2 + \lambda x_2^2 + ... + \lambda_n x_n^2 < \gamma$$

where $\eta = Lx$ with $L' = L^{-1}$ and the vector $x = [x_1, ..., x_n]'$, and λ_i, $1 \leq n$, are eigenvalues of A satisfying $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$. Firstly, let $x_1 \in U(-(\gamma/\lambda_1)^{1/2}, (\gamma/\lambda_1)^{1/2})$. Subsequently, let $x_2 \in U(-((\gamma - \lambda_1 x_1^2)/\lambda_2)^{1/2}, ((\gamma - \lambda_1 x_1^2)/\lambda_2)^{1/2})$. By the same reasoning, it follows that each

$$x_i \in U(-(\tau_i/\lambda_i)^{1/2}, (\tau_i/\lambda_i)^{1/2}),$$

with $\tau_i = (\gamma - \lambda_1 x_1^2 - ... - \lambda_{i-1}^2)$, $1 \leq i \leq n$. Thus the vector x is uniform in $\mathbb{R}^n$. Therefore, the vector $\eta = Lx$ with $\eta' A \eta < \gamma$ is uniform in $\mathbb{R}^n$. ■

Some aspects of Proposition 1 are similar to the principle axis theorem. The result of Proposition 1 gives us a guideline for the generation of a vector which is uniform in $\mathbb{R}^n$ and satisfies eqn.(1).

As in [7], the proposed method consists of two stages. The first stage is to use PRG. Suppose S^{-1} exists. As S is a $n \times n$ matrix, so is S^{-1}. Then, based on the generated number of false validated measurements, the second stage is to generate false measurements satisfying eqn.(1) which are uniformly distributed in the gate. This is accomplished as follows.

Stage II.

1. Obtain the orthogonal matrix L such that

$$L^{-1}S^{-1}L = \begin{bmatrix} \lambda_1 & 0 & 0 & ... & 0 \\ 0 & \lambda_2 & 0 & ... & 0 \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 0 & 0 & ... & 0 & \lambda_n \end{bmatrix}$$

where each λ_i, $1 \leq i \leq n$, is an eigenvalue of the matrix S^{-1} and $\lambda_1 \leq \lambda_2 ... \leq \lambda_n$.

2. l=1;
3. Repeat until $l > Num$
4. Form the vector $x = [x_1 ... x_n]$
5. where
6. $x_1 \in U(-(\gamma/\lambda_1)(1/2), (\gamma/\lambda_1)^{1/2})$; and
7. for $2 \leq i \leq n$, $x_i \in U(-(\tau_i/\lambda_i)^{1/2}, (\tau_i/\lambda_i)^{1/2})$,
8. with $\tau_i = (\gamma - \lambda_1 x_1^2 - ... - \lambda_{i-1} x_{i-1}^2)$.
9. $\mu = Lx$.
10. l=l+1.

Remark: Method B does not involve a verification criterion in Stage II as required in method A; hence it can quickly achieve the objective as long as the number of false validated measurements needed is known.

4 Illustrative example

In this section, we present the results of a computer simulation of an example in order to demonstrate the merits of the proposed method. Monte Carlo simulations have been carried out to evaluate the performance of a PDAF for target tracking in clutter by incorporating the two methods discussed in Section 3. The reader is referred to [1] for general background of a PDAF, as well as for details of the mathematical setup, on which we shall draw freely for the purpose of simulation. To compare the performance of the two methods, we examine 1)the computation time required to produce the result of actual tracking error and estimated error variance

, and 2) the accuracy of the estimated error variance to characterize the tracking error. Simulations have been carried out in a Matlab environment on a Sun workstation in the Department of Electrical Engineering at Royal Military College of Canada.

Assume there is a target with constant velocity. Consider a time-invariant kinematic model for the target.

$$x(k+1) =$$

$$\begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} x(k) + \begin{bmatrix} T^2/2 & 0 \\ T & 0 \\ 0 & T^2/2 \\ 0 & T \end{bmatrix} v(k)$$

$$z(k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} x(k) + w(k)$$

where T represents the sensor sampling period, and $v(k)$ and $w(k)$ are mutually independent white Gaussian noise vectors with zero mean and covariances.

$$Q = cov[v(k)] = \begin{bmatrix} q & 0 \\ 0 & q \end{bmatrix}$$

$$R = cov[w(k)] = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$$

This nearly constant-velocity model can be normalized by choosing $T = 1$. We will use this normalized model for our numerical example. Assume $\lambda = 0.01$, $P_D = 1$ (the target detection probability), $P_G = 0.9997$ (the probability that the target-oriented measurement, assuming the target was detected, falls inside the validation gate), q=0.15, r=1 and a 4-sigma confidence ellipse is chosen for the validation gate in our numerical example.

The initial state for the target is set as $x_t(0) = [10; 1; 50; 1.3]$ where the subscript t refers to the target, and the initial state for the estimation based on the above model is set as as $\hat{x}(0) = x_t(0) + [randn(1)*\sigma_{x1}; randn(1)*\sigma_{x2}; randn(1)*\sigma_{y1}; randn(1)*\sigma_{y2}]$; where σ_{x1}=0.01; σ_{x2}=0.001 σ_{y1}=0.01; σ_{y2}=0.001. The initial error covariance is

$$P(0) = \begin{bmatrix} \sigma_{x1}^2 & 0 & 0 & 0 \\ 0 & \sigma_{x2}^2 & 0 & 0 \\ 0 & 0 & \sigma_{y1}^2 & 0 \\ 0 & 0 & 0 & \sigma_{y2}^2 \end{bmatrix}$$

Let us consider the performance criteria:
the actual average of position rms errors of Monte-Carlo runs $e_0(k) = (1/N \sum_{i=1}^{N} (\hat{x}_1(k) - x_{t,1}(k))^2 + (\hat{y}_1(k) - y_{t,1}(k))^2)^{1/2}$, and the average of position variances of Monte-Carlo runs $e_1(k) = (1/N \sum_{i=1}^{N} (P_{11}(k) + P_{33}(k))^{1/2}$. For this simulation, N has been chosen as 1000.

The simulation results of the PDAF algorithm incorporating method A are shown in figure 1. In this case, it took 4.6 hours to obtain the results of figure 1. Figure 2 shows the simulation results of the PDAF algorithm incorporating method B. It took 10 minutes to obtain the results of figure 2.

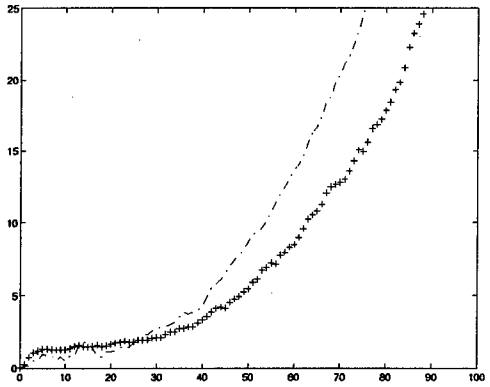


Figure 1: x-axis:time instant; y-axis:error; '-.' line displays $e_0(k)$; '+' line displays $e_1(k)$

Figures 1 and 2 show that $e_1(k)$ can reasonably approximate $e_0(k)$. However, to produce the results in figure 1 takes much longer than those exhibited in figure 2. It can be clearly observed that the divergence of tracking errors occurs in both cases. Furthermore, the actual and the estimated results in both cases are in general agreements, i.e., when the actual rms

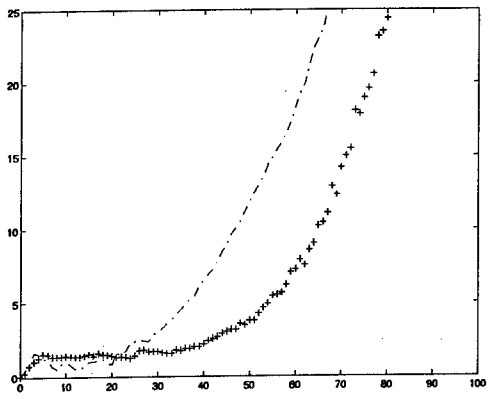


Figure 2: x-axis:time instant; y-axis:error; '-.' line displays $e_0(k)$; '+' line displays $e_1(k)$

error diverges so does the corresponding estimated rms error. The results can be attributed to the heavily cluttered environment due to $q = 0.15$ and can be explained as follows. Suppose there are many false validated measurements in a gate, then the state estimate may deteriorate and its error covariance may increase. Furthermore, this may lead to a larger gate size and more false measurements are likely to fall in the gate at the next time step. As a result, the tracking errors increase as time progresses. In figures 1 and 2, the divergence of actual rms tracking errors takes place at time instant around 25. The divergence of estimated covariance error occurs at time instant around 30 in figure 1, while that of estimated covariance error happens at time instant around 40 in figure 2. Though figure 1 shows better accuracy in the transient region than figure 2, but on the whole, the simulation results indicate that the proposed method is superior in satisfying the assumptions without involving heavy computational loads.

5 Conclusions

We have presented an efficient method for uniformly generating Poisson-distributed measurements in a validation gate. Unlike the approach in [7], the proposed method is computationally inexpensive because it can achieve the objective without running into an indefinite number of iterations. Apart from the structure of the method itself, we have employed a tracking example to demonstrate that a PDAF algorithm incorporating the proposed method in Monte Carlo simulations efficiently and satisfactorily characterizes statistics of tracking errors. In light of the result of the example, the proposed method provides an effective alternative with less computation burden.

Acknowledgement

The authors would like to thank Prof. X. Rong Li for valuable discussion on some aspect of simulations related to this work.

References

[1] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association, New York:Academic*,1988.

[2] S. S. Blackman, *Multiple-Target Tracking with Radar Application, Norwood, MA:Artech House*, 1986.

[3] P. Bratley et.al, *A Guide to Simulation, New York:Springer-Verlag*, 1983.

[4] S. H. Friedberg et. al., *Linear Algebra, Englewood Cliffs:Prentice-Hall, Inc.*, 1979.

[5] T.-J. Ho and M. Farooq, "On the stability of the JPDAF for multiple-target tracking in clutter," *Proceedings of the International Conference on Multisource-Multisensor Information Fusion*, pp.656-661, 1998.

[6] X. R. Li and Y. Bar-Shalom, "Stability evaluation and track life of the PDAF for tracking in clutter," *IEEE Trans. on Automatic Control*, Vol.36, pp.588-602, No.5, May 1991.

[7] X. R. Li, "Generation of random points uniformly distributed in Hyperellipsoids," *Proceedings of IEEE Conference on Control and Decision*, pp.654-658, 1992.

Passive Ranging of a Low Observable Ballistic Missile in a Gravitational Field Using a Single Sensor*

T. Kirubarajan, Y. Bar-Shalom and Yueyong Wang
Electrical and Systems Engineering Department
University of Connecticut
Storrs CT 06269-2157, USA.
ybs@ee.uconn.edu

Abstract *In this paper we present an estimation algorithm for tracking the motion of a low observable target in a gravitational field, for example, an incoming Ballistic Missile, using angle-only measurements. The measurements, which are obtained from a single stationary sensor, are available only for a short time. Also, the target detection probability is low and the false alarm density is high. The algorithm uses the Probabilistic Data Association algorithm in conjunction with Maximum Likelihood estimation to handle the false alarms and the less-than-unity target detection probability. The algorithm also uses the strength of the signals, or Amplitude Information. In addition to the PDA-AI/ML estimator, the Cramer-Rao Lower Bound in clutter is also presented. It is shown that for a ballistic missile in free flight with 0.6 single-scan detection probability, one can achieve a track detection probability of 0.99 with a negligible probability of false track acceptance even at 6dB SNR.*

Keywords: Angle-only tracking, target motion analysis and estimation, probabilistic data association, ballistic missile tracking, low-observable targets.

1 Introduction

A number of tracking algorithms that use radar measurements have been developed for effective defense against tactical ballistic missiles. Various estimators based on the Extended Kalman Filter (EKF) for the reentry phase were implemented in [5] and [7]. However, the EKF, which operates in a recursive manner, would require a high signal-to-noise ratio (SNR) to yield acceptable results. In [8] an algorithm was given for acquisition of low observable ballistic missiles using an electronically scanned array (ESA) radar. An optimal ballistic missile track initiation algorithm based on the Maximum Likelihood (ML) estimator using midcourse observations from passive sensors was presented in [9]. Unlike the case of a radar, in passive localization (from angle-only measurements) the measured range of target is not available, making observability (the ability to estimate the full state of the target) a crucial problem. Passive target tracking in an underwater environment, which is commonly referred to as target motion analysis (TMA) is a widely studied estimation problem of both theoretical and practical interest [2], whereas only a few results about problems in passive ranging of ballistic missiles have been reported [9].

The flight of a Ballistic Missile (BM) consists of three phases namely, boost phase, ballistic phase (midcourse or free-flight phase, in a plane in Earth Centered Inertial (ECI) coordinates) and terminal phase (reentry phase). The passive ranging (also referred to as passive localization) for a BM considered here is to detect and initiate the track before the missile

*Research sponsored by ONR/BMDO Grant N00014-91-J-1950, AFOSR Grant 49620-97-1-0198 and ONR Grant N00014-97-1-0502.

enters the terminal phase using the angle-only measurements from *a single stationary sensor.* The motion of such a ballistic missile is characterized as a free flight in a gravitational field [6]. In this paper we show that it is possible to have complete observability of this motion from a single stationary sensor. Another major concern in such a defense system is that the measurements are available only for a short period of time and the estimator has to obtain an acceptable estimate using those measurements.

To account for the measurement origin uncertainty, the approach called Probabilistic Data Association (PDA) [2], associates probabilistically all the possible measurements to the target of interest. The incorporation of feature measurements in addition to the angle-only measurements into the PDA technique enhances the estimator performance [4]. The feature measurement used in this paper is the measurement amplitude or amplitude information (AI), which is the intensity of the signal at the output of the signal processor. The PDA approach in conjunction with ML, based on a batch of angle and AI measurements, is developed to obtain the track estimate in this paper. The Cramer-Rao Lower Bound (CRLB), which has to incorporate the effect of the false alarms (clutter) and the less-than-unity detection probability, can quantify such a problem's "estimability."

In Section 2 we present the target, sensor and measurement models. In Section 3 we derive the ML estimator based on PDA combined with AI. The numerical implementation of the estimator is also presented. The CRLB in clutter and proof of target observability are given in this section. When an estimate is obtained, a validation procedure is carried out to check if this estimate is acceptable. This is necessary due to the possible local minima. In Section 4, simulation results are presented.

2 Problem Formulation

The focus of this work is to track the motion of a free-flight target, for example, a ballistic missile, using measurements obtained by a passive sensor over a short period of time. This is equivalent to estimating the initial state, namely, the position and velocity of that target.

In this section we present the target motion model, then SNR models of both the target-originated and false measurements is given. Finally, the observability problem is addressed.

2.1 Dynamic Model of Ballistic Missile Flight

For track formation and track extension, it has been common to model the missile motion as a simple quadratic polynomial in each dimension. To achieve maximum range for a given payload, it is very common to maintain a small angle of attack throughout the flight. Nevertheless, the quadratic model falls far short of the mark in modeling the trajectory. However, for limited functions such as track formation, it is usually accurate enough to model short segments within a given missile stage [6]. Satellite surveillance of ballistic missile launches will provide a timely report of each occurrence of a missile launch and launch parameters (missile type, launch time, launch position and heading) [6]. Since the motion of the missile occurs in a plane in ECI coordinates, in this paper we consider a two-dimensinal problem where the angle measurements consist of only the elevation and the sensor is in the same plane as the target.

We assume that n sets of measurements, made at times $t = t_1, t_2, \ldots, t_n$, are available.

For trajectory estimation, the target motion is defined by the 4-dimensional vector

$$x \triangleq \left[\begin{array}{cccc} \xi_0 & \eta_0 & \dot{\xi}_0 & \dot{\eta}_0 \end{array} \right]' \qquad (1)$$

where ξ_0 and η_0 are the coordinates of the target in the vertical and horizontal directions, respectively, at the reference time t_0. The corresponding velocities at time t_0 are $\dot{\xi}$ and $\dot{\eta}$, respectively. The incoming target in a gravitational field keeps moving with its initial velocity and with a known downward acceleration, namely g.

The true elevation angle of the target from the platform (assumed to be located at the origin of the coordinate system) at t_i is given by

$$\theta_i(x) \triangleq f(t_i, x) = \tan^{-1}[\xi(t_i, x)/\eta(t_i, x)] \quad (2)$$

where

$$\xi(t_i, x) = \xi_0 - \dot{\xi}_0\, t_i - 0.5 g t_i^2 \quad (3)$$

$$\eta(t_i, x) = \eta_0 - \dot{\eta}_0\, t_i \quad (4)$$

The possible elevation measurements are in the interval $[0, \pi]$. Since scanning this entire region is not practical, in [4] it was assumed that a cueing region within the field of view of the sensor is available as surveillance region

$$U_\theta \triangleq [\Theta_1, \Theta_2] \subset [0, \pi] \quad (5)$$

The set of measurements at t_i is denoted by

$$Z(i) \triangleq \{z_j(i)\}_{j=1}^{m_i} \quad (6)$$

where m_i is the number of measurements at t_i and the pair of elevation and amplitude measurements $z_j(i)$ is defined by

$$z_j(i) \triangleq [\beta_{ij} \quad R_{ij}]' \quad (7)$$

The cumulative set of measurements during the entire period is

$$Z^n \triangleq \{Z(i)\}_{i=1}^{n} \quad (8)$$

In addition to the above, the following assumptions about the statistical characteristics of the measurements are also made [3]:

1. The measurements at two different sampling instants are conditionally independent, i.e.,

$$p[Z(i_1), Z(i_2)|x] = p[Z(i_1)|x]\; p[Z(i_2)|x] \quad \forall i_1 \neq i_2 \quad (9)$$

where $p[\cdot]$ is the probability density function.

2. A measurement that originated from the target at a particular sampling instant is received by the sensor only once during the corresponding scan with probability P_D and is corrupted by zero-mean additive Gaussian noise with known variance. That is,

$$\beta_{ij} = \theta_i(x) + \epsilon_{ij} \quad (10)$$

where $\epsilon_{ij} \sim \mathcal{N}[0, \sigma_\theta^2]$ is the elevation measurement noise. Due to the presence of false measurements, the index j of the true measurement is not known.

3. The false measurements are distributed uniformly in the surveillance region, i.e.,

$$\beta_{ij} \sim \mathcal{U}[\Theta_1, \Theta_2] \quad (11)$$

4. The number of false measurements at a sampling instant is generated according to a Poisson law with a known expected number of false measurements in the surveillance region [2]. This is determined by the detection threshold at the sensor (the exact equations are given in Section 4).

2.2 SNR Models

We denote by R the signal-plus-noise to noise ratio (SNNR), which is different from the signal-to-noise ratio[1] (SNR), for example, an SNNR value of 7dB corresponds to 6dB SNR [8]. With the noise power normalized to unity, R is then the intensity of the output of the signal processor, consisting of noise only, or target-originated signal plus noise.

The probability density function (pdf) of R when the signal is due to noise only is denoted by $p_0(R)$ and the corresponding pdf when the signal originated from the target is $p_1(R)$. If the average signal-to-noise ratio (SNR) is $\overline{E}$, the pdf of noise-due and target-originated measurements can be written as

[1] The SNR is defined as $10 \log_{10}(\frac{A^2}{\sigma^2})$, where A is the signal amplitude and σ^2 is the variance of noise.

$$p_0(R) = \exp(-R), \quad R \geq 0 \tag{12}$$

$$p_1(R) = \frac{4}{(2+\overline{E})^2} \exp\left(-R + \frac{\overline{E}R}{2+\overline{E}}\right) \left(1 + \frac{\overline{E}R}{2+\overline{E}}\right), \quad R \geq 0 \tag{13}$$

respectively. Here p_1 is a Swerling III model which is believed to be appropriate for ballistic missiles [8]. Note that the noise power in (12) is normalized to unity.

A suitable threshold (low, because we are dealing with low SNR), denoted by τ, is used to declare a detection. Both the probability of detection, P_D and the probability of false alarm P_{FA} (defined for a resolution cell) can be evaluated from the probability density functions of the measurement amplitudes. They are given by

$$P_D = \int_\tau^\infty p_1(R)dR \tag{14}$$

$$P_{FA} = \int_\tau^\infty p_0(R)dR \tag{15}$$

Clearly, in order to increase P_D one has to lower the threshold τ. However, this increases P_{FA} too. Therefore, depending on the SNR we have to select τ so as to compromise between two conflicting requirements.

The density functions given above correspond to the signal at the signal processor output. Those corresponding to the output of the threshold detector are truncated versions of the previous pdfs [8]

$$p_0^\tau(R) = \frac{1}{P_{FA}} \exp(-R), \quad R > \tau \tag{16}$$

$$p_1^\tau(R) = \frac{4}{P_D(2+\overline{E})^2} \exp\left(-R + \frac{\overline{E}R}{2+\overline{E}}\right) \left(1 + \frac{\overline{E}R}{2+\overline{E}}\right), \quad R > \tau \tag{17}$$

where $p_0^\tau(R)$ is the pdf of the amplitudes of the measurements that are due to noise only and $p_1^\tau(R)$ is the pdf of those originated from the target.

Finally, we define the amplitude likelihood ratio ρ, which will be used in the derivation of the estimator, as

$$\rho = \frac{p_1^\tau(R)}{p_0^\tau(R)} \tag{18}$$

3 ML/PDA Estimator

The detections at a sampling instant consist of a number of false measurements and, at most, one target-originated measurement. Even if the target-originated measurement is detected, it cannot be distinguished from the false ones, and thus there is no single measurement that can be used to accurately estimate the target state. In order to resolve this data association problem, an ML estimator based on the PDA technique, which uses *all the measurements in a scan*, is presented next.

3.1 PDA-AI/ML Estimator

If there are m_i detections at t_i, we have the following mutually exclusive and exhaustive events [2]:

$$\varepsilon_j(i) \triangleq \begin{cases} \{z_j(i) \text{ is from target}\}, j = 1, \ldots, m_i \\ \{\text{all measurements are false}\}, j = 0 \end{cases} \tag{19}$$

The pdf of the measurements (6) conditioned on the above events can be written as

$$p\left(Z(i)|\varepsilon_j(i), x\right) = \begin{cases} u^{1-m_i} p(\beta_{ij})\rho_{ij} \\ \quad \prod_{l=1, l\neq j}^{m_i} p_0^\tau(R_{il}), j \neq 0 \\ u^{-m_i} \prod_{l=1}^{m_i} p_0^\tau(R_{il}), \\ \quad j = 0 \end{cases} \tag{20}$$

where $u = U_\theta$ is the area of the surveillance region.

Using the total probability theorem, we can write the likelihood function of x at t_i as

$$p\left[Z(i)|x\right] = u^{-m_i}(1 - P_D) \prod_{j=1}^{m_i} p_0^\tau(R_{ij}) \mu_f(m_i) + \frac{u^{1-m_i} P_D \mu_f(m_i - 1)}{m_i} \prod_{j=1}^{m_i} p_0^\tau(R_{ij}) \sum_{j=1}^{m_i} p(\beta_{ij})\rho_{ij} \tag{21}$$

where $\mu_f(m_i)$ is the Poisson probability mass function (pmf) of the number of false measurements at t_i.

Dividing the above by $p[Z(i)|\varepsilon_0(i), x] \cdot \mu_f(m_i)$ we obtain the *dimensionless* likelihood ratio[2] $\Phi_i[Z(i), x]$ at t_i. Then

$$\begin{aligned}\Phi_i[Z(i), x] &= \frac{p[Z(i)|x]}{p[Z(i)|\varepsilon_0(i), x] \cdot \mu_f(m_i)} \\ &= (1 - P_D) + \frac{P_D}{\lambda} \cdot \\ &\sum_{j=1}^{m_i} \frac{1}{\sqrt{2\pi}\sigma_\theta} \rho_{ij} \exp\left(-\frac{1}{2}\left[\frac{\beta_{ij} - \theta_i(x)}{\sigma_\theta}\right]^2\right) \quad (22)\end{aligned}$$

where λ is the expected number of false alarms per unit area.

Alternatively, we can define the log-likelihood ratio $\phi_i[Z(i), x]$ at t_i as

$$\begin{aligned}\phi_i[Z(i), x] &= \ln\left\{\frac{p[Z(i), x]}{p[Z(i)|\varepsilon_0(i), x]}\right\} \\ &= \ln\left[(1 - P_D) + \frac{P_D}{\lambda} \cdot \right. \\ &\left. \sum_{j=1}^{m_i} \frac{1}{\sqrt{2\pi}\sigma_\theta} \rho_{ij} \exp\left(-\frac{1}{2}\left[\frac{\beta_{ij} - \theta_i(x)}{\sigma_\theta}\right]^2\right)\right] \quad (23)\end{aligned}$$

Using the conditional independence of measurements, the likelihood function of x based on the entire set of measurements can be written in terms of the individual likelihood functions as

$$p[Z^n|x] = \prod_{i=1}^{n} p[Z(i)|x] \quad (24)$$

Then the dimensionless likelihood ratio based on the entire data is given by

$$\Phi[Z^n, x] = \prod_{i=1}^{n} \Phi_i[Z(i), x] \quad (25)$$

From the above, one can write the total log-likelihood ratio $\phi[Z^n, x]$ as

$$\begin{aligned}\phi[Z^n, x] &= \sum_{i=1}^{n} \phi_i[Z(i), x] \\ &= \sum_{i=1}^{n} \ln\left[(1 - P_D) + \frac{P_D}{\lambda} \right. \\ &\left. \sum_{j=1}^{m_i} \frac{1}{\sqrt{2\pi}\sigma_\theta} \rho_{ij} \exp\left(-\frac{1}{2}\left[\frac{\beta_{ij} - \theta_i(x)}{\sigma_\theta}\right]^2\right)\right] \quad (26)\end{aligned}$$

The Maximum Likelihood Estimate (MLE) is obtained by finding the vector $\hat{x}$ that maximizes the above total log-likelihood function.

4 Simulation Results

In this section, we consider a two-dimensional scenario, where the target SNR is as low as 6dB, to illustrate the operation and performance of the PDA-AI/ML estimator. Simulation results are obtained using 100 Monte Carlo runs with the following scenario: The missile enters the sensor surveillance region at $t_0 = 0s$ with initial position $P_0^{\text{true}} = (10^5 m, 10^5 m)$, and initial velocity $v_0^{\text{true}} = (1500 m/s, 1500 m/s)$ and the sensor platform remains stationary at $(0m, 0m)$. It is assumed that the target flight course and the sensor are co-planar, i.e., this is a two dimensional tracking problem.

The sensor's angular aperture is assumed to be $1 \times 10^5 \mu$rad, which consisting of 2000 cells, each of size $C_\theta = 50\mu$rad. Assuming uniform distribution in a cell, the standard deviation of angle measurements is given by $\sigma_\theta = 50/\sqrt{12} = 14.4\mu$rad. The term $\overline{E}$ in equation (13) was taken as 6dB and $P_D = 0.6$ in equation (14). For the given values of $\overline{E}$ and P_D, equations (14) and (15) give the detection threshold $\tau = 3.0866$ and the probability of false alarm in a cell $P_{FA} = 0.0457$.

The expected number of false alarms per unit angle can be calculated as

$$\begin{aligned}\lambda &= \frac{P_{FA}}{\text{volume of angle cell}} \\ &= \frac{0.0475}{50 \times 10^{-6}} \\ &= 914/\text{rad}. \quad (27)\end{aligned}$$

We can also calculate the expected number of false alarms in the entire sensor aperture angle, and the value is found to be 95. This

[2]This normalization is convenient, since the numbers of detections at each scan may be different.

value means that at every sampling time[3] there are about 95 false alarms which exceed the the threshold.

The scans are made at 10Hz for 10s. Fig. 1 shows sample sets amplitude measurements in one run of the Monte Carlo simulations. The target-originated and noise-only measurements are denoted by "." and "*", respectively. However, note that the index of the target-originated measurement is not known to the estimator. It can be seen that target-originated measurements are detected in 55 times out of 100 scans.

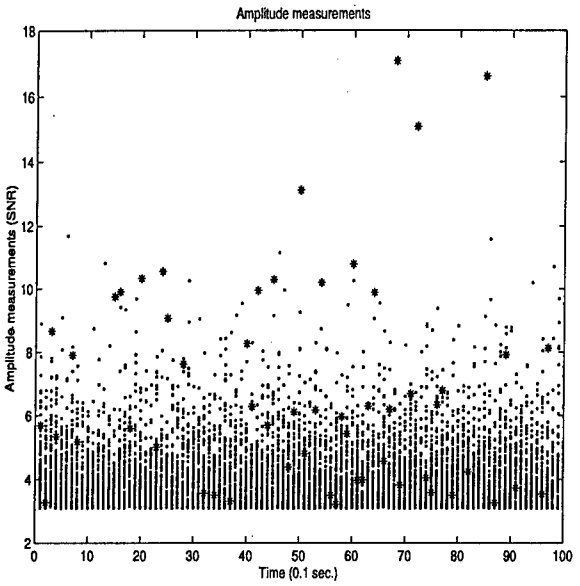


Fig. 1. Amplitude measurements ("*" — target-originated; "." — false alarms)

For the above measurements, the variation of the negative log-likelihood function with position is shown in Fig. 2. It can be seen that the global minimum is located in a narrow valley around (10^5m,10^5m), which makes it difficult to find using numerical techniques. For initialization, a systematic grid search is performed to find an approximate minimum point to start off the quasi-Newton minimization. The grid search procedure is shown in Table 1, where r is the target range and θ is the elevation of the target. These evaluations on the grid points were the most costly part of the numerical calculations – they took up 99.5% of the total time per run, which was 12.2min on a Pentium 400 processor. However, they are parallelizable with a SIMD architecture with linear efficiency. With 200 processors, the total computation time would be 6s.

[3]We assume a staring sensor with all the measurements in frame having the same time tag. The procedure can be modified for a scanning sensor.

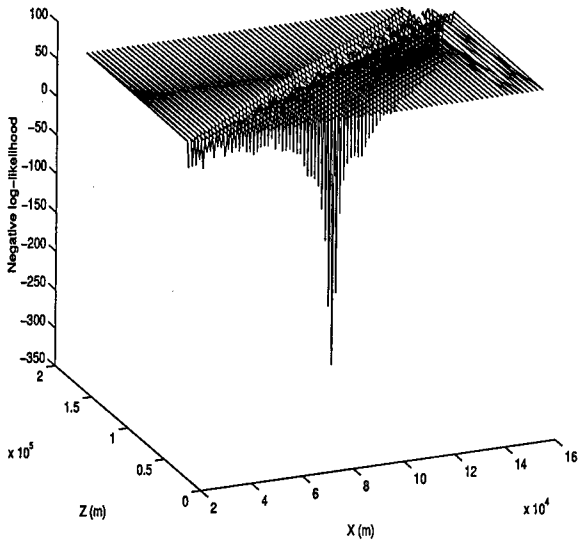


Fig. 2. Variation of negative log-likelihood

Param.	Region	Step-size
r (m)	$10^5 - 3 \times 10^5$	5×10^4
θ (mrad.)	735 – 835	0.2
$\dot{\xi}_0$ (m/s)	300 – 3000	300
$\dot{\eta}_0$ (m/s)	300 – 3000	300

Table 1. Grid search for minimization (25×10^4 points)

The tracks obtained by the maximization are validated using a hypothesis testing technique [4]. The track acceptance threshold τ_1 was set so that the tracks are accepted with 99% probability (the track acquisition probability P_{ACQ}). All the tracks obtained with the maximization procedure were accepted as valid tracks. In Table 2, the average position and velocity estimates $\overline{x}$ and their corresponding standard deviation $\hat{\sigma}$ are given over 100

Monte Carlo runs. The theoretical standard deviations σ_{CRLB} for the scenario under consideration are given for each component.

Param.	x_{true}	$\overline{x}$	$\hat{\sigma}$	σ_{CRLB}
ξ_0 (m)	100,000	100,480	4487	5873
η_0 (m)	100,000	100,480	4489	5875
$\dot{\xi}_0$ (m/s)	1500	1535	310	398
$\dot{\eta}_0$ (m/s)	1500	1535	318	399

Table 2. Results of 100 Monte Carlo runs

In order to check the efficiency of the estimator we need to check its consistancy with the FIM. This is performed by finding the average normalized estimation error squared (NEES) [1] and checking whether it falls within the statistical bounds for acceptance. The NEES is defined as

$$\alpha = [\hat{x}_0 - x_0^{true}]' P^{-1} [\hat{x}_0 - x_0^{true}] \qquad (28)$$

where $P^{-1} = J$ is the FIM. If the estimator is unbiased and the errors are Gaussian with covariance equal to the CRLB, then α defined above is chi-squared distributed with n_x (i.e., 4 in our problem) degrees of freedom. Taking the average over N Monte Carlo runs, the 95% probability bounds on $\bar{\alpha}$ are

$$\frac{\chi^2_{4N}(0.025)}{N} \leq \bar{\alpha} = \frac{1}{N}\sum_{i=1}^{N} \alpha_i \leq \frac{\chi^2_{4N}(0.975)}{N} \qquad (29)$$

where α_i is the NEES in the i^{th} Monte-Carlo run. If the filter is inefficient or biased then $\bar{\alpha}$ will lie above the upper bound. In our simulation the average value of NEES for the accepted tracks is found to be 3.48, which is within the 95% bound [1].

We also carried out a comparison of the best negative log-likelihood ratios (score-of-goodness) the PDA-AI/ML estimator could find in the target-present and target-absent scenarios. In order to get accurate results, we constructed target-absent scenarios by eliminating the corresponding target-present scenario's target-originated measurements. Fig. 3 shows 100 runs' comparison, where target-absent and target-present are denoted by "." and "+", respectively. Among the tracks obtained in the target-absent scenarios none was accepted as a valid track by the acceptance test. Fig. 3 shows that the negative log-likelihood ratios are remarkably well separated between the two types of scenarios. It can be seen that in the target-absent case the negative log-likelihood ratio is around 50, i.e., the estimated trajectories are e^{50} times more likely to come from noise than from a target and, thus, the validation test rejects them. Conversely, in the target-present scenarios the estimated trajectories are e^{220} more likely to be target-originated than noise originated and the test accepts them.

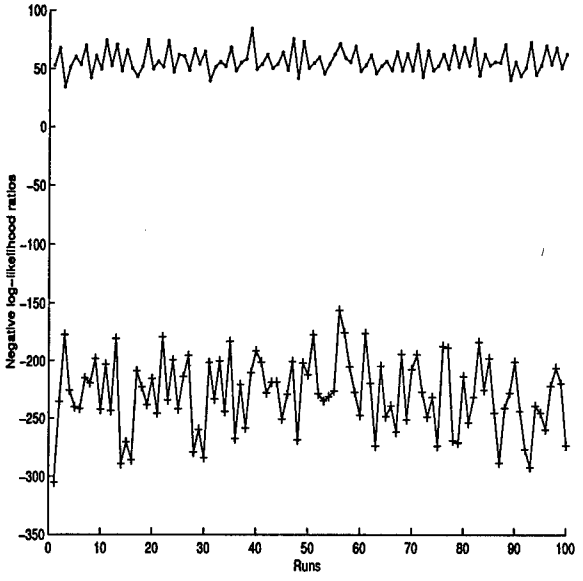


Fig. 3 Best negative log-likelihood ratios ("." target absent; "+" target present)

5 Conclusions

In this paper we presented a PDA-AI/ML algorithm for detecting the track of a low observable target in a gravitational field using angle-only measurements. The measurements, which

are obtained from a single stationary sensor, are available only for a short time. Also, the low target detection probability and high false alarm density present a difficult low-observable environment to work with. The algorithm uses the Probabilistic Data Association (PDA) algorithm in conjunction with Maximum Likelihood (ML) estimation to handle the false alarms and the less-than-unity target detection probability. The algorithm also uses the strength of the signals, or Amplitude Information (AI), modeled as Swerling III type, in the tracking process itself, in addition to using it for thresholding. This is achieved by deriving a combined likelihood based on the angle measurements and the AI, which is then maximized using a parallelizable numerical search.

In addition to the PDA-AI/ML estimator, the Cramer-Rao Lower Bound (CRLB) in clutter is also presented. The proposed estimator is shown to be efficient, that is, it meets the CRLB, even for low-observable fluctuating targets with 6dB average signal-to-noise ratio (SNR). At this SNR, the target detection probability is 0.6 and the expected number of false alarms is 95 per scan. A hypothesis testing-based track validation (track detection) scheme, which confirms the estimated tracks, is also presented in this paper. For the ballistic missile in free flight with 0.6 single-scan detection probability, one can achieve a track detection probability of 0.99 with negligible probability of false track acceptance. The proposed algorithm, which operates in batch mode, can also be used to obtain an initial estimate for a recursive or sliding-window based estimator.

Acknowledgements

A stimulating discussion with H. Weiss is gratefully acknowledged.

References

[1] Y. Bar-Shalom and X. R. Li, *Estimation and Tracking: Principles, Techniques and Software*, Boston, MA: Artech House, 1993. Reprinted by YBS Publishing, 1998.

[2] Y. Bar-Shalom and X. R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*, Storrs (CT 06269-2157): YBS Publishing, 1995.

[3] C. Jauffret and Y. Bar-Shalom, "Track formation with bearing and frequency measurements in clutter", *IEEE Trans. Aerospace and Electronic Systems,* AES-26, (Nov. 1990), 999-1010.

[4] T. Kirubarajan and Y. Bar-Shalom, "Low observable target motion analysis using amplitude information", *IEEE Trans. Aerospace and Electronic Systems,* AES 32-4, (October 1996), 1367-1384.

[5] R. K. Mehra, "A comparison of several nonlinear filters for reentry vehicle tracking", *IEEE Trans. Automatic Control,* AC 16-6, (August 1971), 307-319.

[6] J. G. Rudd, R. A. Marsh and J. A. Roecker, "Surveillance and tracking of ballistic missile launches", *IBM Res. Develop.* Vol.38 No.2, (March 1994).

[7] G. M. Siouris, G. Chen and J. Wang, "Tracking an incoming ballistic missile using an extended interval Kalman filter", *IEEE Trans. Aerospace and Electronic Systems,* AES 33-1, (January 1997), 232-240.

[8] S. Sivananthan, K. Kirubarajan and Y. Bar-Shalom, "A radar power multiplier algorithm for acquisition of low observable ballistic missiles using an electronically scanned array radar", *Proc. 1999 IEEE Aerospace Conference, Snowmass, CO, March 1999*

[9] M. Yeddanapudi, Y. Bar-Shalom, K. R. Pattipati and S. Deb, "Ballistic missile track initiation from satellite observations", *IEEE Trans. Aerospace and Electronic Systems,* AES 31-3, (July 1997), 1054-1071.

Bayesian Networks for Target Identification and Attribute Fusion with JPDA

P. Korpisaari and J. Saarinen
Signal Processing Laboratory
Tampere University of Technology
P.O.Box 553
33101 Tampere
Finland
Email: pkorpi@cs.tut.fi
Tel: +358 3 3653392
Fax: +358 3 3653095

Abstract An observation-to-target association problem is addressed in this paper. The Joint Probabilistic Data Association method (JPDA) is expressed with the concepts of Bayesian networks. This is done by assigning a manydimensional association vector to a root variable of the network. We propose a generalized network structure which enables applying attribute information in the context of JPDA. A case example shows how this approach can be used for target identification and target tracking purposes.

Keywords: JPDA , Bayesian networks, Attribute fusion

1 Introduction

Modern sensors produce measurements that contain attribute and kinematic information. Kinematic information describes the position-related state of the targets. Attribute information describes pieces of information that can be used for target identification. Attribute information is usually at different levels of abstraction which makes the identification procedure difficult.

Data association is a key problem in multitarget tracking. A well-known approach is to apply Joint Probabilistic Data Association method (JPDA) [2] to resolve association ambiquities in the case of closely located multiple targets. JPDA defines a set of feasible events that are used as hypotheses to explain the association event under consideration. We express the JPDA in a special form of Bayesian networks. This enables a direct extension to fuse attribute information in the framework provided by JPDA. Set of feasible events are illustrated by an association matrix. This matrix defines discrete-valued association vectors which describe feasible association events. The association vector is defined at each time instant for validated measurements. Finally, the association vector is used as a manydimensional root variable in the Bayesian network which contains target's kinematic state vector as a leaf variable. Thus, kinematic measurements induce a joint probability distribution for feasible association vectors in the root variable. This probability distribution can be used for computing measurement-to-target probabilities.

We apply hierarchical attribute structure to describe their internal depencies. Attribute hierarchy is implemented as a singly-connected Bayesian network. This approach enables measurements at different levels of abstraction. Furthermore, this makes possible a use of incomplete attribute measurements. The attribute network is integrated into larger Bayesian network which performs JPDA data asociation procedure for measurements containing both kinematic and at-

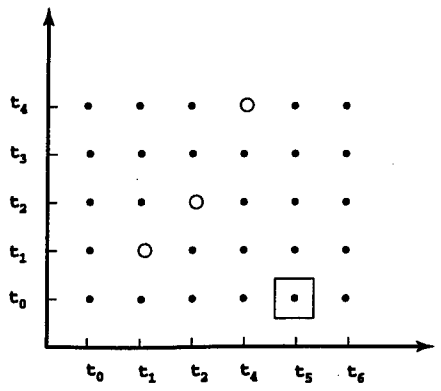


Figure 1: Association vector $\underline{\boldsymbol{\theta}}$.

tribute information.

The proposed generalized association network performs the data association task in the cluttered multitarget case. The method enables fusion of attribute and kinematic information. Additionally, the method allows hidden variables that are unobservable but are used for explaining the attribute dependencies.

2 Joint Probabilistic Data Association

Joint Probabilistic Data Association (JPDA) [1] is a well-known method for resolving association ambiquities in the case of multiple targets. We list here briefly the main principles of the algorithm.

Let $\mathcal{T}$ denote a set of possible sources containing false alarm and T targets.

$$\mathcal{T} = \{t_i; i = 0, \ldots, T\} \qquad (1)$$

where t_0 refers to false alarm hypothesis and t_j $(j = 1 \ldots T)$ refers to jth target.

2.1 Feasible events

An association event describes one particular way to correlate measurements to tracks. Feasible event is an association event which fulfills the following conditions [2]:

- A measurement z_i can be originated only from one source t_j $(j = 0 \ldots T)$.
- A target t_j $(j = 1 \ldots T)$ can not produce more than one measurement at time instant k.

Note that a number of detections produced by the false alarm t_0 is not limited.

2.2 Association vector

An association vector $\underline{\boldsymbol{\theta}}$ contains m elements. m equals to number of detections received at the processing time instant. $\underline{\boldsymbol{\theta}}$:s ith element θ_i indicates the source (target or false alarm) to which the ith measurement will be associated in that particular event.

$$\underline{\boldsymbol{\theta}} = [\theta_1, \theta_2, \ldots, \theta_m] \qquad (2)$$

Possible value of θ_i are false alarm t_0 and targets that are validated to detection z_i denoted as $\left\{t_i^1, t_i^2 \ldots t_i^l\right\} \subset \mathcal{T}$ where l is a number of validated targets. Thus, the association vector is defined on the following grid:

$$\Omega = \left\{t_0, t_1^1, t_1^2 \ldots t_1^{l_1}\right\} \times \left\{t_0, t_2^1, t_2^2 \ldots t_2^{l_2}\right\} \times \ldots \\ \ldots \times \left\{t_0, t_m^1, t_m^2 \ldots t_m^{l_m}\right\} \qquad (3)$$

The first condition of feasible events is directly satisfied since for a given vector $\underline{\boldsymbol{\theta}}$ only one value is assigned to each component θ_i. The second condition implies that several association points in Ω are not feasible and thus impossible. These points correspond to events where one target t_j $(j \neq 0)$ would be associated to more than one detection. This condition is fulfilled if all nonzero elements of $\underline{\boldsymbol{\theta}}$ are unequal:

$$\theta_i \neq \theta_j \quad \text{if } \theta_i \neq 0 \text{ and } \theta_j \neq 0 \ \forall\, i \neq j \qquad (4)$$

A two-dimensional grid for two detections z_1 and z_2 is illustrated in figure 1. The feasible points in Ω are drawn as black solid points and infeasible points that do not satisfy the condition (4) are drawn with empty circles. The squared point corresponds to the association vector that defines the following detection-to-target association pairs: $z_1 \leftrightarrow t_5$ and $z_2 \leftrightarrow t_0$.

3 JPDA as Bayesian Networks

A key idea of JPDA is to calculate probabilities of joint association events. These probabilities are conditioned on the observations

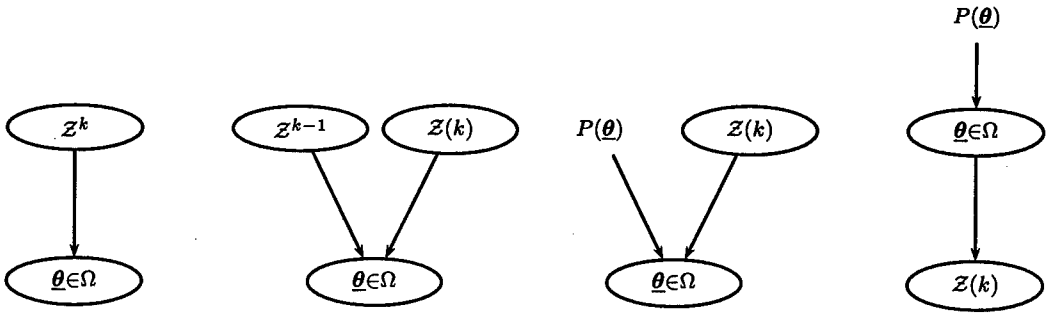


Figure 2: Basic principles of JPDA.

$\mathcal{Z}^k$. This is illustrated with the first graph in Fig. 2. $\mathcal{Z}^k$ is split into current observation set $\mathcal{Z}(k)$ and past observations $\mathcal{Z}^{k-1}$. These two observation sets are assumed to be independent. The past observations do not have direct impact to current association assignments. Such an influence which is not directly related to current observations is represented by association events' prior probabilities $P(\underline{\theta})$. Finally, an arc reversal operation produces the final graph in Fig. 2 which illustrates the basic principles of JPDA.

As described above the JPDA algorithm can be represented solely as a Bayesian network. A root variable of the network is m-dimensional association vector $\underline{\theta}$ which is defined on a grid Ω determined by the conditions of feasible events. Association vector's probability $p(\underline{\theta})$ is a product of its elements' probabilities. The elements θ_i may be represented as children of $\underline{\theta}$ as it is shown in Fig. 3. By its definition θ_i defines a cause for the detection $z_i(k)$. Such a causality is described in Bayesian networks by setting $z_i(k)$ to a direct child of θ_i. These two nodes are connected via conditional probability density function $p(z_i(k)|\theta_i)$ which in this case is continuous conditioned on the discrete variable θ_i. This connection is actually a collection of l_i+1 distinct probability density functions. l_i is a number of validated targets and the one extra pdf is due to false alarm hypothesis which is always present as a possible candidate for θ_i. The pdf corresponding to false alarm hypothesis is uniformly distributed and the others are normally distributed. Thus, JPDA is a Bayesian network which finally performs marginalization in order to determine the observation-to-target association probabilities from the determined $p(\underline{\theta})$.

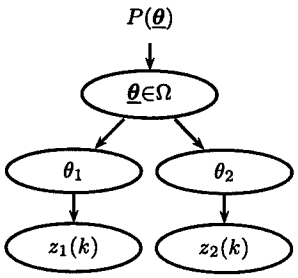


Figure 3: Bayesian networks for JPDA.

A link from position estimates and observations to JPDA is given by the conditional probability function $p(z_i(k)|\theta_i)$. In a similar way additional non-kinematic observations may be caused by the same source defined by θ_i. These observations are linked to θ_i also by a conditional probability function. We assume that the additional observation is an aircraft type. Since it is independent from position it can be represented as an additional children to θ_i. Both the association vector and the aircraft type are discrete and thus the conditional pdf between these two variables is purely discrete.

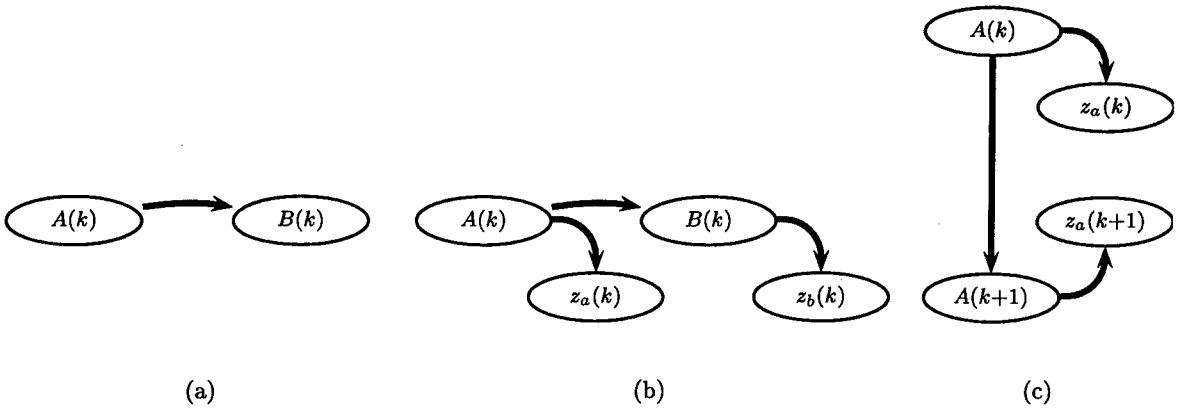


Figure 4: Hierarchical attribute structure:(a) dependent attributes (b) observations for dependent attributes (c) temporal dependency.

4 Attributes' causalities

4.1 Hierarchical attributes

An attribute may have a direct influence to another attribute, i.e. a value a_i of an attribute A affects to variable B's probability distribution function $p(B)$. Thus, B is a child of A in Bayesian networks. This kind of dependency relation is illustrated in Fig. 4(a). We model attributes' internal dependencies with a hierarchical tree which is essentially a singly connected Bayesian network. A root variable of the tree is aircraft type and all other variables are used to explain this variable. The proper estimation of the root variables value is a target identification part of the attribute fusion problem.

4.2 Sensor model

Another dependency that has to be taken into account is observation's relation to attribute's correct values. This is essentially a sensor modelling task. The detection's mixing matrix links observation to the corresponding attribute node. The link is implemented in a similar way as attributes' internal dependencies with conditional probability tables. However, it has a different interpretation from the semantical point of view. The sensor model is illustrated in Fig. 4(b).

4.3 State evolution model

A third model of dependency is a temporal causality between two adjacent values of the same attribute as illustrated in Fig. 4(c). In the case of position variable this dependency is modelled with linear dynamic equation and it is evaluated with Kalman filter. All attributes are assumed to remain constant.

5 Generalized association network

We extend the Bayesian networks performing the JPDA algorithm (Fig. 3) by adding an extra child to each θ_i node. This new child node is an aircraft type. As it was the case in positional detections this node is connected to the JPDA algorithm by the conditional probability function $p(z_a(k)|\theta_i)$. In order to define this probability we utilize the following causality chain. θ_i defines a target t_i which has its own target type estimate presented as a probability distribution $p(a|t_i)$. Target's correct aircraft type has a direct impact to the detected aircraft type distribution. This conditional pdf is denoted as $p(z_a(k)|a)$. Thus

$$p(z_a(k)|t_i) = p(z_a(k)|a)p(a|t_i)p(t_i) \quad (5)$$

Now, since t_i is actually a given value of θ_i the above equation can written as follows:

$$p(z_a(k)|\theta_i) = p(z_a(k)|a)p(a|\theta_i)p(\theta_i) \quad (6)$$

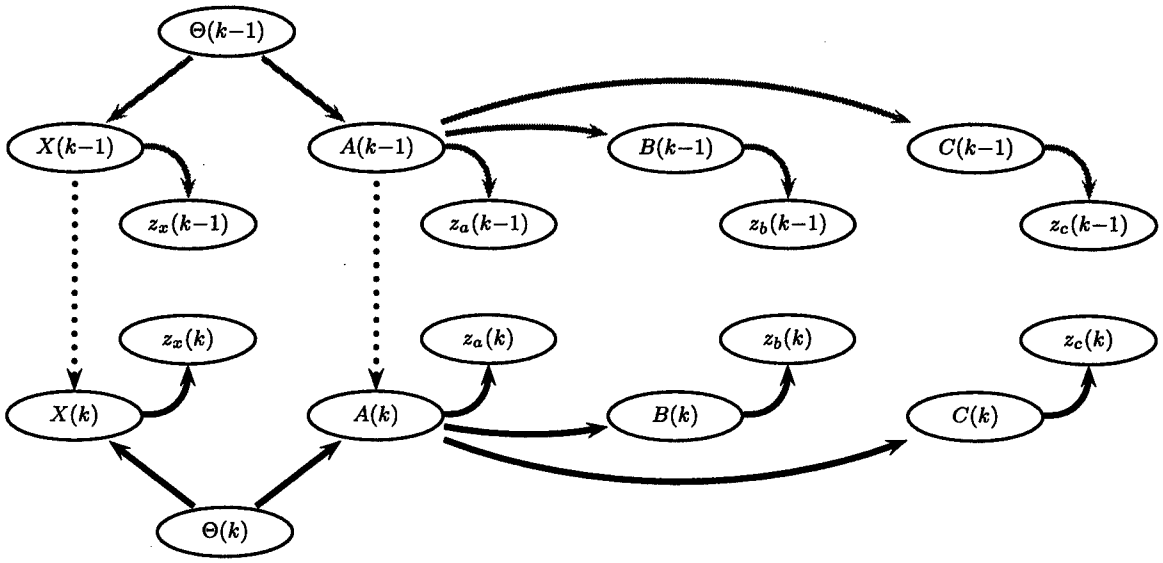


Figure 5: Generalized association network.

Probabilities $p(z_a(k)|a)$ are calculated by Bayesian formula as follows:

$$p(z_a(k)|a) = \frac{p(a|z_a(k))p(z_a(k))}{\sum_A p(a|z_a(k))p(z_a(k))} \quad (7)$$

Assuming that all aircraft type observations are equally likely apriori and applying the above formula yields:

$$p(z_a(k)|\theta_i) = \frac{p(a|z_a(k))}{\sum_A p(a|z_a(k))} p(a|\theta_i)p(\theta_i) \quad (8)$$

Moreover, we model the internal dependencies of different attributes in the context of JPDA algorithm. A hierarchical attribute tree is attached to aircraft type node. This tree and the possible attribute detections are used to determine the probability distribution function of aircraft type node based on the given attribute detections. This probability $p(a|z_a(k))$ is then applied to the Eq. 8. The sensor models are also attached to the network. Finally the tenporal dependencies are taken into account in such a way that the aircraft type probability distribution $P_A(k-1)$ is used as a priori distribution for the root variable of the attribute network. All these dependencies are illustrated in Fig. 5.

6 Data association in cluttered multitarget environment

Target tracking has the following steps for each time instant:

- Prediction
- Validation
- Association
- Correction

Prediction and correction are performed with Kalman filter. Validation is carried out based on the position information only. At the validation phase a new root variable $\underline{\boldsymbol{\theta}}$ will be initialized for the Bayesian network. This vector has the same number of components as is the number of received observations during the scan period. The validation defines the possible values for each component of $\underline{\boldsymbol{\theta}}$. Additionally a grid defining the set of feasible events will be set up during the validation phase. At the association stage the constructed network is evaluated and a joint probability distribution $p(\underline{\boldsymbol{\theta}})$ will be determined. The marginal distributions are determined for association probabilities. As these probabilities have been determined

the correction phase follows the basic equations of JPDA and Kalman filter.

7 Simulations

A multitarget tracking system that utilizes the generalized association network presented above has been implemented. The system is capable to fuse positional detections with attribute detections. The attribute hierarchical tree used in simulations is illustrated in Fig. 6. We present here two examples describing the central properties of the proposed system.The simulation was carried out with two crossing targets. The targets were of different aircraft type and thus the conditional probabilities in their attribute hierarchy trees were different. These probabilities have been used for both simulating the observation data and for reasoning purposes.

The two crossing targets used in the simulation are illustrated in Fig. 7. The Fig. 8 illustrates tracking results with JPDA algorithm that utilizes only the positional detections. In Fig. 9 the same situation has been tracked with generalized association network. The track switch that happens in the case of ordinary JPDA does not occur in the case of additional attribute information. This illustrates the system's capabilities to handle additional information and hence producing better tracking results.

Another property of the proposed system is that in addition to position tracking purposes, evidencies on attributes may be yield. For example, in Fig. 10 the evolution of target identification probabilities are presented. In this simulation we used four different aircraft types and the correct aircraft type of the target is illustrated with the bold line.

8 Conclusions

We presented a Bayesian network connected with Joint Probabilistic Data Association algorithm. The proposed method is capable to utilize different kind of attributes and use the additional information related to them in order to gain better tracking performancies. The network models attributes' internal

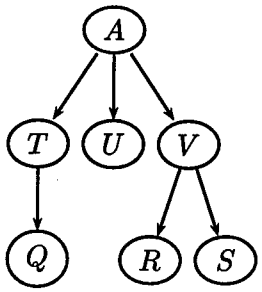


Figure 6: An attribute hierarchy used in simulations.

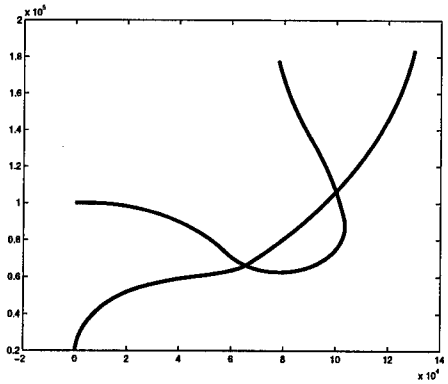


Figure 7: Two crossing targets.

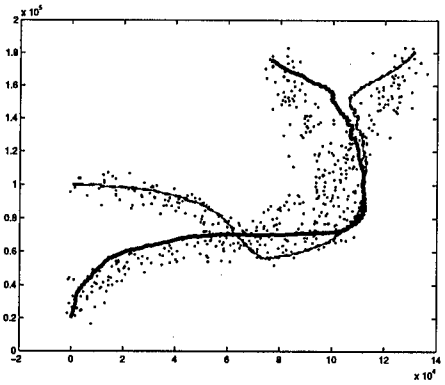


Figure 8: Track switch.

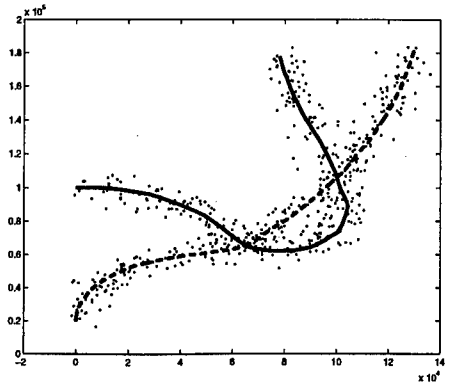


Figure 9: Non-switching tracks.

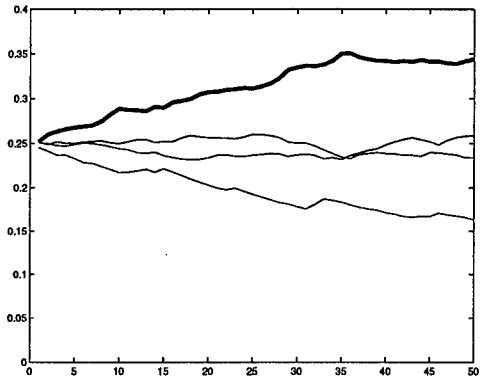


Figure 10: Type identification probabilities.

dependencies with a hierchical tree. It also contains the sensor models and gives relative easy basis to implement even more sophisticated dependencies.

References

[1] T. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Sonar tracking of multiple targets using joint probabilistic data association," *IEEE Journal of Oceanic Engineering*, vol. 8, pp. 173–184, July 1983.

[2] Y. Bar-Shalom and T. Fortmann, *Tracking and Data Association.* San Diego, California: Academic Press, 1988.

An Adaptive IMM Estimator for Aircraft Tracking

Emil Semerdjiev [1] **Ludmila Mihaylova** [1]

Bulgarian Academy of Sciences
Central Laboratory for Parallel Processing
'Acad. G. Bonchev' Str., Bl. 25-A, 1113 Sofia, Bulgaria
Phone: (359 2) 979 66 20; Fax: (359 2) 707 273
E-mail: lsm@bas.bg, signal@bas.bg

X. Rong Li *

University of New Orleans
Department of Electrical Engineering
New Orleans, LA 70148
Phone: 504-280-7416, Fax: 504-280-3950
E-mail: xli@uno.edu

Abstract–*An adaptive Interacting Multiple-Model (IMM) estimator using a small number of models is proposed for maneuvering aircraft tracking. It estimates the difference between the true target control parameter and the value currently used in the IMM models to improve the estimator's performance. The algorithm performance is compared with the performance of a standard IMM estimator for some maneuver scenarios via Monte Carlo simulations.*

Key Words: target tracking, IMM, adaptive estimation

1. Introduction

In recent years the design of reliable and effective multiple-model (MM) algorithms for maneuvering target tracking is a subject of extensive research (see e.g., [1-7]). These algorithms are used to overcome problems caused by structural and parametric uncertainties. The Interacting Multiple-Model (IMM) algorithm is the most commonly used MM estimation algorithm among them. The lack of knowledge about the target's control parameters is overcome in it by introducing a set of fixed control parameters. This set is expected to cover the range of possible parameter changes. A set of models represents the system behavior in each fixed control value. Kalman filters based on these models are running in parallel and their estimates are finally fused [1-3, 6] to compute the overall estimate. When the range of the expected control parameter is wide, however, IMM needs a large number of models to provide consistent estimation.

One promising solution to this problem is to use variable-structure estimation algorithms [5-7]. An alternative, nontrivial solution is proposed in this paper. It requires a minimal number of models (one for rectilinear motion, one for right turn and one for left turn) to cover the range of all possible target maneuvers. The proposed adaptive IMM algorithm estimates the difference between the control parameter assumed in the current model and its real value. The method has been applied at first for marine targets tracking in [10], where the range of the control parameter is very narrow. To cover the very wide respective range for air targets an additional adaptation mechanism is applied. It is concerned with the IMM transition probabilities and the fudge factor and the noise covariance matrices of the maneuvering models. The algorithm's performance is evaluated by Monte Carlo simulations and the effectiveness is illustrated by a comparison with 3- and 5-model standard IMM algorithm versions.

2. Aircraft Models

The target motion is described in the horizontal plane xOy by the commonly used model [8]:

$$\dot{X} = V \sin\varphi,$$
$$\dot{Y} = V \cos\varphi,$$
$$\dot{v} = g n_T,$$
$$\dot{\varphi} = -g n_N^* / v,$$

where $n_N^* = n_N \sin\gamma$, $\gamma = arc\cos(1/n_N)$; (X, Y) are aircraft mass center coordinates, V and φ are aircraft velocity and heading, n_N and n_T are nor-

1 Partially supported by the Bulgarian National Science Fund - Contract No. I-808/98.
* Supported by ONR via Grant N00014-97-1-0570, NSF via Grant ECS-9734285, and LEQSF via Grant (1996-99)-RD-A-32.

mal and tangential g-load factors (NLF and TLF), γ is roll angle; $g = 9.81$ is the load factor.

The respective discrete-time model has the form

$$X_k = X_{k-1} + Tv_{k-1}\sin\varphi_{k-1},$$
$$Y_k = Y_{k-1} + Tv_{k-1}\cos\varphi_{k-1},$$
$$\varphi_k = \varphi_{k-1} + Tgn^*_{N,k-1}/v_{k-1},$$
$$v_k = v_{k-1} + Tgn_{T,k},$$

where $n^*_{N,k} = n_{N,k}\sin\gamma_k, \gamma_k = arc\cos(1/n_{N,k})$; k is the current discrete time; T is the radar sampling interval. The state vector $x_k = \begin{pmatrix} X_k & Y_k & \varphi_k & v_k \end{pmatrix}'$ should be estimated in the presence of unknown control parameters n_N and n_T based on radar measurements y_k, modeled as:

$$y_k = Hx_{i,k} + w_k,$$

where H is measurement matrix and w_k is white Gaussian noise with a covariance matrix R:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},\ R = \begin{bmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_Y^2 \end{bmatrix}.$$

3. Extended Model and Adaptive IMM

The proposed IMM algorithm uses three models to cover the possible target motions in the horizontal plane: rectilinear motion ($i = 1$), right turn ($i = 2$) and left turn ($i = 3$):

$$X_{i,k} = X_{i,k-1} + Tv_{i,k-1}\sin\varphi_{i,k-1}, \quad (1)$$
$$Y_{i,k} = Y_{i,k-1} + Tv_{i,k-1}\cos\varphi_{i,k-1}, \quad (2)$$
$$\varphi_{i,k} = \varphi_{i,k-1} + Tg\left(n^*_{N_i} + \Delta n^*_{N_i,k-1}\right)/v_{i,k-1}, \quad (3)$$
$$v_{i,k} = v_{1,k-1} + Tgn_{T,k}, \quad (4)$$
$$\Delta n^*_{N_i,k} = \Delta n^*_{N_i,k-1}. \quad (5)$$

The i-th difference $\Delta n^*_{N_i,k-1}$ is a measure of the mismatch between the NLF currently used by i-th model and its true value. The extended state vector has the form $x_{i,k} = \begin{pmatrix} X_{i,k} & Y_{i,k} & \varphi_{i,k} & v_{i,k} & \Delta n^*_{N_i,k} \end{pmatrix}'$. It is also presumed for all IMM filters that $n_{T,k} \equiv 0$.

The EKF for the i-th model have the recursion:

$$\hat{x}_{i,k/k} = \hat{x}_{i,k/k-1} + K_{i,k}\gamma_{i,k}, \quad (6)$$
$$\hat{x}_{i,k/k-1} = f\left(\hat{x}_{i,k-1/k-1}, n^*_{N_i,k-1}\right), \quad (7)$$
$$\gamma_{i,k} = y_k - H_i\hat{x}_{i,k/k-1}, \quad (8)$$
$$P_{i,k/k-1} = \phi_i f_i^x P_{i,k-1/k-1}\left(f_i^x\right)' + Q_{i,k}, \quad (9)$$
$$S_{i,k} = H_i P_{i,k/k-1} H_i' + R, \quad (10)$$
$$K_{i,k} = P_{i,k/k-1} H_i' S_i^{-1}, \quad (11)$$
$$P_{i,k/k} = P_{i,k/k-1} - K_{i,k} S_i K_{i,k}', \quad (12)$$

where $\hat{x}_{i,k/k}$ and $\hat{x}_{i,k/k-1}$ are the filtered and the predicted estimates of x_k; $f_i^{\hat{x}} = \partial f_i / \partial \hat{x}_i$ given by:

$$f_i^{\hat{x}} = \begin{bmatrix} 1 & 0 & T\hat{v}_{i,k/k-1}\cos\hat{\varphi}_{i,k/k-1} & T\sin\hat{\varphi}_{i,k/k-1} & 0 \\ 0 & 1 & -T\hat{v}_{i,k/k-1}\sin\hat{\varphi}_{i,k/k-1} & T\cos\hat{\varphi}_{i,k/k-1} & 0 \\ 0 & 0 & 1 & -\frac{Tg\left(n^*_{N_i} + \Delta\,\hat{n}^*_{N_i,k/k-1}\right)}{\hat{v}^2_{i,k/k-1}} & \frac{Tg}{\hat{v}_{i,k/k-1}} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

γ_i and $S_{i,k}$ are the filter residual process and its covariance matrix, $P_{i,k/k}$ and $Q_{i,k}$ is the error and the system noise covariance matrices, $K_{i,k}$ is the filter gain matrix, $\phi_i > 1$ is a fudge factor (FF).

The estimate $\Delta\hat{n}^*_{N_i,k/k}$ has no significant physical meaning directly but it contains useful information about the maneuver's starting/final times and intensity. It allows us to develop an adaptive mechanism for estimation consistency improvement. They are arranged below according to their impact.

a) The FF is adaptively changed according to the final estimate $\hat{x}(5)$ in the subsequent times k, k-1:

$$\phi_{i,k} = \begin{cases} \phi_0 = 1.06, & \text{for } i = 1; \\ 1 + \dfrac{\left|\hat{x}_k(5) + \hat{x}_{i,k-1}(5)\right|}{4n^*_{N,\max}} \le \phi_0, & \text{otherwise} \end{cases} \quad (13)$$

This adaptive FF is introduced to improve the common filter consistency.

b) The fifth diagonal element of the process noise covariance matrix $Q_{i,k}$, $i = 1,2,3$ of the EKF is adaptively changed to provide faster response to the maneuvers:

$$Q_{i,k} = diag\left(\sigma_X^2 \quad \sigma_Y^2 \quad \sigma_\varphi^2 \quad \sigma_v^2 \quad \left|\hat{x}_{k/k-1}(5) - \hat{x}_{i,k/k-1}(5)\right|\right) \quad (14)$$

c) The transition probabilities are computed as follows:

$$Pr_{11}(k) = Pr_{11}(0)e^{-\left|\hat{x}_k(5) - \hat{x}_{1,k}(5)\right|},$$
$$Pr_{jj}(k) = Pr_{jj}(0)e^{-\left|\hat{x}_k(5) - \hat{x}_{j,k}(5)\right|/n^*_{N,max}}, j=2,3; \quad (15)$$
$$Pr_{ij}(k) = Pr_{il}(k) = \left(1 - Pr_{ii}(k)\right)/2; i \ne j \ne l,$$
$$i,j,l \in [1,3].$$

where $n^*_{N,max}$ (it is set = 7) is the maximal expected value of the NLF $n^*_{N,k}$, and the standard IMM transition probability matrix has the form:

$$Pr_{ii}(k) = Pr_{ii}(0) = const, \qquad (16)$$
$$Pr_{ij}(k) = Pr_{il}(k) = (1 - Pr_{ii}(k))/2; i \neq j \neq l,$$
$$i, j, l \in [1, i_{max}].$$

This adaptive transition probability matrix provides faster system mode transition.

4. Performance Evaluation

The performance of the adaptive IMM filter (denoted below as IMM3a) is evaluated by Monte Carlo simulation for 200 independent runs. Its performance is compared with that of a 3- and a 5-model standard IMM (denoted as IMM3 and IMM5). IMM3 and IMM5 use: the model (1)-(4) and the EKF equations (6)-(12), where it is set $\Delta\hat{n}^{*}_{N_i,k/k-1} \equiv 0$ and in the matrix $f_i^{\hat{x}}$ the last row and column are excluded.

It is preset:

$$\sigma_X = \sigma_Y = 100m, \sigma_\psi = 3^o, \sigma_v = 10m/s,$$
$$\sigma_{\Delta n_N} = 0.2, T = 1s. \qquad (17)$$
$$n^{*}_{N_i} = n_{N_i}\ sin(arccos(1/n_{N_i}))$$

IMM3 and IMM3a use $n_N = (1 \quad 4 \quad -4)'$, whereas IMM5 use $n_N = (1 \quad 3 \quad -3 \quad 6 \quad -6)'$.

Example 1.

To demonstrate the significant role of the proposed adaptation, the IMM3 and IMM5 are compared with a simplified version of IMM3a (denoted as IMM3as), where the adaptive mechanisms (13)-(15) and (16) are not included; that is, all algorithms use constant transition probabilities (16), a common FF $\phi = 1.06$ and covariance matrices $Q_{i,k}$ with diagonal elements given in (17). All the IMM algorithms are running with the following initial conditions: $X_0 = Y_0 = 0\,m$, V_0=350 m/s, $\varphi = 0°$. The initial error covariance matrix $P(0)$, the initial mode probability vector $\mu(0)$ and the transition probability matrices Pr are:

$P(0) = diag\{10000, 10000, 9, 4000, 0.5\}$ for IMM3as,

$P(0) = diag\{10000, 10000, 9, 4000\}$ for IMM3,

where $diag\{.\}$ denotes a diagonal matrix.

For IMM3 and IMM3as it is preset:

$$Pr = \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}; \qquad \mu(0) = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix},$$

+For the IMM5 it is preset:

$$Pr = \begin{bmatrix} 0.6 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.05 & 0.8 & 0.05 & 0.05 & 0.05 \\ 0.05 & 0.05 & 0.8 & 0.05 & 0.05 \\ 0.05 & 0.05 & 0.05 & 0.8 & 0.05 \\ 0.05 & 0.05 & 0.05 & 0.05 & 0.8 \end{bmatrix}, \qquad \mu(0) = \begin{pmatrix} 1/5 \\ 1/5 \\ 1/5 \\ 1/5 \\ 1/5 \end{pmatrix}$$

A target maneuver with n_N=7 is the worst case for IMM3a and IMM3as.The true target trajectory and the NLF change are given in Figs. 1, 2.

The Normalized Estimation Error Squared (NEES) [2] is the most informative and integral measure of performance. So, the respective NEES plots for all algorithms are computed and presented in Fig. 3 ('1' - IMM3, '2' - IMM3as, '3' - IMM5). Here and below the NEES is computed for the first four components of the state vector.

Obviously the standard 3-model IMM does not provide consistent estimates during the maneuver at all, while the consistency of IMM3as is better than the IMM5.

Example 2.

The designed IMM3a (denoted by '1') is compared with IMM5 (denoted by '2') for three types of maneuvers: a fast maneuver with n_N=7, a moderate maneuver with n_N=3 and a weak maneuver with n_N=1.2. A noise covariance matrix is introduced in the EKF equation (9), with elements $j = \overline{1,4}$, given in (17). The fifth element of Q in IMMa is adaptively changed according to (14). Obviously, the presence of separate models with n_N=3 and n_N=6 gives advantages to the IMM5 over the IMM3 and IMM3a in the first two test scenarios.

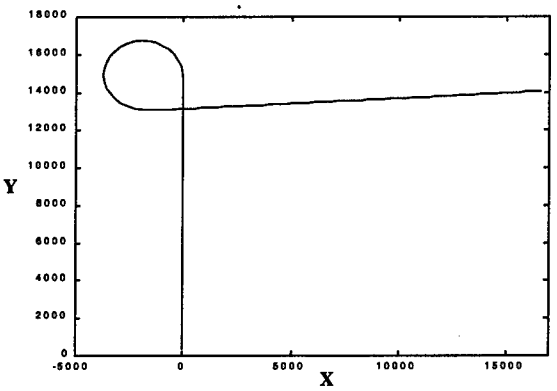


Fig. 1 True aircraft trajectory

The respective results for the fast maneuver are shown in Figs. 4-14. The NEES is given in Fig. 4. The mean errors (ME) and the root-mean square

errors (RMSE) of the state vector are shown in Figs. 5-8 and Figs.9-11, respectively. The average model probabilities are presented in Figs. 12-13 and the average FF behavior is given in Fig. 14. These results show that IMM3a have better consistency, accuracy and faster response to abrupt maneuvers.

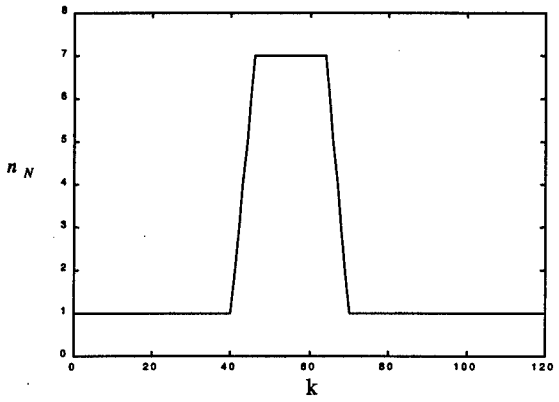


Fig. 2 True aircraft normal load factor

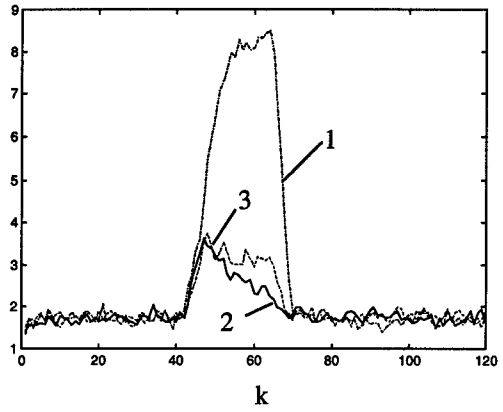


Fig. 3 Normalized Estimation Error Squared

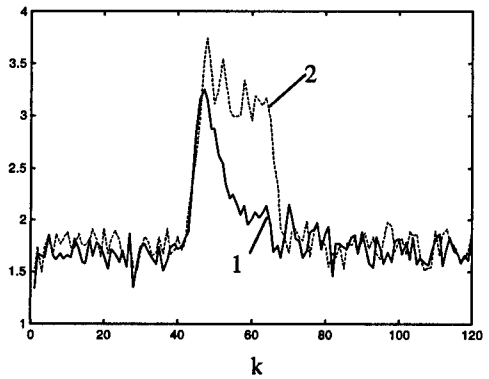


Fig. 4 Normalized Estimation Error Squared

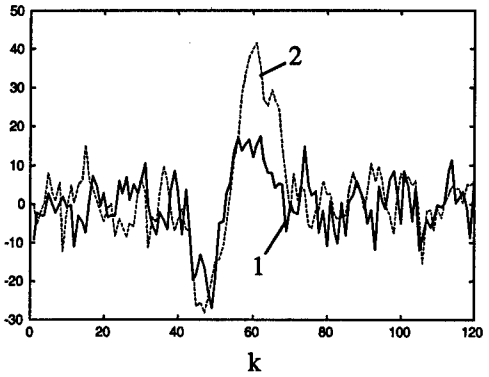


Fig. 5 X Position ME

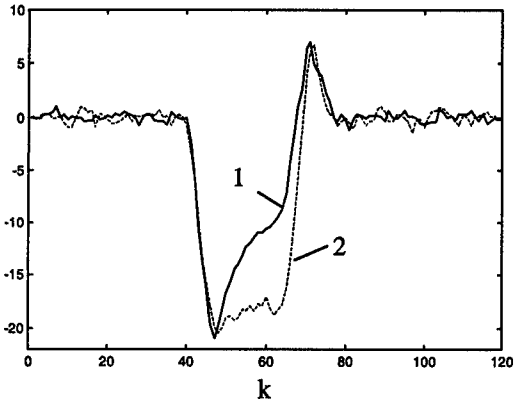


Fig. 6 Heading ME

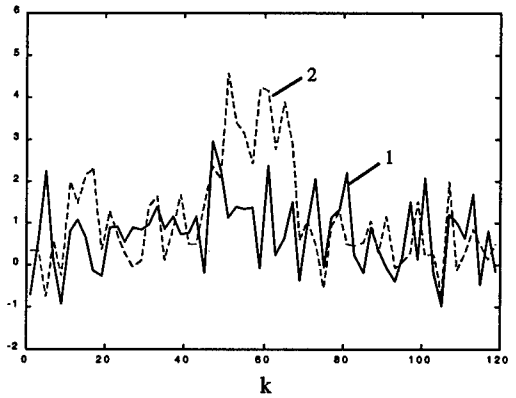


Fig. 7 Velocity ME

The true trajectory and the true NLF change for the moderate maneuver are represented in Figs. 15-16. The NEES and average FF behavior are given in Figs. 17 and 18, respectively.

The same inferences can be drawn for the next two scenarios. The true trajectory and the true NLF change for the weak maneuver are presented in Figs. 19-20. The NEES and average FF behavior are given in Figs. 21 and 22, respectively.

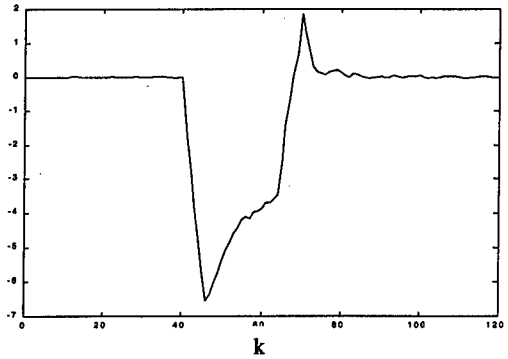


Fig. 8 $\Delta\hat{n}^{*}_{N_i,k/k}$ ME

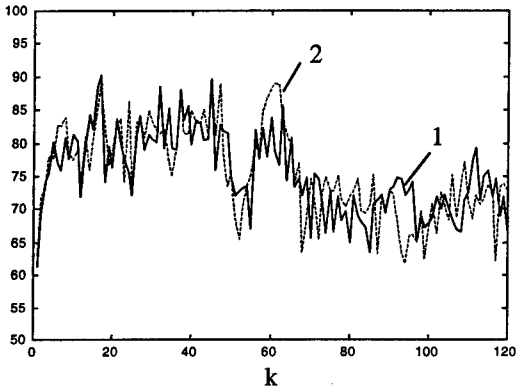


Fig. 9 X Position RMSE

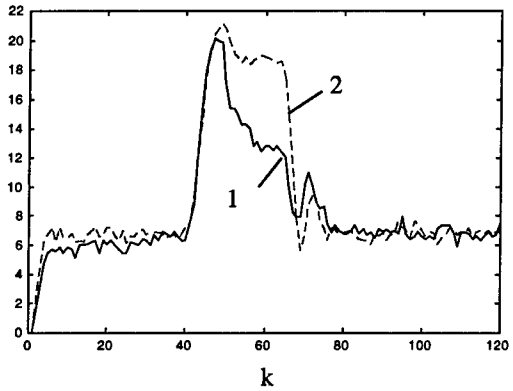


Fig. 10 Heading RMSE

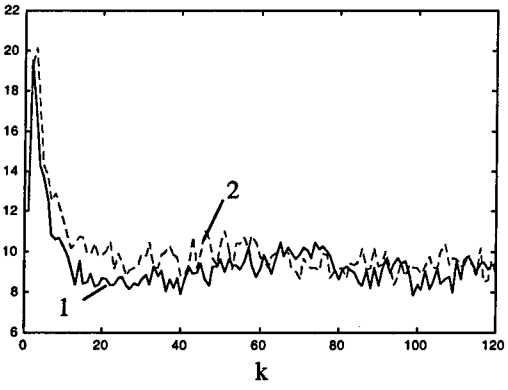


Fig. 11 Velocity RMSE

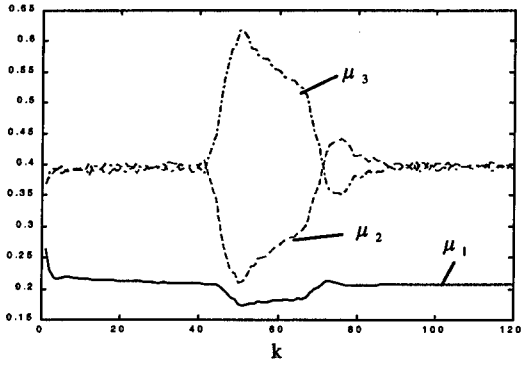


Fig. 12 Mean of Model Probabilities for IMM3

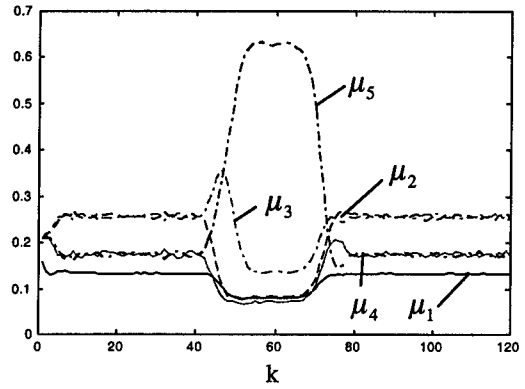


Fig. 13 Mean of Model Probabilities for IMM5

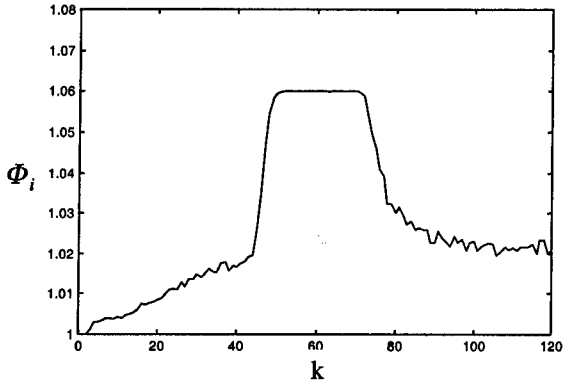


Fig. 14 Average fudge factor

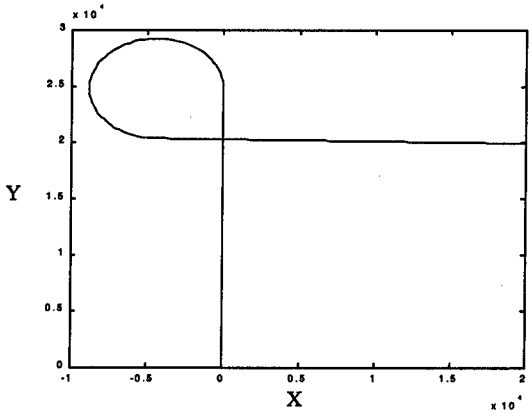


Fig. 15 True aircraft trajectory for n_N =3

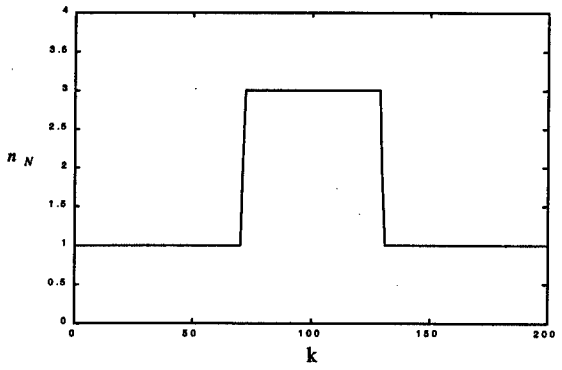


Fig. 16 True aircraft normal load factor

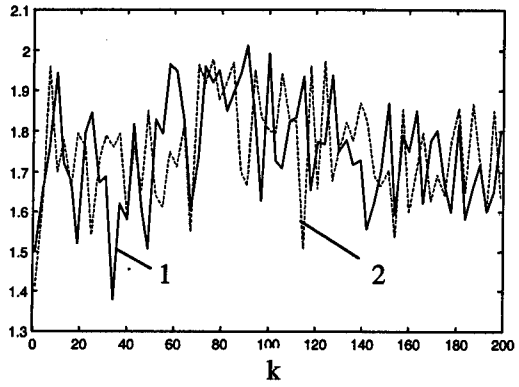


Fig. 17 Normalized Estimation Error Squared

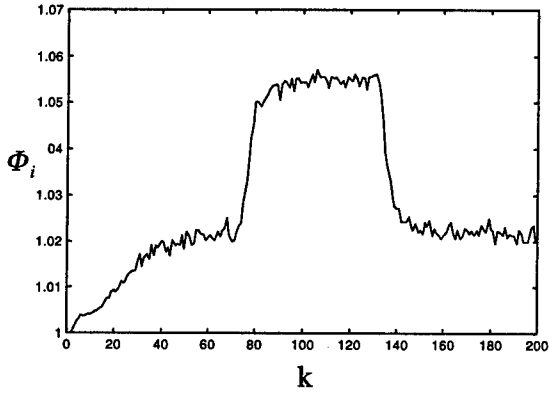


Fig. 18 Average fudge factor

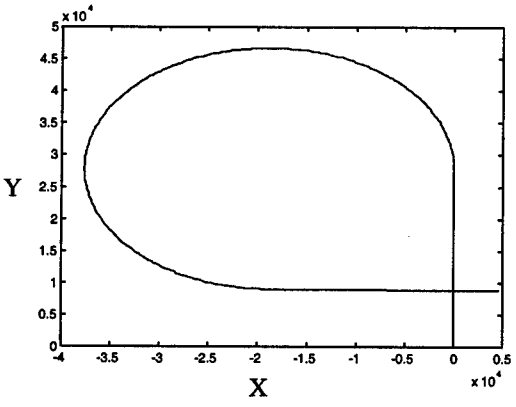


Fig. 19 True aircraft trajectory for $n_N = 1.2$

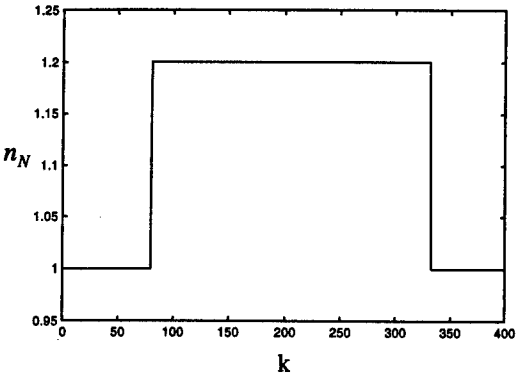


Fig. 20 True aircraft normal load factor

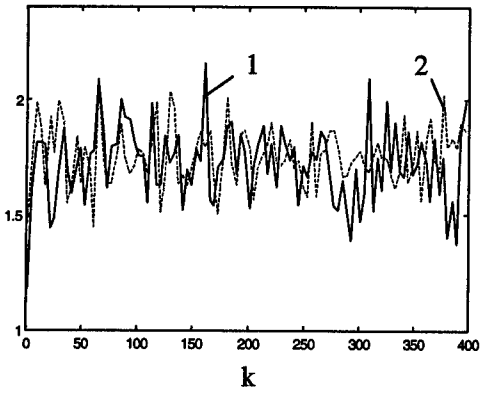


Fig. 21 Normalized Estimation Error Squared

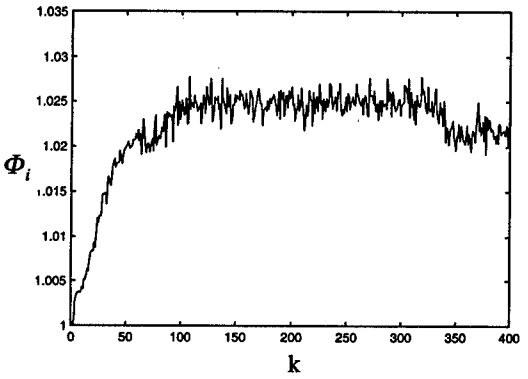


Fig. 22 Average fudge factor

6. Conclusions

An adaptive IMM algorithm using a small number of models and covering a wide range of possible aircraft maneuvers is proposed. It estimates the difference between the real control parameter and its value used in the current model in real time. This makes it possible to introduce an additional adaptation mechanism to cover a wide range of possible air target maneuvers. This mechanism tunes the IMM transition probabilities, the EKF's fudge factor and

the EKF's covariance. The algorithm's efficiency is demonstrated for the worst cases maneuvers.

References

[1] Bar-Shalom, Y., Ed., *Multitarget-Multisensor Tracking: Applications and Advances*, vol. 2, Artech House, 1992.

[2] Bar-Shalom, Y., and X. R. Li. *Estimation and Tracking Principles, Techniques and Software*, Artech House, 1993.

[3] Bar-Shalom, Y., and X. R. Li. *Multitarget-Multisensor Tracking: Principles and Techniques*, YBS Publishing, 1995.

[4] Jilkov, V., P. Angelova and Tz. Semerdjiev. Design and Comparison of Mode-Set Adaptive IMM for Maneuvering Target Tracking, *IEEE Trans. Aerosp. and Electr. Systems*, 35(1): 343-350, Jan. 1999.

[5] Jilkov, V. P., L.Mihaylova, X. R. Li, An Alternative IMM Solution to Benchmark Radar Tracking Problem, 1998 Intern. Conf. on Information Fusion (Fusion'98), vol.2, 924-929, Las Vegas, Nevada, USA, July 6-9, 1998.

[6] Li, X. R. *Hybrid Estimation Techniques.* In *Control and Dynamic Systems.* (Ed., C. T. Leondes), vol. 76, 213-287, Academic Press, 1996.

[7] Li, X. R., and Y. Bar-Shalom. Multiple-Model Estimation with Variable Structure, *IEEE Trans. Automatic Control,* 41(4): 478-493, 1996.

[8] Ostoslavsky, I., Strajeva, I., *Flight Dynamics.* Flying Aircraft Trajectories. Mashinostroenie, Moscow, 2nd edition, 1969 (in Russian).

[9] Semerdjiev, E., L. Mihaylova, Tz. Semerdjiev. Manoeuvring Ship Model Identification and IMM Tracking Algorithm Design, in Proc. 1998 Intern. Conf. on Information Fusion (Fusion'98), Vol. 2, 968-973, Las Vegas, Nevada, USA, July 6-9, 1998.

[10] Semerdjiev, E., L. Mihaylova, Adaptive IMM Algorithm for Manoeuvring Ship Tracking, in Proc. 1998 Intern. Conf. on Information Fusion (Fusion'98), Vol. 2, 974-979, Las Vegas, Nevada, USA, July 6-9, 1998.

Tracking Maneuvering Targets Using Geographically Separated Radars

Hiroshi Kameda Shingo Tsujimichi Yoshio Kosuge

Mitsubishi Electric Corporation, Information R & D Center
5-1-1 Ofuna, Kamakura, Kanagawa 247-8501, JAPAN,

Abstract *This paper proposes a maneuvering target tracking algorithm using geographically separated radars. This filtering algorithm is discussed in terms of tracking performance, tracking success rate and tracking accuracies as compared with other conventional methodology. Through several simulations, validity of this algorithm has been confirmed.*

Keywords: Multi-Target Multi-Sensor Tracking, Geographically Separated Radars, Interacting Multiple Models

1 Introduction

Measurement-to-track or hit-to-track data association is an essential technique in track maintenace algorithms. The joint probabilistic data association(JPDA)[1][2] filter has been reported to be suitable for the above technique. The JPDA filter updates a track with a weighted sum of feasible hits at every scan. The weights are calculated by finding all of the possible hit-to-track combination hypotheses, along with all possible hypotheses of hit-to-track associations which include track misses. However, as the measurement conditions, such as crossing angles of crossing targets and range between radar and targets, become severe, the JPDA cannot give full performance due to radar resolutions and measurement errors.

The purpose of this paper is to enhance the conventional JPDA filter. In order to attain this objective, first, geographically separated radars are applied to reduction of the influences of radar resolution and measurement errors. Measurements from various points are received for tracking. Measurements arrive as raw data; that is, individual position measurements such as slant range, elevation and azimuth for each radar accompanied by a time stamp and an estimated standard deviation. A major problem encountered in using these radars is that tracking algorithm accepts measurements from many different locations and coordinate conversion plays a very important role in tracking algorithm[3]. The JPDA updates are accomplished for one set of measured parameters at a time with an appropriate measurement matrix computed for each measurement point with the consideration of coordinate conversion.

Next, the interacting multiple model (IMM) algorithm is applied to JPDA for tracking multi-targets maneuvering in three dimensions. The applicability of the original IMM algorithm was investigated and confirmed through simulations[4]. In addition, in the presence of clutter, the IMM has to be complemented in order to take into account the uncertainty of measurements origin. It was shown by Houles and Bar-Shalom that the probabilistic data association(PDA) logic is an efficient solution for this aspect[5]. In tracking long range targets, however, the IMM cannot give full performance for maneuvering targets due to the measurement errors. In order to overcome this problem, we apply the IMM algorithm to JPDA where distributed radars are fused.

This paper is organized as follows. Coordinate systems used in this paper are shown in the next section. Section 4 presents our methodology using distributed radars. Section 5 discusses numerical performance results of the described methodologies using simulation data. The conclusion is given in the final section.

2 Coordinate Systems

Fig.1(a) shows the geocentric-equatorial coordinate system, which is centered at the center of the earth. In this system, z is directed along the axis of the Earth rotation and x and y lie in the equatorial plane, with x pointing Greenwich. Radar centered N-E-U coordinate system is shown in Fig. 1(b). In this system, z is directed along the axis of the local vertical and x and y lie in the local horizontal plane with x pointing east and y pointing north. We

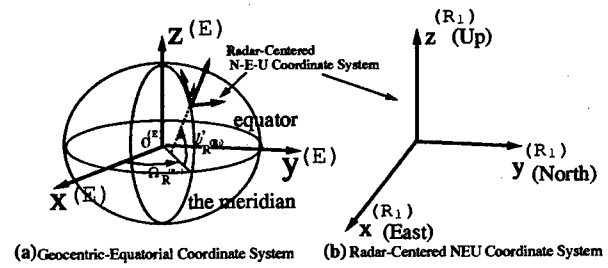


Figure 1: Definition of Coordinate Systems.

define position vector in geocentric-equatorial coordinate system as:

$$\boldsymbol{x}^{(E)} = \left[x^{(E)}, y^{(E)}, z^{(E)}\right]^T. \quad (1)$$

On the other hands, position vector in radar centered N-E-U coordinate system of radar l is defined in eqn.(2).

$$\boldsymbol{x}^{(R_l)} = \left[x^{(R_l)}, y^{(R_l)}, z^{(R_l)}\right]^T \quad (2)$$

The transformation from $\boldsymbol{x}^{(E)}$ to $\boldsymbol{x}^{(R_l)}$ is given by the following equation:

$$\begin{bmatrix} x^{(R_l)} \\ y^{(R_l)} \\ z^{(R_l)} \end{bmatrix} = \boldsymbol{T}_{ER_l} \begin{bmatrix} x^{(E)} \\ y^{(E)} \\ z^{(E)} \end{bmatrix} - \boldsymbol{r}_{ER_l}, \quad (3)$$

where each element of $\boldsymbol{T}_{ER_l}$ and $\boldsymbol{r}_{ER_l}$ consists of geodetic longitude and latitude of radar l.

3 JPDA using geographically separated radars

3.1 Feature and subject

In tracking long range targets using single radar, measurement errors in cross-range direction are much larger compared with range direction (see Fig.2). Besides, it is seldom possible to resolve two targets due to an insufficient angle resolution.

Then, if measurement data at the same time from radar 1 and radar 2 are supplied to the tracking point and tracking process is done, the problem of the measurement error and angle resolution can be reduced as shown in this figure. As a result, the improvement of the tracking performance can be expected. In using

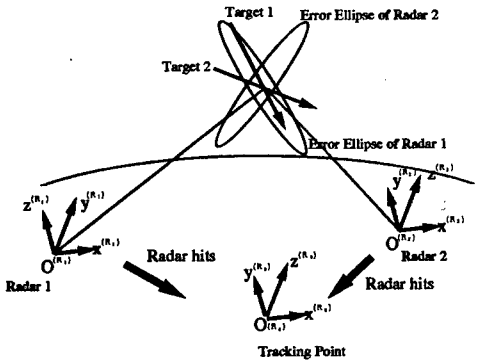


Figure 2: Target Tracking Using Geographically Separated Radars.

measurement data from geographically separated radars, coordinate axes of tracking point and each radar do not correspond due to the roundness of the earth. In order to construct tracking algorithm, this disagreement must be considered in constructing of the measurement equation of tracking filter.

3.2 Data association

Denote a scan of hits of radar $l(l = 1, 2, \cdots)$ as:

$$\boldsymbol{Z}_{k,l} = \left[\boldsymbol{z}^l_{k,1}, \boldsymbol{z}^l_{k,2}, \cdots \boldsymbol{z}^l_{k,m_{k,l}}\right] \quad (4)$$

where $\boldsymbol{z}^l_{k,i}(i = 1, 2, \cdots, m_{k,l})$ is an individual hit of radar l on scan k and $m_{k,l}$ is the number of hits on scan k.

Denote the history of hits up to scan k as $\boldsymbol{Z}^{k,l}$:

$$\boldsymbol{Z}^{k,l} = [\boldsymbol{Z}_{1,l}, \boldsymbol{Z}_{2,l}, \cdots \boldsymbol{Z}_{k,l}]. \quad (5)$$

Furthermore, denote the history of the number of hits of radar l up to scan k as:

$$\boldsymbol{M}_{k,l} = [m_{1,l}, m_{2,l}, \cdots, m_{k,l}]. \quad (6)$$

Denote the set of data association hypotheses for radar l at scan k as:

$$\boldsymbol{X}^{k,l} = \left[X^{k,1,l}, X^{k,2,l}, \cdots, X^{k,\alpha_k,l}\right] \quad (7)$$

where α_k is the total number of hypotheses. The joint hypotheses $X^{k,i,l}(i = 1, 2, \cdots \alpha_k)$ are defined as:

$$X^{k,i,l} = \cap_{j=1}^{m_{k,l}} X^{k,i,l}_{jn_j} \quad (8)$$

where $X^{k,i,l}_{jn_j}$ is the event that hit j originated from target $n_j(n_j = 0, 1, \cdots, L)$. Index $n_j = 0$ stands for clutter. Each joint event can be represented by a data association matrix shown as eqn.(9).

$$\Omega(\boldsymbol{X}^{k,l}) = \left[W^{k,l}_{jn}\right] \quad (9)$$

where $W^{k,l}_{j0} = 1$ means that hit j could originate from clutter, $W^{k,l}_{jn} = 1$ if hit j is inside the validation gate of target n and $W^{k,l}_{jn} = 0$ if hit j is outside the validation gate of target n for $j = 1, 2, \cdots, m_{k,l}$, and $n = 0, 1, 2, \cdots, L$. Based on the data association matrix, data association hypotheses $X^{k,i,l}$ are generated in matrix form as the following equation:

$$\Omega(\boldsymbol{X}^{k,i,l}) = \left[W^{k,i,l}_{jn}\right] \quad (10)$$

where $W^{k,i,l}_{jn} = 1$ means that data association hypothesis $X^{k,i,l}_{jn}$ is possible and $W^{k,i,l}_{jn} = 0$ if $X^{k,i,l}_{jn}$ is impossible.

According to the above data association hypotheses $X^{k,i,l}$, the following indicators can be defined:

$$\delta_n(X^{k,i,l}) = \sum_{j=1}^{m_{k,l}} W^{k,i,l}_{jn}(n = 1, 2, \cdots L) \quad (11)$$

$$\delta_n(X^{k,i,l}) = 1 \text{ (if target n is detected)} \quad (12)$$

$$\delta_n(X^{k,i,l}) = 0 \text{ (if target n is not detected)} \quad (13)$$

$$\tau_j(X^{k,i,l}) = \sum_{n=1}^{L} W^{k,i,l}_{jn} \quad (14)$$

$$\tau_j(X^{k,i,l}) = 1$$
(if hit j is associated with target) (15)

$$\tau_j(X^{k,i,l}) = 0$$
(if hit j is not associated with target) (16)

$$\Psi(X^{k,i,l}) = \sum_{j=1}^{m_{k,l}} \left[1 - \tau_j(X^{k,i,l})\right] \quad (17)$$

where $\Psi(X^{k,i,l})$ is the total number of clutter in data association hypothesis $X^{k,i,l}$.

4 Multiple Model JPDA

In this section, JPDA using distributed radars for maneuvering targets is shown.

4.1 Modeling

The kinematic model of target n $(n = 1, 2, \cdots, L)$ in Cartesian coordinates is described as follows:

$$\boldsymbol{x}^n_{k,a^n} = \Phi_{k-1,a^n}\boldsymbol{x}^n_{k-1,a^n} + \boldsymbol{w}^n_{k-1,a^n} \quad (18)$$

where

$\boldsymbol{w}^n_{k,a^n}$:acceleration noise of the target n,
and $E\left[\boldsymbol{w}^n_{k,a^n}\right] = \mathbf{0}$, $E\left[\boldsymbol{w}^n_{k,a^n}\boldsymbol{w}^{n^T}_{k,a^n}\right] = \boldsymbol{Q}^n_{k,a^n}$.

where suffix k means scan k, and a denotes the targets kinematic model number. With regard to model number a, a six-dimensional vector consisting of position and velocity for each coordinate is considered as kinematic model 1

and a nine-dimensional vector adding accelerations to model 1 is applied as kinematic model 2.

Transition matrix $\Phi_{k,a^n}(a=1,2)$ is

$$\Phi_{k-1,1^n}=\begin{bmatrix} I_{3\times 3} & \Delta t I_{3\times 3} \\ O_{3\times 3} & I_{3\times 3} \end{bmatrix}, \tag{19}$$

$$\Phi_{k-1,2^n}=\begin{bmatrix} I_{3\times 3} & \Delta t I_{3\times 3} & \frac{(\Delta t)^2}{2} I_{3\times 3} \\ O_{3\times 3} & I_{3\times 3} & \Delta t I_{3\times 3} \\ O_{3\times 3} & O_{3\times 3} & I_{3\times 3} \end{bmatrix} \tag{20}$$

where $I_{3\times 3}$ is unity matrix and Δt is sampling interval.

$$Q^n_{k,1^n}=q_1^n\begin{bmatrix} \frac{(\Delta t)^4}{4} I_{3\times 3} & \frac{(\Delta t)^3}{2} I_{3\times 3} \\ \frac{(\Delta t)^3}{2} I_{3\times 3} & (\Delta t)^2 I_{3\times 3} \end{bmatrix}. \tag{21}$$

$$Q^n_{k,2^n}=q_2^n\begin{bmatrix} \frac{(\Delta t)^4}{4} I_{3\times 3} & \frac{(\Delta t)^3}{2} I_{3\times 3} & \frac{(\Delta t)^2}{2} I_{3\times 3} \\ \frac{(\Delta t)^3}{2} I_{3\times 3} & (\Delta t)^2 I_{3\times 3} & \Delta t I_{3\times 3} \\ \frac{(\Delta t)^2}{2} I_{3\times 3} & \Delta t I_{3\times 3} & I_{3\times 3} \end{bmatrix} \tag{22}$$

where q_1 and q_2 are the variances of the acceleration noise of kinematic models 1 and 2, respectively.

Assuming that x_k^n is the state vector of target n after the integration of the above kinematic models, the relationshipe between x_k^n and x_{k,a^n}^n is described by the following equation :

$$x_{k,a^n}^n = D_a x_k^n \tag{23}$$

where D_a is a constant matrix for adjusting of dimensions of x_{k,a^n}^n and x_k^n.

Denote the dimension p of x_k^n as :

$$p = max\{p_1, p_2\} \tag{24}$$

where p_1 and p_2 are dimensions of kinematic models 1 and 2, respectively. Each kinematic model is modeled as a stationary Markov process and the transition probabilities are denoted by $P_{a^n b^n}$,

$$P_{a^n b^n} = Pr\left[\Psi_{k,a^n} | \Psi_{k-1,b^n}\right] \tag{25}$$

where Ψ_{k,a^n} is the event that kinematic model a is true. Denote the set of events Ψ_{k,a^n} as

$$\Psi^{k,a} = \left[\Psi_{k,a^1}, \Psi_{k,a^2}, \cdots, \Psi_{k,a^L}\right]. \tag{26}$$

Denote the transition probabilities of the set of events by P^{ab},

$$P^{ab} = Pr\left[\Psi^{k,a} | \Psi^{k-1,b}\right] \tag{27}$$

and assuming that $P_{a^n b^n}$ is independent of target n,

$$P^{ab} = \prod_{n=1}^{L} P_{a^n b^n} \tag{28}$$

The observation vector from radar l is a three-dimensional vector as follows:

$$z_k^l = \left[x_{ko}^{(R_l)}, y_{ko}^{(R_l)}, z_{ko}^{(R_l)}\right]^T, \tag{29}$$

where o means Observation and (R_l) denotes the observation vector in N-E-U coordinate system of radar l.

The observation equation can be written as:

$$z_k^l = h(x_{k,a^n}^n) + v_{k,l} \tag{30}$$

where

z_k^l:observed vector on scan k of radar l,

$v_{k,l}$:observation noise on scan k of radar l,

and $E[v_{k,l}] = \mathbf{o}$, $E\left[v_{k,l} v_{k,l}^T\right] = R_{k,l}$.

$$h(x_{k,a^n}^n) = T_{ER_l} T_{ER_0}^T H_{k,a_0} x_{k,a^n}^n + T_{ER_l}(r_{ER_0} - r_{ER_l}) \tag{31}$$

$H_{k,a,0}$ is measurement matrix at tracking point:

$$H_{k,1,0} = \begin{bmatrix} I_{3\times 3} & 0 I_{3\times 3} \end{bmatrix}. \tag{32}$$

$$H_{k,2,0} = \begin{bmatrix} I_{3\times 3} & 0 I_{3\times 6} \end{bmatrix}. \tag{33}$$

Measurement matrix for radar l at tracking point $H_{k,a}^l$ is obtained as follows:

$$H_{k,a,l} = T_{ER_l} T_{ER_0}^T H_{k,a,0}, \tag{34}$$

$$H_{k,1,l} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & 0 & 0 & 0 \\ h_{21} & h_{22} & h_{23} & 0 & 0 & 0 \\ h_{31} & h_{32} & h_{33} & 0 & 0 & 0 \end{bmatrix}, \tag{35}$$

$$H_{k,2,l} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & 0 & 0 & 0 & 0 & 0 & 0 \\ h_{21} & h_{22} & h_{23} & 0 & 0 & 0 & 0 & 0 & 0 \\ h_{31} & h_{32} & h_{33} & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{36}$$

where $h_{ij}(i = 1, \cdots, 3 : j = 1, \cdots, 3)$ is the funtion of geodetic longitudes and latitudes of radars and tracking point.

4.2 Prediction

Predicted state vector based on each kinematic model in N-E-U coordinate system on scan k, $\hat{\boldsymbol{x}}^n_{k,a^n}(-)$ is assumed to be state estimated value based on $\boldsymbol{M}^l_{k-1}$, $\boldsymbol{Z}^{k-1,l}$ and Ψ_{k,a^n} and the error covariance of $\hat{\boldsymbol{x}}^n_{k,a^n}(-)$ is defined as $\boldsymbol{P}^n_{k,a^n}(-)$.

$\hat{\boldsymbol{x}}^n_{k,a^n}(-)$ and $\boldsymbol{P}^n_{k,a^n}(-)$ are calculated as follows:

$$\begin{aligned}\hat{\boldsymbol{x}}^n_{k,a^n}(-) &= E\left[\boldsymbol{x}^n_k|\Psi_{k,a^n},\boldsymbol{Z}^{k-1,l},\boldsymbol{M}^l_{k-1}\right],\\ &= \sum_{b=1}^{2} Pr\left[\Psi_{k-1,b^n}|\Psi_{k,a^n},\boldsymbol{Z}^{k-1,l},\boldsymbol{M}^l_{k-1}\right]\boldsymbol{D}_b\\ &\times\left(\boldsymbol{\Phi}_{k-1,b^n}\hat{\boldsymbol{x}}^n_{k-1,b^n}(+)\right),\end{aligned} \tag{37}$$

$$\begin{aligned}\boldsymbol{P}^n_{k,a^n}(-) &= E\left[(\boldsymbol{x}^n_k-\hat{\boldsymbol{x}}^n_{k,a^n}(-))(\boldsymbol{x}^n_k-\hat{\boldsymbol{x}}^n_{k,a^n}(-))^T\right.\\ &\left.|\Psi_{k,a^n},\boldsymbol{Z}^{k-1,l},\boldsymbol{M}^l_{k-1}\right],\\ &= \sum_{b=1}^{2} Pr\left[\Psi_{k-1,b^n}|\Psi_{k,a^n},\boldsymbol{Z}^{k-1,l},\boldsymbol{M}^l_{k-1}\right]\boldsymbol{D}_b^T\\ &\times\left[\boldsymbol{\Phi}_{k-1,b^n}\left\{\boldsymbol{P}^n_{k-1,b^n}(+)\right.\right.\\ &+\left(\hat{\boldsymbol{x}}^n_{k-1,b^n}(+)-\sum_{b=1}^{2} Pr\left[\Psi_{k-1,b^n}|\Psi_{k,a^n},\boldsymbol{Z}^{k-1,l},\boldsymbol{M}^l_{k-1}\right]\right.\\ &\left.\times\hat{\boldsymbol{x}}^n_{k-1,b^n}(+)\right)\\ &\times\left(\hat{\boldsymbol{x}}^n_{k-1,b^n}(+)-\sum_{b=1}^{2} Pr\left[\Psi_{k-1,b^n}|\Psi_{k,a^n},\boldsymbol{Z}^{k-1,l},\boldsymbol{M}^l_{k-1}\right]\right.\\ &\left.\left.\times\hat{\boldsymbol{x}}^n_{k-1,b^n}(+)\right)^T\right\}\boldsymbol{\Phi}^T_{k-1,b^n}+\boldsymbol{Q}^n_{k-1,b^n}\Big]\boldsymbol{D}_b.\end{aligned} \tag{38}$$

where $E[\cdot]$ means average.

$$\begin{aligned}&Pr\left[\Psi_{k-1,b^n}|\Psi_{k,a^n},\boldsymbol{Z}^{k-1,l},\boldsymbol{M}^l_{k-1}\right]\\ &=\frac{P_{ab}\beta^{n,l}_{k-1,b^n}}{\sum_{b=1}^{2}P_{ab}\beta^{n,l}_{k-1,b^n}}\end{aligned} \tag{39}$$

Next, predicted state vector and its error covariance with each mixed estimate obtained by the above equations are calculated as follows:

$$\begin{aligned}\hat{\boldsymbol{x}}^n_k(-) &= E\left[\boldsymbol{x}^n_k|\boldsymbol{Z}^{k-1,l},\boldsymbol{M}^l_{k-1}\right],\\ &= \sum_{a=1}^{2}\sum_{b=1}^{2}\beta^{n,l}_{k-1,b^n}P_{a^nb^n}(\boldsymbol{D}_a^T\hat{\boldsymbol{x}}^n_{k,a^n}(-))\end{aligned} \tag{40}$$

where $P_{a^nb^n}$ is transition probability.

$$\begin{aligned}\boldsymbol{P}_k(-) &= E\left[(\boldsymbol{x}_k-\hat{\boldsymbol{x}}_k(-))(\boldsymbol{x}_k-\hat{\boldsymbol{x}}_k(-))^T|\boldsymbol{Z}^{k-1,l},\boldsymbol{M}^l_{k-1}\right],\\ &= \sum_{a=1}^{2}\sum_{b=1}^{2}\beta^{n,l}_{k-1,b^n}P_{a^nb^n}\boldsymbol{D}_a\left[\boldsymbol{P}^n_{k,a^n}(-)\right.\\ &+\left.(\hat{\boldsymbol{x}}^n_{k,a^n}(-)-\boldsymbol{D}_a\hat{\boldsymbol{x}}^n_k(-))\,(\hat{\boldsymbol{x}}^n_{k,a^n}(-)-\boldsymbol{D}_a\hat{\boldsymbol{x}}^n_k(-))^T\right]\\ &\times\boldsymbol{D}_a^T\end{aligned} \tag{41}$$

4.3 Validation gate

At each scan, a validation gate, centered around the predicted measurement of the target, is set up to select the measurements to be associated probabilistically with the target. The validation region is :

$$\left(\boldsymbol{z}^l_k-\hat{\boldsymbol{z}}^{n,l}_k(-)\right)^T\boldsymbol{S}^{n,l^{-1}}_k(-)\left(\boldsymbol{z}^l_k-\hat{\boldsymbol{z}}^{n,l}_k(-)\right)\le d \tag{42}$$

where $\boldsymbol{S}^{n,l}_k$ is the covariance of the innovation corresponding to the true measurement and $\hat{\boldsymbol{z}}^{n,l}_k(-)$ is the predicted measurement of the target.

$$\begin{aligned}\hat{\boldsymbol{z}}^{n,l}_k(-) &= E\left[\boldsymbol{z}^l_k|\boldsymbol{Z}^{k-1,l},\boldsymbol{M}^l_{k-1}\right]\\ &= h(\hat{\boldsymbol{x}}^n_k(-))\end{aligned} \tag{43}$$

$$\begin{aligned}\boldsymbol{S}^{n,l}_k(-) &= E\left[(\boldsymbol{z}^l_k-\hat{\boldsymbol{z}}^{n,l}_k(-))(\boldsymbol{z}^l_k-\hat{\boldsymbol{z}}^{n,l}_k(-))^T|\right.\\ &\left.\boldsymbol{Z}^{k-1,l},\boldsymbol{M}^l_{k-1}\right]\\ &= \boldsymbol{H}_{k,l}\boldsymbol{P}^n_k(-)\boldsymbol{H}^T_{k,l}+\boldsymbol{R}_{k,l}\end{aligned} \tag{44}$$

where $\boldsymbol{H}_{k,l}$ is $\boldsymbol{H}_{k,2,l}$ and d is a gate size parameter from the chi-square distribution with 3 degrees of freedom.

4.4 Filtering

The probability of the individual joint events on scan k, $\beta^l_{k,i,a,b}$, is defined by $\boldsymbol{Z}^{k,l}$ and $\boldsymbol{M}^l_k$.

$$\begin{aligned}\beta^l_{k,i,a,b} &= Pr\left[X^{k,i,l},\Psi^{k,a},\Psi^{k-1,b}|\boldsymbol{Z}^{k,l},\boldsymbol{M}^l_k\right]\\ &= \frac{\gamma^l_{k,i,a,b}\beta^{b^l}_{k-1}}{\sum_{a=1}^{2^L}\sum_{b=1}^{2^L}\sum_{i=1}^{\alpha_k}\gamma^l_{k,i,a,b}\beta^{b^l}_{k-1}}\end{aligned} \tag{45}$$

$$\begin{aligned}\gamma^l_{k,i,a,b} &= \eta_k^{\Psi(X^{k,i,l})}\prod_{j:\tau_j(X^{k,i,l})=1} g(\boldsymbol{z}^l_{k,j};\boldsymbol{z}^{n,l}_{k,a^n}(-),\boldsymbol{S}^{n,l}_{k,a^n}(-))\\ &\times\prod_{n:\delta_n(X^{k,i,l})=1}P^n_D\prod_{n:\delta_n(X^{k,i,l})=0}(1-P^n_DP_{GK})P^{ab}\end{aligned} \tag{46}$$

where $g(\boldsymbol{z}^l_{k,j}; \boldsymbol{z}^{n,l}_{k,a^n}(-), \boldsymbol{S}^{n,l}_{k,a^n}(-))$ is the normal density function with $\boldsymbol{z}^{n,l}_{k,a^n}(-)$ mean and covariance equal to the covariance matrix of $\boldsymbol{S}^{n,l}_{k,a^n}$ for the track to which hit $\boldsymbol{z}^l_{k,j}$ is assigned. P_D is the probability of detecting the target, Ψ is the number of clutter hits, and η_k and P_{GK} are the probabilities that clutter and the target exist in the validation gate, respectively. $\beta^l_{k,i,a,b}$ is the a posteriori probability of which target kinematic model is true and which measurement is true. Eqns. (47) $\sim$ (50) can be obtained using eqn.(45).

$$\begin{aligned}\beta^{n,l}_{k,i,a^n} &= Pr\left[X^{k,i,l}, \Psi_{k,a^n} | \boldsymbol{Z}^{k,l}, \boldsymbol{M}^l_k\right] \\ &= \sum_{\substack{n'=1 \\ n' \neq l}}^{L} \sum_{a^{n'}=1}^{2} \sum_{b=1}^{2^L} \beta^l_{k,i,a,b} \end{aligned} \quad (47)$$

$$\beta^{a^l}_k = Pr\left[\Psi^{k,a} | \boldsymbol{Z}^{k,l}, \boldsymbol{M}^l_k\right] = \sum_{i=1}^{\alpha_k} \sum_{b=1}^{2^L} \beta^l_{k,i,a,b} \quad (48)$$

$$\beta^{n,l}_{k,a^n} = Pr\left[\Psi_{k,a^n} | \boldsymbol{Z}^{k,l}, \boldsymbol{M}^l_k\right] = \sum_{i=1}^{\alpha_k} \beta^{n,l}_{k,i,a^n} \quad (49)$$

$$\beta^{n,l}_{k,i} = Pr\left[X^{k,i,l} | \boldsymbol{Z}^{k,l}, \boldsymbol{M}^l_k\right] = \sum_{a^n=1}^{2} \beta^{n,l}_{k,i,a^n} \quad (50)$$

$$\beta^{n,l}_{k,j} = \sum_{i=1}^{\alpha_k} \beta^{n,l}_{k,i} W^{k,i,l}_{j,n} (j = 1, 2, \cdots, m_{k,l}) \quad (51)$$

$\beta^{n,l}_{k,j}$ is the a posteriori probability that the measurement $\boldsymbol{z}^l_{k,j}$ originated from the target n.

$$\beta^{n,l}_{k,0} = 1 - \sum_{j=1}^{m_{k,l}} \beta^{n,l}_{k,j} \quad (52)$$

If the new set of the measurement $\boldsymbol{Z}_{k,l}$ is obtained, the update equation based on each kinematic model is:

$$\begin{aligned}\hat{\boldsymbol{x}}^n_{k,a^n}(+) &= E\left[\boldsymbol{x}^n_{k,a^n} | \Psi_{k,a}, \boldsymbol{Z}^{k,l}, \boldsymbol{M}^l_k\right] \\ &= \hat{\boldsymbol{x}}^n_{k,a^n}(-) + \boldsymbol{K}^n_{k,a^n} \nu^{n,l}_{k,a^n} \end{aligned} \quad (53)$$

where $\boldsymbol{K}^n_{k,a^n}$ is the filter gain matrix and $\nu^{n,l}_{k,a^n}$ is the combined innovation:

$$\begin{aligned}&\boldsymbol{K}^n_{k,a^n} = \boldsymbol{P}^n_{k,a^n}(-) {\boldsymbol{H}_{k,a,l}}^T \\ &\times \left[\boldsymbol{H}_{k,a,l} \boldsymbol{P}^n_{k,a^n}(-) \boldsymbol{H}^T_{k,a,l} + \boldsymbol{R}_{k,l}\right]^{-1} \end{aligned} \quad (54)$$

$$\nu^{n,l}_{k,j,a^n} = \boldsymbol{z}^l_{k,j} - h(\hat{\boldsymbol{x}}^n_{k,a^n}(-)) \quad (j = 1, \cdots, m_{k,l}) \quad (55)$$

$$\nu^{n,l}_{k,a^n} = \sum_{j=1}^{m_{k,l}} \beta^{n,l}_{k,j} \nu^{n,l}_{k,j,a^n} \quad (56)$$

The updated error covariance matrix is given by:

$$\begin{aligned}&\boldsymbol{P}^n_{k,a^n}(+) = E\left[(\boldsymbol{x}^n_k - \hat{\boldsymbol{x}}^n_{k,a^n}(+))(\boldsymbol{x}^n_k - \hat{\boldsymbol{x}}^n_{k,a^n}(+))^T \right. \\ &\left. | \Psi_{k,a}, \boldsymbol{Z}^{k,l}, \boldsymbol{M}^l_k \right] \\ &= \beta^{n,l}_{k,0} \boldsymbol{P}^n_{k,a^n}(-) + (1 - \beta^{n,l}_{k,0}) \boldsymbol{P}^{n'}_{k,a^n}(+) \\ &+ \boldsymbol{K}^n_{k,a^n} \left[\sum_{j=1}^{m_{k,l}} \beta^{n,l}_{k,j} \nu^{n,l}_{k,j,a^n} \nu^{n,l^T}_{k,j,a^n} - \nu^{n,l}_{k,a^n} \nu^{n,l^T}_{k,a^n}\right] \boldsymbol{K}^{n^T}_{k,a^n} \end{aligned} \quad (57)$$

$$\boldsymbol{P}^{n'}_{k,a^n}(+) = \left(\boldsymbol{I} - \boldsymbol{K}^n_{k,a^n} \boldsymbol{H}_{k,a,l}\right) \boldsymbol{P}^n_{k,a^n}(-) \quad (58)$$

where the dimension of unity matrix $\boldsymbol{I}$ depends on the kinematic model. In the case of model 1, $\boldsymbol{I}$ is 6×6 unity matrix. In another model, it is 9×9 unity matrix.

The combination of the model-conditioned updated state vector and error covariances is calculated as follows:

$$\begin{aligned}\hat{\boldsymbol{x}}^n_k(+) &= E\left[\boldsymbol{x}^n_k | \boldsymbol{Z}^{k,l}, \boldsymbol{M}^l_k\right], \\ &= \sum_{a=1}^{2} \beta^{n,l}_{k,a^n} (\boldsymbol{D}^T_a \hat{\boldsymbol{x}}^n_{k,a^n}(+)) \end{aligned} \quad (59)$$

$$\begin{aligned}&\boldsymbol{P}_k(+) = E\left[(\boldsymbol{x}^n_k - \hat{\boldsymbol{x}}^n_k(+))(\boldsymbol{x}^n_k - \hat{\boldsymbol{x}}^n_k(+))^T | \boldsymbol{Z}^{k,l}, \boldsymbol{M}^l_k\right], \\ &= \sum_{a=1}^{2} \beta^{n,l}_{k,a^n} \boldsymbol{D}_a \left[\boldsymbol{P}^n_{k,a^n}(+) + (\hat{\boldsymbol{x}}^n_{k,a^n}(+) - \boldsymbol{D}_a \hat{\boldsymbol{x}}^n_k(+)) \right. \\ &\left. \times (\hat{\boldsymbol{x}}^n_{k,a^n}(+) - \boldsymbol{D}_a \hat{\boldsymbol{x}}^n_k(+))^T\right] \boldsymbol{D}^T_a \end{aligned} \quad (60)$$

The feature of this method is the fact that the volume of the validation gate changes according to the kinematic model probabilities as shown eqn.(41)[6].

5 Numerical Results

The validity of IMM-JPDA is examined through Monte Carlo simulations of 50 runs.
(1) Target Trajectories

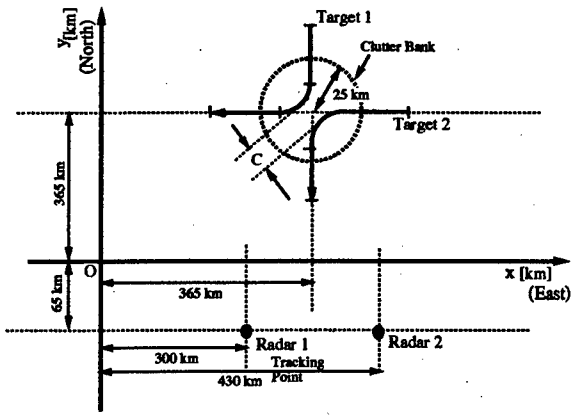


Figure 3: Simulation Scenario 1

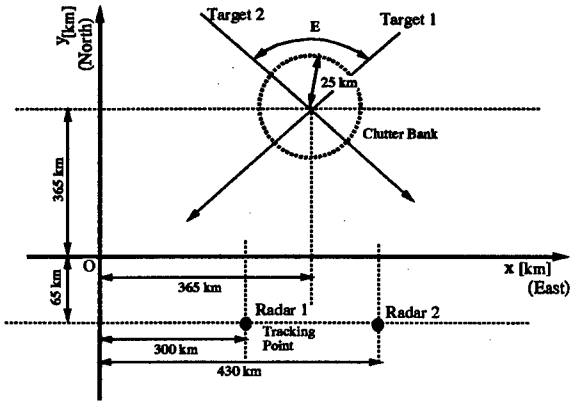


Figure 4: Simulation Scenario 2

Two types of target trajectories are applied in computer simulation. Figs.3 and 4 show simulation scenario examples. First, two targets start as shown in Fig.3. The kinematic characteristics of the target paths are as follows. Targets start moving along two straight lines at a velocity of 340m/s; subsequently targets perform a 90° turn with maneuver acceleration of 5g. The minimum distance C between the two tracks are assumed as 100m, 200m, 300m, 400m and 500m.

Next, two targets start as shown in Fig.4. The constant velocities are 340 m/s and these targets cross each other at 500s at 30, 20 and 10 deg crossing angle E.

(2) Measurement System

Radars having a range measurement standard deviation of 100 meters are located as shown in Figs. 3 and 4. The angle measurement standard deviation is assumed to be 0.85 deg. A sampling interval is set to be 6s throughout the trajectories. The P_{fa} (Probability of false alarm) is 1.0×10^{-6} for clean environment and the P_{fa} is 0.01 for clutter environment assuming clutter being uniformly distributed in the clutter bank shown in Figs. 3 and 3 and the number of clutter being Poisson distributed. The P_D(Probability of detection) is set to 0.9 for both environments.

(3) Tracking Algorithm

Tracking performance of IMM-JPDA was compared with conventional JPDA. Both JPDAs use measurements of two distributed radars shown in Figs. 3 and 4. The JPDA is based on constant velocity model and it has process noises with a standard deviation of positions and velocities, corresponding to $49m/s^2$ of acceleration. This value of the process noise was selected to show the maximum performance in tracking the targets of trajectory scenario 1 at the minimum distance of 500m under the clean environment.

The IMM algorithm consists of two kinematic models. The first model is a six-dimensional constant velocity model with white noise acceleration. The second model is a nine-dimensional constant acceleration model with white noise acceleration. The first model has process noise with a standard deviation of 0.01 m/s^2 and the second model has process noise with a standard deviation of 19.6 m/s^2. The assumed gate size parameter d is 12.83 for Both algorithms.

(4) Tracking Success Rate

Tables 1 and 2 show tracking success (which means that each track ends on the same target for which it started) rates of both JPDA algorithms. IMM-JPDA shows preferable tracking success rate on average compared with JPDA in both scenarios 1 and 2.

Table 1: Tracking Success Rates(Scenario1)[%]

	Clean Environment		Clutter Environment	
C[m]	IMM JPDA	JPDA	IMM JPDA	JPDA
500	92.0	64.0	88.0	60.0
400	92.0	60.0	86.0	56.0
300	90.0	54.0	86.0	52.0
200	84.0	50.0	80.0	48.0
100	78.0	44.0	70.0	40.0

Table 2: Tracking Success Rates(Scenario2)[%]

	Clean Environment		Clutter Environment	
E[deg]	IMM JPDA	JPDA	IMM JPDA	JPDA
30°	88.0	60.0	82.0	58.0
20°	78.0	52.0	70.0	46.0
10°	74.0	42.0	68.0	38.0

(5)Tracking Accuracies

Figs. 5 and 6 show RMS position errors of target 1 by two methods under clean environment. IMM-JPDA method shows better results also in tracking accuracies around the minimum distance point(500s) and the crossing point(500s).

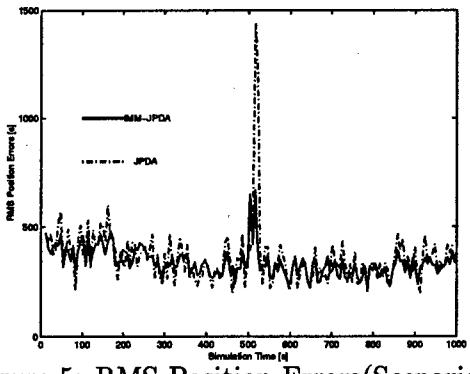


Figure 5: RMS Position Errors(Scenario 1)

6 Conclusion

The JPDA filter using distributed radars to track maneuvering targets has been presented.

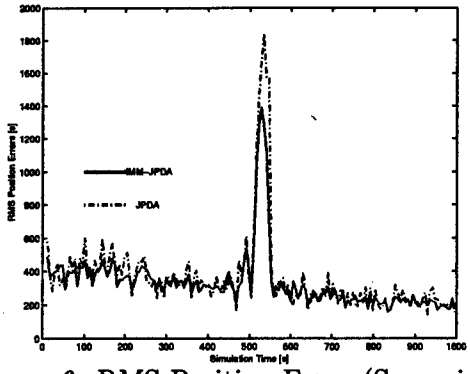


Figure 6: RMS Position Errors(Scenario 2)

The tracking performance of this method and that of the conventional JPDA were evaluated and compared with respect to tracking success rates and tracking accuracies through a Monte Carlo simulation. Our computer simulation results indicated that the enhanced method showed better performance on average.

References

[1] Bar-Shalom, Y. and Fortman, T. E., Tracking and Data Association, Academic Press, New York, 1988.

[2] Bar-Shalom, Y. and Li, X., Estimation and Tracking : Principles, Techniques, and Software, Artech House, MA., 1993.

[3] Bar-Shalom, Y., Multisensor Tracking: Advanced Applications, Artech House, MA., 1989.

[4] Blom, H. A. P., and Bar-Shalom Y. : "The interacting multiple model algorithm for systems with markovian switching coefficients", IEEE Transactions on Automatic Control, 33(Aug. 1988), 780-783.

[5] Houles, A., and Bar-Shalom Y. : "Multisensor tracking of a maneuvering target in clutter", IEEE Transactions on Aerospace and Electronics Systems, 25(Mar. 1989), 176-188.

[6] Kosuge, Y. : "Maneuvering target tracking using multiple maneuver model joint probabilisitc data association", Proc. IECON, '91 pp.2059-2064, 1991.

PARTICLE METHODS FOR MULTIMODAL FILTERING

Christian Musso, Nadia Oudjane
ONERA DTIM. BP 72 92322, France . {musso,oudjane}@onera.fr

Abstract : We present a quick method of particle filter (or bootstrap filter) with local rejection which is an adaptation of the kernel filter. This filter generalizes the regularized filter. The conditional density of the state is recursively estimated. The proposed filter allows a precise correction step in a given computational time. In the context of the 2D tracking problem with angle and/or range measurements, simulations show a better behavior of this filter compared with the Kalman filter and with classical bootstrap filter. We present also some results of a multi model particle filter which can track maneuvering targets.

Keywords : particle filter, boostrap, tracking, non-linear filtering, Monte-Carlo, rejection

1. INTRODUCTION

We consider a target following a noisy dynamical equation which is partially observed, (notations are the same as in [1])

$$X_t = F(X_{t-1}) + V_t \qquad (1)$$
$$Y_t = H(X_t) + W_t \qquad (2)$$

where $F : R^d \longrightarrow R^d$ and $H : R^d \longrightarrow R^q$ are given functions, $(V_t), (W_t)$ are iid variables with densities,

$$V_t \sim p(v)dv, \qquad W_t \sim q(v)dv \qquad (3)$$

X_0, the initial state of density p_0, is assumed independent of $(V_t), (W_t)$. (X_t) is a markov chain

$$X_t / (X_{t-1} = x_{t-1}) \sim p(x - F(x_{t-1}))dx \qquad (4)$$

(Y_t) are independent conditionally to (X_t), and each (Y_t) is, conditionally to X_t, independent from $X_s (s \neq t)$,

$$Y_t / (X_t = x_t) \sim q(y - H(x_t))dy \qquad (5)$$

The Extended Kalman filter (EKF) is widely used to estimate recursively the mean and the variance of the state X_t given the passed measure $Y^t = (Y_1, ..., Y_t)$. The EKF assumes that the conditional density is Gaussian. But, when F or H is highly non-linear, or in case of multimodality, the EKF is inefficient. The goal of the non-linear filtering (NLF) is to estimate the whole law of the state X_t given the measures Y^t. For example, in the tracking context, we will be able to estimate precisely the probability of the presence of a target in any portion of the state space and consenquently to estimate the position of the target. For this filter there is no hypothesis concerning the linearity of F and H and no conditions about the nature of the noise V and W. We want to estimate recursively the conditional density, denoted by $f_{t/t}(x / y^t)$. Suppose we know $f_{t-1/t-1}$, and a new measure y_t is available. Bayes rules give easily the formulation of $f_{t/t}(x / y^t)$ in two steps,

$$f_{t/t-1}(x / y^{t-1}) = \int p_t(x / x_{t-1}) f_{t-1/t-1}(x_{t-1} / y^{t-1}) dx_{t-1} \qquad (6)$$
$$f_{t/t}(x / y^t) \propto q_t(y_t / x) f_{t/t-1}(x / y^{t-1}) \qquad (7)$$

The first step (6) is the prediction step using the dynamical law. The predicted density $f_{t/t-1}(x / y^{t-1})$ is the expectation of $p_t(x / X_{t-1})$ where X_{t-1} follows $f_{t-1/t-1}$. The second step (7) is the correction step. $q_t(y / x_t) = q(y - H(x_t))$ is the likelihood at the point x_t. When the noise W_t is Gaussian, q can be expressed as

$$q_t(y / x_t) = \frac{1}{(2\pi)^{d/2}\sqrt{\det(\Omega)}} \exp\left[\frac{-1}{2}(y - H(x_t))' \Omega^{-1} (y - H(x_t))'\right] \qquad (8)$$

The correction step confronts the new measure with the predicted density. The recursion begins with the assumed known initial density $f_{0/0}(x) = p_0(x)$. A natural way to solve (6) and (7) is to discretize the state space. This can be done when the dimension (d) of this space is low (d<4). Otherwise the computing cost is high. [2]. Another way is to use Monte-Carlo methods.

2. CLASSICAL PARTICLE FILTER ALGORITHM

The aim of the particle filter, called also bootstrap filter or Monte-Carlo filter, is to generate recursively a sample (particles) which follows approximatevely the conditional density $f_{t/t}(x/y^t)$. Suppose we have, at time (t-1) a N-sample $(x_{t-1/t-1}^{(1)},...,x_{t-1/t-1}^{(N)})$ according to $f_{t-1/t-1}$. The integral (6) can be approximated by the empirical expectation. In other words, the density $f_{t-1/t-1}$ is approximated by the empirical measure $(1/N)\sum_{j=1}^{N}\delta_{x_{t-1/t-1}^{(j)}}(x)$ which puts uniformly the mass on the particles. (6) and (7) become,

$$\hat{f}_{t/t-1}(x/y^{t-1}) = \frac{1}{N}\sum_{j=1}^{N} p_t(x/x_{t-1/t-1}^{(j)}) \qquad (9)$$

$$\hat{f}_{t/t}(x/y^t) \propto \sum_{j=1}^{N} q_t(y_t/x)p_t(x/x_{t-1/t-1}^{(j)}) \qquad (10)$$

The error of this approximation is independent of the dimension (d) and is of order $1/\sqrt{N}$ (law of large numbers). Therefore, unlike discretization methods, Monte-Carlo methods can, theorically , deal with large (d) with a reasonable computing cost. There is two classical ways to generate sample from (9) and (10).

2.1 The weighted resampling method (SIR)

These filters called SIR filter (Sampling Importance Resampling) ([3],[4]) or IPF (Interacting Paticle Filter) ([5]) or CONDENSATION (Conditional Density Propagation) ([6]) first generate a sample from (9). This can be done by,

1 Generate I uniform on {1,...,N} (11)

2 Generate X according to $p_t(x/x_{t-1/t-1}^{(I)})$ (12)

Step 1 is a bootstrap algorithm and step 2, for fixed $x_{t-1/t-1}^{(I)}$, gives a predicted particle $x_{t/t-1}^{(I)}$ according the dynamical law (1). This algorithm produces an iid sample $(x_{t/t-1}^{(1)},..,x_{t/t-1}^{(N)})$. Then in (7), $f_{t/t-1}(x/y^{t-1})$ is approximated by the empirical density $(1/N)\sum_{j=1}^{N}\delta_{x_{t/t-1}^{(j)}}(x)$,

$$\hat{f}_{t/t}(x) = \sum_{j=1}^{N} w_j \delta_{x_{t/t-1}^{(j)}}(x) \qquad (13)$$

where $w_j = q_t(y_t/x_{t/t-1}^{(j)}) / \sum_{j=1}^{N} q_t(y_t/x_{t/t-1}^{(j)})$ is the weight of each particle proportional to the likelihood. Generating a sample from (13) can be done by (correction step),

1 Generate I , $P(I=i) = w_i$ (multinomial) (14)

2 Put $X = x_{t/t-1}^{(I)}$ (15)

The most likely predicted particle are the most duplicated. (11) (12) (14) (15) produce quickly a new iid sample $(x_{t/t}^{(1)},...,x_{t/t}^{(N)})$.

2.2 The rejection method (RM)

The RM [1],[4] generate the predicted particle $(x_{t/t-1}^{(1)},...,x_{t/t-1}^{(N)})$ with (11) and (12), like in SIR. But approximation (13) is avoided. The RM produces a exact sample according to (10). It is easy to check that the following algorithm generates this sample ,

1. Generate I uniform on {1,...,N}

2 Generate X ~ $p_t(x/x_{t-1/t-1}^{(I)})$ and U uniform on [0,1] (16)

3. If $q_t(y_t/X) \geq c_t U$, accept X, $x_{t/t}^{(j)}$=X and j=j+1 (17)

where $c_t \geq \sup_x q_t(y_t/x)$. Steps 1,2,3 are repeated to get the desired size of the sample. For the rejection method, the correction step is exact for fixed N, unlike the weighted sample method. However, in this form, the computing cost is high. Indeed, the probability that (16)+(17) produce a sample X (acceptance probability) is proportional to c_t^{-1}. The maximum of q(y/.) can be high (see

for example (8) $c_t^{-1} = (2\pi)^{d/2}\sqrt{\det(\Omega)}$). SIR and RM have a serious drawback. In case of low dynamical noise, we observe that in multiplying the high weighted particles, the prediction step will explor poorly the state space. There is with time a degeneracy phenomenon. The particle clouds will concentrate on a few points of the state space. The discrete nature of the (weak) approximations reduce the exploring capacity. Therefore, it is useful to generate a sample from a smooth distribution which approximate sthe underlying distribution $f_{t/t}$ which is assumed to be smooth.

2.3 The kernel and the regularized methods

Hurzeler and Kunsch [1] have introduced the kernel filter (KF) which uses local rejection and density kernel estimation. The new algorithm proposed in this paper (L2RPF) is an adaptation of the KF. Regularized particle methods (RPF) that have been proposed in ([6]...[9]), deal with weighted sample methods. The RPF (version where regularization is made after correction) estimates $f_{t/t-1}$ by a non-parametric density estimation using a kernel (K). (13) becomes,

$$\hat{f}_{t/t}(x) = \frac{1}{h}\sum_{j=1}^{N} w_j K[h^{-1}A_t(x - x_{t/t-1}^{(j)})] \qquad (18)$$

where h is the bandwith, K the kernel which is itself a density, A_t^{-1} the root of S the covariance matrix of the particles $x_{t/t-1}^{(j)}$ ($A_t A_t' = S^{-1}$). The algorithm,

1. Generate I according $P(I = i) = w_i$ (19)
2. Generate $Z \sim K(x)dx$ (20)
3. Put $X = x_{t/t-1}^{(I)} + hA_t^{-1}Z$ (21)

produces a sample according to (18). The regularization (21) improve the exploring capacity. Note that h=0 gives the SIR (14), (15). K and h are chosen in order to minimize the L^2 error , MSE(K,h) $= \int (\hat{f}_{t/t}(x) - f_{t/t}(x))^2 dx$ [10], [11], among the even kernels of L^2 norm equal to 1)

$$K(x) = \frac{1}{2}c_d^{-1}(d+2)(1-\|x\|^2) \text{ if } \|x\| \le 1 \qquad (22)$$

0 otherwise. c_d is the volume of the unity sphere. The optimal h is,

$$h = A(K)N^{-1/(d+4)}/2 \text{ with}$$
$$A(K) = \left[8c_d^{-1}(d+4)(2\sqrt{\pi})^d\right]^{1/(d+4)} \qquad (23)$$

It is important to whiten the particle before the regularisation because h is the same in all directions. Note that the MSE depends now on the dimension (d) with the optimal h.

3.THE L2RPF FILTER

3.1 Description of the filter

The Local Rejection Regularised Particle Filter allows a precise correction step in a given computational time. Given $(x_{t/t-1}^{(j)})_{j=1...,N}$ and a scalar α_t, we generate a corrected sample with the following algorithm, (24), (25), (26)

1.Generate I , $P(I = i) \propto c_{t,i}(\alpha_t)$
2.Generate $Z \propto K(x)dx$, U uniform on [0,1]
3. Put $X = x_{t/t-1}^{(I)} + hA_t^{-1}Z$
4 If $q_t(y_t / X) \ge \alpha_t\, c_{t,I}(\alpha_t)$U, we accept X, $x_{t/t}^{(j)}$=X and j=j+1 (27)

The coefficients $c_{t,i}(\alpha_t)$ (computed below) satisfy,

$$c_{t,j}(\alpha_t) \ge \sup_{x \in \Sigma_j(\alpha)} q_t(y_t / x) \qquad (28)$$

$\Sigma_j = \{x / (x - x_{t/t-1}^{(j)})' S^{-1}(x - x_{t/t-1}^{(j)}) \le \alpha_t^2 h^2\}$ is a local ellipsoid centered on the particle $x_{t/t-1}^{(j)}$. α_t is a control parameter between 0 and 1.

Proposition 3.1 : the L2RPF algorithm produce a sample according to

$$\hat{f}_{t/t}^{\alpha_t}(x) \propto \sum_{i=1}^{N}[c_{t,i}(\alpha_t)\min(1, \frac{q_t(y_t / x)}{\alpha_t c_{t,i}(\alpha)})]$$
$$K[h^{-1}A_t(x - x_{t/t-1}^{(i)})] \qquad (29)$$

Indeed, with « ≤ » in R^d in the « coordinate by coordinate » sense and putting $g(x) = q_t(y_t / x)$, the 3 independent variables being I, U and Z, we have

$$P(X \le x) \propto \sum_{i=1}^{N} \iint_{\substack{g(x^i+hA^{-1}z) \ge \alpha c_i(\alpha) u \\ x^i+hA^{-1}z \le x}} K(z)\ c_i(\alpha)\, dz\, du \propto$$

$$\sum_{i=1}^{N} \int_{x^i+hA^{-1}z \le x} [c_i(\alpha) \min(1, \frac{g(x^i+hA^{-1}z)}{\alpha c_i(\alpha)})]\, K(z)\ dz$$

(Putting $zz = x^i + hA^{-1}z$, it becomes)

$$\propto \sum_{i=1}^{N} \int_{zz \le x} [c_i(\alpha) \min(1, \frac{g(zz)}{\alpha_t c_i(\alpha_t)})]$$

$$K(h^{-1}A(zz - x^i))dzz$$

We obtain (29) after derivating the last expression w.r.t « x ».

Proposition 3.2 : the acceptance probability P_a of the L2RPF is with

$$c = (\sum_{i=1}^{N} c_{t,i}(\alpha))^{-1}, \qquad (30)$$

$$P_a = c \sum_{i=1}^{N} \int [c_i(\alpha_t) \min(1, \frac{g(x_{t/t-1}^{(i)} + hA^{-1}z)}{\alpha_t c_i(\alpha_t)})]$$

$$K(z)dz \qquad (31)$$

$$P_a(\alpha) \approx c \sum_{i=1}^{N} c_i(\alpha_t) \min(1, \frac{g(x_{t/t-1}^{(i)})}{\alpha_t c_i(\alpha_t)}) \qquad (32)$$

(31) is computed like in (29). (32) is obtained using an expansion of $g(x_{t/t-1}^{(i)} + hA^{-1}z)$ around $h = 0$. This approximation is in general precise. If we put α=1 in (29) the « min » is $c_{t,i}^{-1}(\alpha) q_t(y_t / x)$, because c_i is a local maximum (28),

$$\hat{f}_{t/t}^{\alpha_t=1}(x) \propto \sum_{i=1}^{N} [q_t(y_t / x)\ K[h^{-1}A_t(x - x_{t/t-1}^{(i)})]$$

which is the KF with the exact correction. In this case P_a is minimal (32), the computational cost is maximal. If we put α=0 in (29) the « min » is 1 and $c_i(\alpha = 0) = w_i$ (Σ_j reduces to a particle). We obtain the RPF (18). In this case P_a=1, the computational cost is low. Note that $P_a(\alpha)$ decrease swhen α increases At each time, the choice of α is done by the following manner : we keep the maximal value of α such as $P_a(\alpha) \ge P_a^{\min}$ (with a coarse discretization of [0,1]). $P_a^{\min}$ is given by the computing capability . The higher α is, better the correction. When α is chosen, we put $N_e = N / P_a(\alpha_t)$ the number of test-samples which enter in the loop (24-27). In practice, α_t is close to 0 for the first measures, then increases to 1 when the particles concentrate on likely regions of the state-space. Now we present a fast method to compute $c_{t,i}(\alpha)$

3.2 computing the $c_{t,i}(\alpha_t)$ coefficient

By Lagrangian methods, we can see that the coordinates $(x_1,...,x_d)$ of a point in Σ_j verifies(1≤i≤d), (33)

$$x_i^{\min} = x_i^j - \alpha h \sqrt{S_{ii}} \le x_i \le x_i^j + \alpha h \sqrt{S_{ii}} = x_i^{\max}$$

where S_{ii} is on the diagonal of S. Let C_j being the hyper-cube $\{x / x_i^{\min} \le x_i \le x_i^{\max}, 1 \le i \le d\} (\Sigma_j \subset C_j.)$.

$c_{t,i}(\alpha_t)$ will be the maximum of g on C_j. Assume that the measure function $(H_k(.), k = 1,...,q)$ (2) is locally decreasing for one coordinate or increasing for an other. For example if we measure an angle $H_k(x)$=arctg(x_1/x_2), H_k increases when x_1 increases and H_k decreases when x_2 increases (if x_1, x_2 >0). The extreme values of H_k are, (34)

$$H_k^{\min} = H_k(x_m^{extr}) \le H_k(x) \le H_k(x_M^{extr}) = H_k^{\max}$$

where $x_{(.)}^{extr}$ equals $x_i^{\min}$ or $x_i^{\max}$. Suppose that the q components of the measure noise W are independent and that q(.) (5) decreases around the origine, it can be seen that the maximum of the likelihood on C_j ($c_{t,j}$= sup(g(x)) is, (35),

$$\max_{x \in C_j} \prod_{k=1}^{q} q_{W_k}(y_k - H_k(x)) = \prod_{k=1}^{q} q_{W_k}(y_k - H_k^{extr})$$

where $H_k^{extr} = H_k^{\min}$ if $y_k \le H_k^{\min}$, $H_k^{extr} = y_k$ if $H_k^{\min} \le y_k \le H_k^{\max}$ and $H_k^{extr} = H_k^{\max}$ if $y_k \ge H_k^{\max}$

4. SIMULATIONS

We present three 2D-tracking problems. L2RPF is applied in each problem execpt

the last one. The computing cost is about 30 times bigger than the EKF with P_a^{min}=0.2. The number of particles (N) is 5000. The number of Monte-Carlo (MC) is 50. In these problems, the dynamical noise level is equal to zero. So, we can easily compute the Cramer-Rao Lower Bound.

4.1 Bearing only

The target has a uniform straight motion $(x_t^1, \dot{x}^1, x_t^2, \dot{x}^2)$=(10km, -10m/s, 10km, 10m/s) (figure (1)). The observer (xo_t, yo_t) is on the origine at time 0. He has a 2 legs motion: (50m/s 0m/s) speed for the first (during 100s) and (-50m/s 50m/s) for the second. During 200s, the observer measures every second a noisy angle with standard deviation (std)=0.5°, $H(X_t) = arctg((x_t^1 - xo_t)/\ (x_t^2 - yo_t))$. The initial estimate X(0/0) (center of the cloud) of the target is a Gaussian variable centered on the true position with covariance matrix, P(0/0)=diag(5km, 30m/s, 5km, 30m/s). Figure (1) shows the estimated trajectory (center of the cloud) of the L2RPF. Figure (3) shows the evolution of the acceptance probability with the corresponding control parameter α (Figure (2)). For each MC trial we compute the trajectory estimation error in order to obtain the std (for the 50 trial) of the target position. As you can see on figure (5), the std of the horizontal position of the target is very close to the Cramer-Rao Lower Bound (CRLB) (without bias (figure (4)). In this context, (observability problem) the EKF has diverged 5 times over 50. .

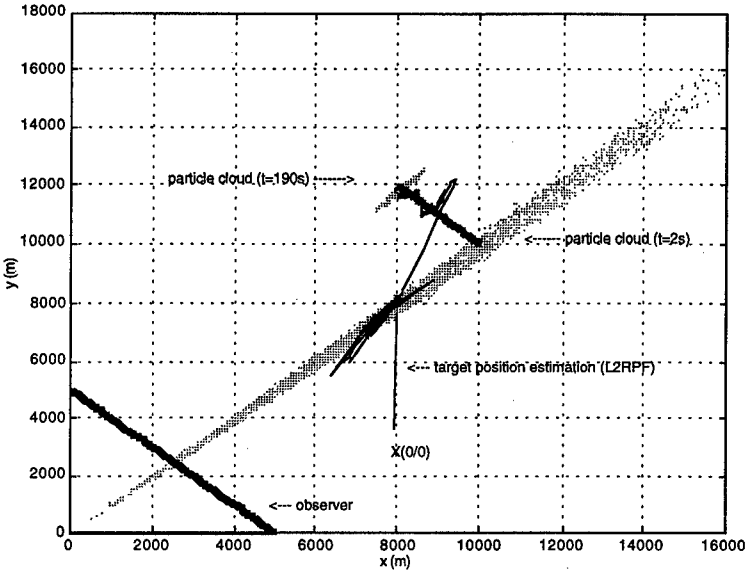


Figure (1) : true and estimated trajectories

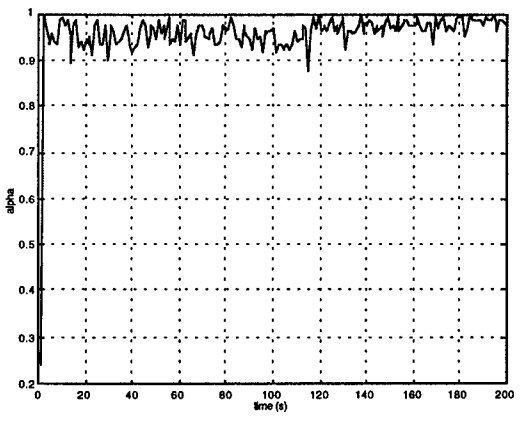


Figure (2) : control parameter evolution

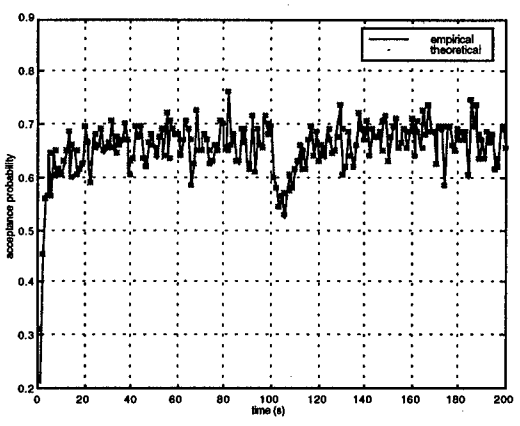


figure (3) : acceptance probability

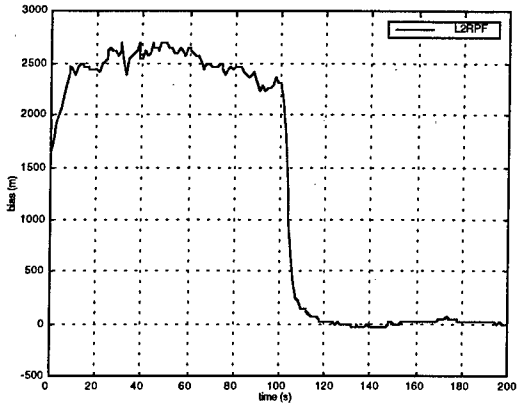
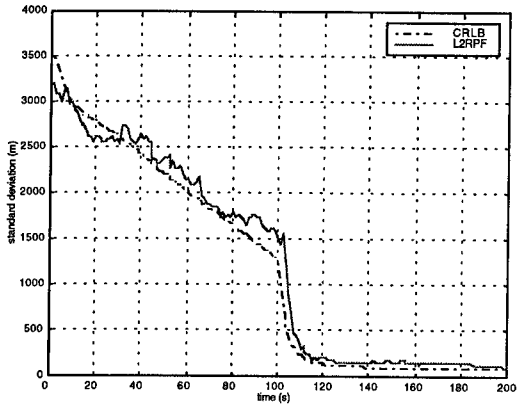


Figure (4) : bias for x-position



Figure(5) : standard deviation for x-position

4.2 Range and Bearing

The target has a uniform straight motion $(x_t^1, \dot{x}^1, x_t^2, \dot{x}^2)$=(5km, -20m/s, 5km, -20m/s). The observer is on the origin. During 200s, the observer measures every second a noisy angle with std=1° and a range with std=1m (very precise), $Y_t = [arctg(x_t^1 / x_t^2), \sqrt{(x_t^1)^2 + (x_t^2)^2}]$. The initial estimate X(0/0) of the target is a Gaussian variable centered on the true position with covariance matrix, P(0/0)=diag(0.5km, 50m/s, 0.5km, 50m/s). Results with the 50 MC are shown below, L2RPF and EKF are compared. Figure (6) shows the x-position estimator bias. And we observe on Figure (7) that, unlike the EKF, the L2RPF converge rapidly to the CRLB. RPF has been performed. The results are comparable with the L2RPF for the std. But the variance of the clouds are bigger with the RPF. Error estimation is more precise with the L2RPF.

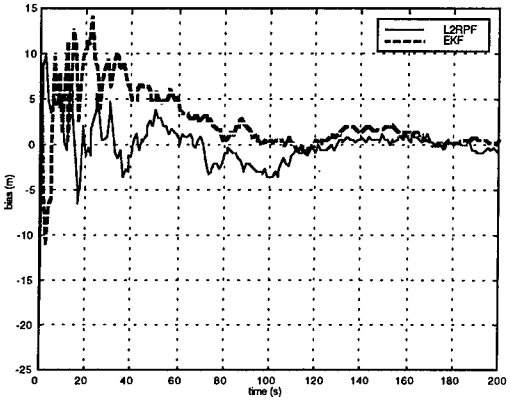


Figure (6) : x-position bias

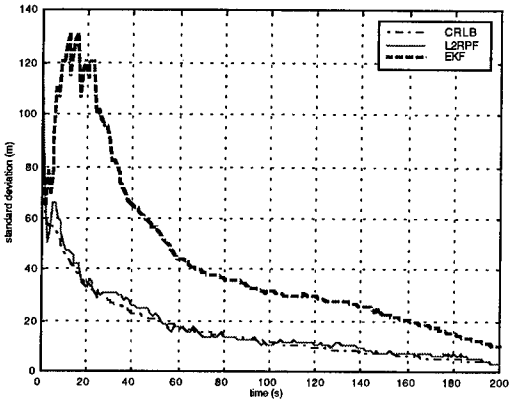


Figure (7) : x-position std

4.3 Multiple Model Particle Filter (MMPF)

The MMPF is presented in [12]. By means of the formalism of Interacting Multiple Model [13] where the dynamical model θ_t of the target has to be estimated among some fixed models. This is a case of multi-modality. We suppose that $\{\theta_t\}$is a discrete Markov chain with a given transition matrix. Therefore, we can apply the theory of the particle filter with the new augmented state E_t,

$$E_t = (X_t, \theta_t) \qquad (36)$$

Assuming that, given θ_{t-1}, θ_t is independent of X_{t-1}, prediction step (6) is given by,

$$P(E_t = e_t / y^{t-1}) = \int P(X_t / \theta_t, X_{t-1}) P(\theta_t / \theta_{t-1}) P(E_{t-1} = e_{t-1} / y^{t-1}) de_{t-1} \quad (37)$$

If we have a sample $\{E^i_{t-1}\}$ from E_{t-1} / y^{t-1}, the following algorithm produces a predicted sample $\{E^i_{t/t-1}\}$ according to (37), for each particle $E^i_{t-1} = (X^i_{t-1}, \theta^i_{t-1})$,

1 Generate θ^i_t according to $p(\theta_t / \theta_{t-1} = \theta^i_{t-1})$
2 Generate $X^i_{t/t-1}$ according to $p(X_t / \theta^i_{t-1} X^i_{t-1})$

Correction step is done with the RPF version ((18)...(21)). The updated sample state $\{X^i_{t/t}\}$ is given by the marginal ditribution of $\{E^i_{t/t}\}$.

In our simulation, the target can have 2 motions : uniform straight motion (USM) and a turn with constant velocity (2 state models). Figure (8) shows the geometry. The observer placed on the origin measure sbearing (std=1°) and range (std=20m) every 10s. The duration of the first USM is 600s, the duration of the turn is 800s, and the duration of the last USM is 600s. The initial USM mode probability is 0.99, and the transition markov matrix $p(\theta_t / \theta_{t-1})$ is $\begin{bmatrix} 0.98 & 0.02 \\ 0.02 & 0.98 \end{bmatrix}$. The initial estimate X(0/0) of the target is a Gaussian variable centered on the true position with covariance matrix, P(0/0)=diag(450m, 63m/s, 42m, 60m/s). The number of Monte-Carlo is 100.
Classical IMM filter and the MMPF are compared. MMPF estimates the angular turn rate (dimension of the state=5). The IMM knows this rate (otherwise for this context, IMM is not stable). Nevertheless, the behavior of the 2 filters are comparable. Probabilities of the USM mode are shown in Figure (9), they follow the change of the dynamic. On Figure (10) we can see a good angular turn rate estimation for the MMPF.

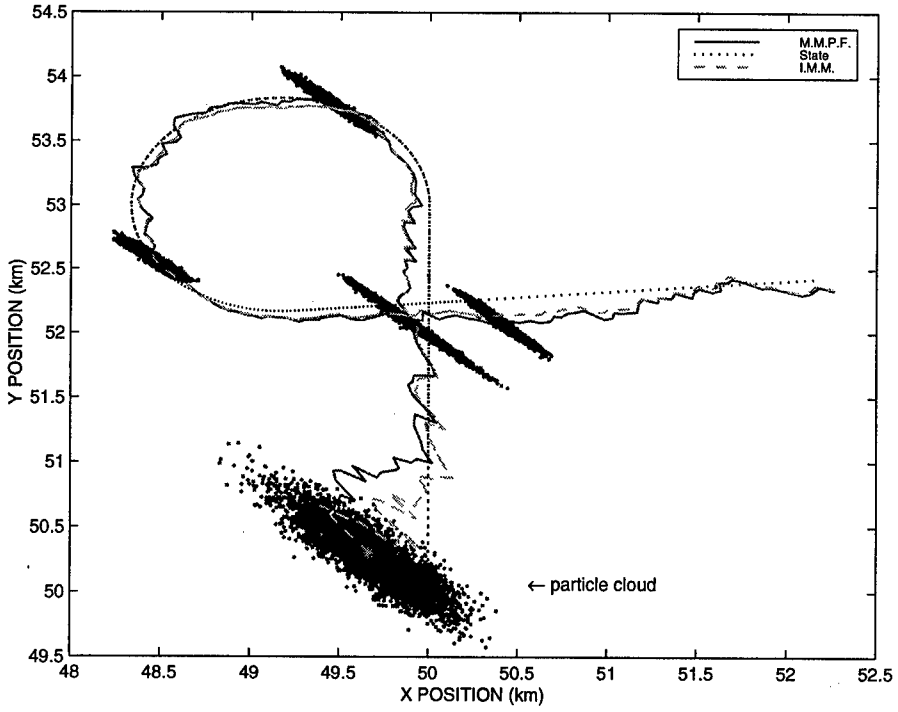
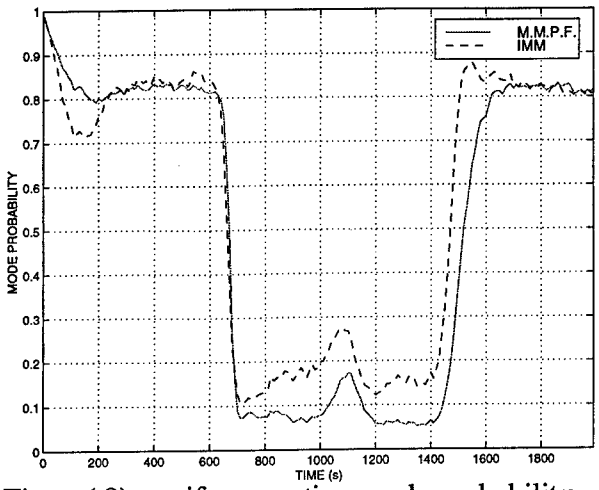


Figure (8) : true and estimated trajectories for the 2 filters



Figure(9) : uniform motion mode probability

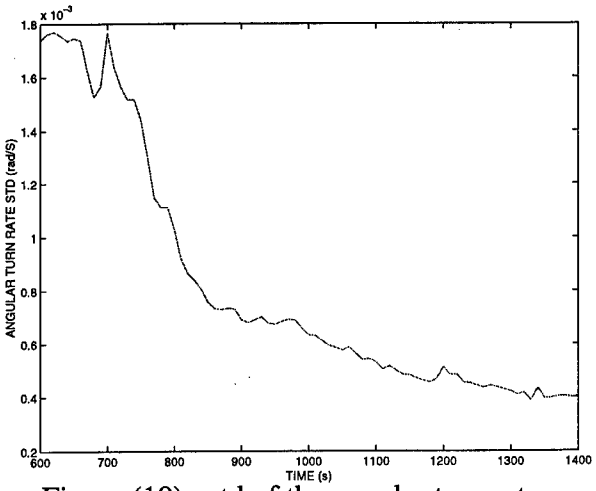


Figure (10) : std of the angular turn rate

REFERENCES

[1]Markus Hurzeler, Hans R. Kunsch. « Monte Carlo approximations for general state space models » Research Report No. 73 April 1995. Seminar fur Statistik. CH-8092 Zurich; Switzerland

[2] CAI Zhiqiang, F. Le Legland, Zhang Huilong, « An adaptative local grid refinement method for nonlinear filtering », rapport de recherche 2679, INRIA, Oct. 1995

[3] N.Gordon, D. Salmond, A.Smith. « Novel approach to nonlinear/ non-Gaussian Bayesian state estimation ». IEE Proceedings, Part F, vol 140, pp. 107-113, April. 1993

[4] M. K. Pitt, N. Shephard. « Filtering via simulation : auxiliary particle filter » Oct, 22 1997

[5]P. Del Moral. « Nonlinear filtering : interacting particle solution », Markov Processes and related Fields, vol. 2, no4, pp 555-580. 1996

[6] M. Isard, A. Blake. « Contour tracking by stochastic propagation of conditional density ». Proc European Conf. Computer Vision. pp 343 Cambridge, UK, 1996

[6]C. Musso, N. Oudjane. « regularisation schemes for branching particle systems as a numerical solving method of the nonlinear filtering problem ». ISSC 1998. Dublin (Ireland) June 25-26 1998

[7] N. Oudjane, C. Musso. « regularised particle schemes applied to the tracking problem ». IRS 98. Munchen (Germany), September 15-17, 1998

[8] F. Le Legland, C. Musso, N. Oudjane. « An analysis of regularized interacting particle methods for nonlinear filtering ». Preprints of the 3rd IEEE European Workshop on Computer--Intensive Methods in Control and Data Processing, Prague 1998. pp 167-174. September 1998

[9] C. Musso, , N. Oudjane. « methodes statistiques en filtrage particulaire appiqué au pistage ». International Conference on Radar Systems. Brest. France Mai 1999

[10] B. W. Silverman « Density Estimation for Statistics and Data Analysis ». Chapman & Hall 1986

[11] L. Devroye, Laszlo Gyorfi. « Nonparametric Density Estimation ». John Wiley & Sons 1984.

[12] N. Oudjane, C. Musso. « Multiple Model Particle Filter ». 17ème colloque Gretsi sur le traitement du signale et des images. Vannes 13-17 Septembre 1999

[13] A.P. Henk Blom and Yaakov Bar-Shalom. « The interacting multiple model algorithm for systems with markovian switching coefficients. IEEE Transactions On Automatic Control 780:783, 1998

An Algorithm for Quasi-hierarchy Fusion Estimation with Transforming Observation Values

Hongyan Sun, Kezhong He, Bo Zhang

The State Key Lab of Intelligent Technology & Systems
Dept. of Computer Science & Technology
Tsinghua University
Beijing, 100084, P.R.China
Hongyansun@263.net, Tel:++86-10-62783191

Shiyi Mao

Dept. of Electrical Engineering
Beijing University of
Aeronautics& Astronautics
Beijing, 100083, P.R.China

Abstract *An algorithm of quasi-hierarchy fusion estimation with transforming observation values is given in multisensor systems that the target state model is linear, the observation model is non-linear and this non-linear function has a inverse function. Second, its properties are analyzed and the quasi-hierarchy fusion formulas of target state with better properties are further obtained. Third, the realization architecture of above algorithm is also presented. Finally, the algorithm is applied to multi-radar tracking systems. And a new preprocessing method is used in the target state quasi-hierarchy fusion estimation with transforming observation values in multi-radar tracking systems, i.e. it is made use of to increase one dimension (angular velocity) in observation data and reestablished observation model, then the quasi-hierarchy fusion estimation with transforming observation values is made. Thus, the feasibility of this algorithm is proved.*

Key words: *fusion, state estimation, target tracking.*

1. Introduction

In multi-radar tracking systems, radar observation values (position, azimuth, and position rate of change) are obtained under polar-coordinate systems, but its target state tracking is completed under Cartesian coordinate systems. So, radar observation values are a non-linear function of target state. Thus, it is proposed in multi-radar tracking systems that the multisensor system state estimation under the non-linear observation model.

For the single sensor systems that the target state model is linear, the observation model is non-linear and this non-linear function has a inverse function, the system state estimation is obtained by transforming observation values and general Kalman filtering.

Here, the target state estimation of non-linear multisensor systems is studied.

Data fusion is a new available technology in multisensor data processing. And data fusion is the key in multisensor multitarget tracking systems. It combines data from multiple(and possibly diverse)sensors, perform the track with the higher quality than of the single sensor, provide more useful information than the sum based on data from the separate sensor.

The hierarchical fusion is an important method of multisensor data processing. In multisensor multitarget tracking systems, It is frequently referred to the sensor level tracking where each local sensor maintains its own track file based only on its own data. The tracks from the various sensor are transmitted to a single central processor which is responsible for fusing the tracks to form a central track file, which may be fed back to the each sensor. This approach overcomes the large communication and high computation loads that the centralized fusion

approach has to do. And it is much easier to implement than the distributed sensor network approach. The main disadvantage of the approach is the need for two types of algorithms: one for sensor level tracking and the other for data fusion.[3]

Based on characteristics of above the non-linear systems and advantages of multisensor hierarchy fusion, an algorithm for quasi-hierarchy fusion estimation with transforming observation values is proposed in this paper. First, quasi-hierarchy fusion estimation formulas with transforming observation values are given in multisensor systems that the target state model is linear, the observation model is non-linear and this non-linear function has a inverse function. Second, its properties are analyzed and the quasi-hierarchy fusion formulas of target state with better properties are further obtained. Third, the realization architecture of above algorithm is also presented. Finally, for multi-radar tracking systems, a new preprocessing method is used in the target state quasi-hierarchy fusion estimation with transforming observation values in multi-radar tracking systems, i.e. it is made use of to increase one dimension (angular velocity: an angle rate of change which is a observable function of most radial velocity and azimuth) in observation data and reestablished observation model, then the quasi-hierarchy fusion estimation with transforming observation values is made. It decreases the load of usual preprocessing method that changes the position and azimuth under polar-coordinate systems into the component of velocity of the target state with two sensor observation values or a sensor prediction value.

2. Mathematics Model

If systems state model is described by

$$X(t+1) = F(t)X(t) + G(t)W(t), \quad t = 1,2,... \quad (1)$$

where $X(t) \in R^n$ is the target state vector at time t , $F(t) \in R^{n \times n}$ is the known and possibly time-varying matrix, $G(t) \in R^{n \times q}$ is the driven matrix, $W(t) \in R^q$ is a zero-mean white Gaussian vector noise.

Given observation systems of N sensors, the ith sensor observation model is described by

$$Z_i(t) = h_i[X(t),t] + V_i(t), \quad i = \overline{1,N} \qquad (2)$$

where $Z_i(t) \in R^n$ is the observation vector of ith sensor, $h_i[X(t),t] \in R^n$ is the non-linear vector function of $X(t)$ of the ith sensor, and $h_i[X(t),t] = [h_{i1}[X(t),t],...,h_{in}[X(t),t]]^T$ $V_i(t) \in R^n$ is the vector noise. The priori information of systems is as following:

(1) $h_i^{-1}[.,t]$ ($i = \overline{1,N}$) exists,

(2) $W(t), V_i(t)$ ($i = \overline{1,N}$) are zero-mean white noises and are assumed to be independent of each other and the initial state $X(t_0)$, i.e. for $t > t_0$,

$$E\{W(t)\} = 0, \quad E\{W(t)W^T(\sigma)\} = Q(t)\partial(t-\sigma)$$

$$E\{V_i(t)\} = 0, \quad E\{V_i(t)V_j^T(t)\} = R_i(t)\partial_{ij}(t-\sigma)$$

$$E\{W(t)V_i^T(t)\} = 0$$

$$E\{X(t_0)W^T(t)\} = E\{X(t_0)V_i^T(t)\} = 0$$

(3) Initial state is subject to the canonical distribution function, i.e.

$$E\{X(t_0)\} = \overline{X}(0)$$

$$E\{[X(t_0) - \overline{X}(0)][X(t_0) - \overline{X}(0)]^T\} = P(0)$$

The global observation equations based on N sensors are described by

$$Z(t) = h[X(t),t] + V(t) \qquad (3)$$

where $Z(t) \in R^{nN}$ is the observation vector which

consists of N sensor's, $h[X(t),t] \in R^{mN}$ is the non-linear vector function of and $X(t)$

$h[X(t),t] \triangleq [h_1[X(t),t]^T,...,h_N[X(t),t]^T]^T$,

$V(t) \in R^{mN}$ is the vector noise and

$V(t) = [[V_1(t)]^T,...,[V_N(t)]^T]^T$, its covariance matrix is described by

$R(t) = diag[R_1(t),..., R_N(t)]$

The above models (1) and (2) as well as (1)and (3) formulates the single sensor and multisensors system models, respectively.

3. Transforming Observation Values

Given the model (2), $h_i[X(t),t]$ has a inverse function, i.e. $h_i^{-1}[.,t]$ exists, so a new observation value $\eta_i(t)$ may be obtained:

If $V_i(t) = 0$, then $\eta_i(t) = h_i^{-1}[Z_i(t),t] = X(t)$ (4)

If $V_i(t) \neq 0$, then the noise $V(t)$ and true value $X(t)$ will all influence $\eta_i(t)$, so let $V_i(t) = \overline{V}_i(t)$, where $\overline{V}_i(t)$ is the zero-mean Gassian noise, i.e. if the covariance matrix of $V_i(t)$ is very small, then

$$\eta_i(t) = X(t) + \overline{V}_i(t) \tag{5}$$

where

$$\overline{V}_i(t) = \frac{\partial h_i^{-1}[Z_i(t),t]}{\partial [Z_i(t)]} dZ_i$$

and its approximately covariance matrix of $\overline{V}_i(t)$ is

$$\overline{R}_i(t) = [\frac{\partial h_i^{-1}[Z_i(t),t]}{\partial [Z_i(t)]} dZ] R_i(t) [\frac{\partial h_i^{-1}[Z_i(t),t]}{\partial [Z_i(t)]} dZ]^T$$

It is a function of observation values $Z_i(t)$ and $R_i(t)$. Let

$$Z_i(t) = \hat{Z}_i(t|t-1) = h_i[\hat{X}(t|t-1),t]$$

where $\hat{X}(t|t-1)$ is the fusion prediction estimation value, then

$$\overline{R}_i(t) = [\frac{\partial h_i^{-1}[\hat{Z}_i(t|t-1),t]}{\partial [\hat{Z}_i(t|t-1)]} dZ] R_i(t) [\frac{\partial h_i^{-1}[\hat{Z}_i(t|t-1),t]}{\partial [\hat{Z}_i(t|t-1)]} dZ]^T$$

and $i = \overline{1,N}$.

Thus, the global linear observation model based on N sensors is by

$$\eta(t) = X(t) + \overline{V}(t), \tag{6}$$

where $\eta(t) = [[\eta_1(t)]^T,...,[\eta_N(t)]^T]^T$ is the observation value after transforming observation value $Z(t) \in R^{mN}$,and

$\overline{V}(t) = [[\overline{V}_1(t)]^T,...,[\overline{V}_N(t)]^T]^T$ is the zero-mean white noise, i.e.

$\overline{V}(t) \sim (0, \overline{R}(t))$, $\overline{R}(t) = diag(\overline{R}_1(t),..., \overline{R}_N(t))$.

So, the equations (5) and (6) are a single sensor and multisensor systems observation models after transforming observation value $Z_i(t) \in R^m$ and $Z(t) \in R^{mN}$ respectively.

4. An Algorithm for Quasi-hierarchy Fusion and its Application

4.1 Quasi-hierarchy fusion equations

The ith (I=1,2,...,N) sensor Kalman filtering estimation is obtained with model (1) and (5) as following:

$$\hat{X}_i(t|t) = \hat{X}_i(t|t-1) + P_i(t|t)[\overline{R}_i(t)]^{-1}[\eta_i(t) - \hat{X}_i(t|t-1)] \tag{7a}$$

$$P_i^{-1}(t|t) = P_i^{-1}(t|t-1) + [\overline{R}_i(t)]^{-1} \tag{7b}$$

$$\hat{X}_i(t|t-1) = F(t-1)\hat{X}_i(t-1|t-1) \tag{7c}$$

$$P_i(t|t-1) = F(t-1)P_i(t-1|t-1)[F(t-1)]^T + G(t-1)Q(t-1)[G(t-1)]^T \tag{7d}$$

The global Kalman filtering of all sensors is

obtained with models (1) and (6) as following:

$$\hat{X}(t|t)=\hat{X}(t|t-1)+P(t|t)H^T[\overline{R}(t)]^{-1}[\eta(t) - H\hat{X}(t|t-1)] \quad (8a)$$

$$P^{-1}(t|t)=P^{-1}(t|t-1)+H^T[\overline{R}(t)]^{-1}H. \quad (8b)$$

$$\hat{X}(t|t-1)=F(t-1)\hat{X}(t-1|t-1) \quad (8c)$$

$$P(t|t-1)=F(t-1)P(t-1|t-1)[F(t-1)]^T + G(t-1)Q(t-1)[G(t-1)]^T \quad (8d)$$

Note:where

$H=[I_1,I_2,...,I_N]^T, I_i(i=1,2,...,N)$

is a unit matrix of $n\times n$ orders. And

$\overline{R}_i(t),(i=\overline{1,N}),\overline{R}(t)$ is the function of fusion prediction estimation value $\hat{X}(t|t-1)$, thus the fusion state estimation value is the non-linear function of $\hat{X}(t|t)$ and $\hat{X}(t|t-1)$ after transforming observation values. Hence this state estimation is called quasi-hierarchy fusion estimation.

The quasi-hierarchy fusion estimation based on the state estimation of N sensors is as following: for (7a),(7b)and (8a), it is obtained that

$$\hat{X}(t|t)=\hat{X}(t|t-1)+P(t|t)\sum_{i=1}^{N}\{P^{-1}(t|t)[X_i(t|t)-\hat{X}(t|t-1)]- -P_i^{-1}(t|t-1)[\hat{X}_i(t|t-1)-\hat{X}(t|t-1)]\} \quad (9a)$$

for (7b) and (8b), it is obtained that

$$P^{-1}(t|t)=P^{-1}(t|t-1)+\sum_{i=1}^{N}[P_i^{-1}(t|t)- -P_i^{-1}(t|t-1)] \quad (9b)$$

where

$$\hat{X}(t|t-1)=F(t-1)\hat{X}(t-1|t-1) \quad (9c)$$

$$P(t|t-1)=F(t-1)P(t-1|t-1)[F(t-1)]^T + G(t-1)Q(t-1)[G(t-1)]^T \quad (9d)$$

Thus, equations (9a)-(9b) are the quasi-hierarchy fusion estimation of the non-linear multisensor systems after transforming observation values.

4.2. Properties

(1) The quasi-hierarchy fusion estimation (9a)-(9b)is the same as of the standard linear multi-sensor systems[1] in construction. But they have the essense difference. The formar is the non-linear function of the state fusion predicaton $\hat{X}(t|t-1)$ and the hierarchical fusion estimation $\hat{X}(t|t)$ and it is a quasi-estimation. But what of the later is linear optimal estimaton[1].

(2) The fusion covariance matrix $P(t|t)$ can't be operated out off the computer for it is the function of fusion predication estimation $\hat{X}(t|t-1)$ and fusion estimation $\hat{X}(t|t)$ (see (7b)and (9b)).

(3) $P(t|t)$ and $P(t|t-1)$ only represent the linear model estimation accuracy after transforming observation values, while the fusion estimation accuracy is also dependent on the error from transforming observation values. That is when only the norm of $\overline{X}(t|t-1)=X(t)-\hat{X}(t|t-1)$ is small enough, quasi-hierarchy fusion estimation (9a)-(9d) has a very small error.

(4) Another form of (9a)-(9b)

$$\hat{X}(t|t)=\hat{X}(t|t-1)+P(t|t)\sum_{i=1}^{N}\{P^{-1}(t|t)[X_i(t|t)-\hat{X}(t|t-1)]- -[\overline{R}_i(t)]^{-1}[\hat{X}_i(t|t-1)-\hat{X}(t|t-1)]\} \quad (10a)$$

$$P^{-1}(t|t)=P^{-1}(t|t-1)+\sum_{i=1}^{N}[\overline{R}_i(t)]^{-1} \quad (10b)$$

The formula (10a) indicates that the quasi-hierarchy fusion estimation with transforming observation values equals the weighted sum of fusion prediction and the fusion track innovation The fusion track innovation is defined by the difference between two parts. First equals the sum that it is weighed by the sensor quasi-prediction covariance matrix that the difference between each local sensor quasi-filtering estimation and its quasi-prediction estimation, second equals the sum that it is weighed by the

sensor observation covariance matrix(after transforming observation values) that the difference between each local sensor quasi-filtering estimation and its quasi-fusion prediction estimation.

The formula (10b) indicates that it equals sum of the fusion prediction estimation covariance matrix inverse and each local sensor observation covariance matrix inverse (after transforming observation values) that the covariance matrix inverse of the quasi-hierarchy fusion estimation with transforming observation values.

When the fusion estimation formulas (10a)-(10b) and (9a)-(9b) are used, each local sensor filtering estimation only transmitted to the central processing agent at every time in the multisensor systems, while each local sensor prediction estimation, the fusion prediction estimation and their covariance matrix are operated by their state models in central processing agent. And each local sensor observation covariance matrix is once transmitted to the central processing agent and is computed as varied-time.

In a word, the fusion estimation (10a)-(10b),(9a)-(9d) further decreases transmission loads than (9a)-(9d), and each local sensor covariance matrix is not transmitted in the multisensor systems.

4.3. Realization Architecture

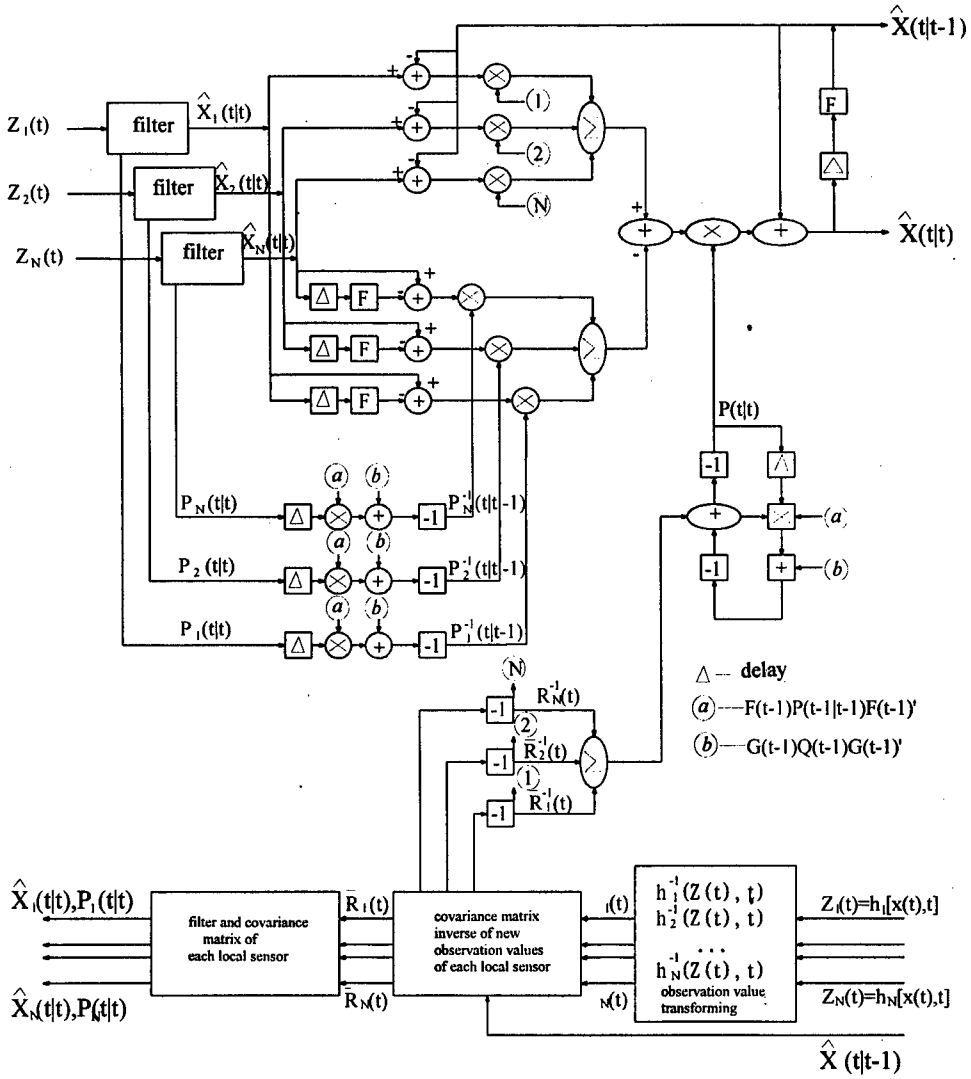


Fig.1 the quasi-hierarchy fusion estimation with transforming observation values

4.4. Application Example

In multi-radar tracking systems, given two dimensions observations, the radar observation position and the position rate of change are described $\rho_i(t)$ and $\dot{\rho}_i(t)$,respectively, the azimuth is described $\theta_i(t)$,$(i=\overline{1,N})$. The non-linear equation of observation systems is

$$Z_i(t)=\begin{pmatrix}\rho_i(t)\\ \theta_i(t)\\ \dot{\rho}_i(t)\end{pmatrix}=\begin{pmatrix}\sqrt{x^2(t)+y^2(t)}\\ arctg\dfrac{y(t)}{x(t)}\\ \dfrac{x(t)\dot{x}(t)+y(t)\dot{y}(t)}{\sqrt{x^2(t)+y^2(t)}}\end{pmatrix}+\begin{pmatrix}n_{\rho_i}(t)\\ n_{\theta_i}(t)\\ n_{\dot{\rho}_i}(t)\end{pmatrix}=h_i[X(t),t]+n_i(t) \quad (11)$$

where $n_{\theta_i}(t), n_{\rho_i}(t), n_{\dot{\rho}_i}(t)$ is described observation noises of azimuth, position and position rate of change, respectively. And they are independent of each other.

In formula (11), the target state vector is of four dimensions $((x(t), y(t), \dot{x}(t), \dot{y}(t))^T)$,but its observation value from each local sensor is of three dimensions $(\rho_i(t), \theta_i(t), \dot{\rho}_i(t))$,$(i=\overline{1,N})$, so $h_i[X(t),t]$ in the formula (11) is not the function of one-to-one. For solved the inverse function $h_i^{-1}[Z(t),t]$,a new method are used here, i.e. to increase one dimension of radar observation value in formula (11):

$$\dot{\theta}_i(t)=[\frac{\dot{y}(t)x(t)-\dot{x}(t)y(t)}{x^2(t)+y^2(t)}]_i+n_{\dot{\theta}_i}(t)$$

thus,

$$\bar{Z}_i(t)=\begin{pmatrix}\rho_i(t)\\ \theta_i(t)\\ \dot{\rho}_i(t)\\ \dot{\theta}_i(t)\end{pmatrix}=\begin{pmatrix}\sqrt{x^2(t)+y^2(t)}\\ arctg\dfrac{y(t)}{x(t)}\\ \dfrac{x(t)\dot{x}(t)+y(t)\dot{y}(t)}{\sqrt{x^2(t)+y^2(t)}}\\ \dfrac{x(t)\dot{y}(t)-\dot{y}(t)x(t)}{x^2(t)+y^2(t)}\end{pmatrix}+\begin{pmatrix}n_{\rho_i}(t)\\ n_{\theta_i}(t)\\ n_{\dot{\rho}_i}(t)\\ n_{\dot{\theta}_i}(t)\end{pmatrix}=\bar{h}_i[X(t),t]+\bar{n}_i(t) \quad (12)$$

where $n_{\dot{\theta}_i}(t)$ is the noise of $\dot{\theta}_i(t)$,and

(4.5.1)

$n_v(t), n_\theta(t), n_{\dot{\theta}}(t), n_{\dot{\rho}}(t)$ are independent of each other. Thus, the observation value transformation of the formula (12) is

$$x(t)=\rho(t)\cos\theta_i(t) \quad (13a)$$

$$y(t)=\rho(t)\sin\theta_i(t) \quad (13b)$$

$$\dot{x}(t)=\dot{\rho}(t)\cos\theta_i(t)-[\rho_i(t)\sin\theta_i(t)]\dot{\theta}_i(t) \quad (13c)$$

$$\dot{y}(t)=\dot{\rho}_i(t)\sin\theta_i(t)+[\rho_i(t)\cos\theta_i(t)]\dot{\theta}_i(t) \quad (13d)$$

so the new observation value is

$$\eta_i(t)=\begin{pmatrix}x(t)\\ \dot{x}(t)\\ y(t)\\ \dot{y}(t)\end{pmatrix}+\begin{pmatrix}\bar{V}_{ix}(t)\\ \bar{V}_{i\dot{x}}(t)\\ \bar{V}_{iy}(t)\\ \bar{V}_{i\dot{y}}(t)\end{pmatrix}=X(t)+\bar{V}_i(t) \quad (14)$$

where $\bar{V}_i(t)$ is the observation noise after transforming observation values.

Known from the formulas (13a)-(13d), the observation error in Cartesian coordinate systems is of non-Gaussian distribution. And tracking filering estimation is also non-linear. So the observation error needs to change to Gaussian's.

Let the observation error vector $(\Delta\rho_i(t), \Delta\theta_i(t), \Delta\dot{\rho}_i(t), \Delta\dot{\theta}_i(t))^T$ in polar-coordinate systems is much smaller than the target truth value $(\rho_i(t), \theta_i(t), \dot{\rho}_i(t), \dot{\theta}_i(t))^T$,

then derivation of formulas (13a)-(13d) with respect to time t is

$$\begin{pmatrix}\Delta[x(t)]\\ \Delta[\dot{x}(t)]\\ \Delta[y(t)]\\ \Delta[\dot{y}(t)]\end{pmatrix}=$$

$$\begin{pmatrix} & 0 & -\rho_i(t)\sin\theta_i(t) & 0\\ -\dot{\theta}_i(t)\sin\theta_i(t) & \cos\theta_i(t) & -\dot{\rho}_i(t)\theta_i(t)-\rho_i(t)\dot{\theta}_i(t)\cos\theta_i(t) & -\rho_i(t)\sin\theta_i(t)\\ \sin\theta_i(t) & 0 & \rho_i(t)\cos\theta_i(t) & 0\\ \dot{\theta}_i(t)\cos\theta_i(t) & \sin\theta_i(t) & \dot{\rho}_i(t)\cos\theta_i(t)-\rho_i(t)\dot{\theta}_i(t)\sin\theta_i(t) & \rho_i(t)\cos\theta_i(t)\end{pmatrix}$$

$$\times\begin{pmatrix}\Delta[\rho_i(t)]\\ \Delta[\dot{\rho}_i(t)]\\ \Delta[\theta_i(t)]\\ \Delta[\dot{\theta}_i(t)]\end{pmatrix}$$

where $(\Delta[x(t)],\Delta\dot{x}(t),\Delta[y(t)],,\Delta[\dot{y}(t)])^T$ has the linear relationship with $(\Delta\rho_i(t),\Delta\theta_i(t),\Delta\dot{\rho}_i(t),\Delta\dot{\theta}_i(t))^T$ .Therefor, $(\Delta[x(t)],\Delta\dot{x}(t),\Delta[y(t)],\Delta[\dot{y}(t)])^T$ is subject to Gaussian distribution, and its covariance matrix is

$$\overline{R}_i(t)=\begin{pmatrix}[\sigma_x^2(t)] & [\sigma_{x\dot{x}}^2(t)] & [\sigma_{xy}^2(t)] & [\sigma_{x\dot{y}}^2(t)]\\ [\sigma_{\dot{x}x}^2(t)] & [\sigma_{\dot{x}}^2(t)] & [\sigma_{\dot{x}y}^2(t)] & [\sigma_{\dot{x}\dot{y}}^2(t)]\\ [\sigma_{xy}^2(t)] & [\sigma_{\dot{x}y}^2(t)] & [\sigma_{y}^2(t)] & [\sigma_{y\dot{y}}^2(t)]\\ [\sigma_{x\dot{y}}^2(t)] & [\sigma_{\dot{x}\dot{y}}^2(t)] & [\sigma_{y\dot{y}}^2(t)] & [\sigma_{\dot{y}}^2(t)]\end{pmatrix}(15)$$

where

$$[\sigma_x^2(t)]_i=\sigma_{\rho_i}^2(t)\cos^2\theta_i(t)+\rho_i^2(t)\sigma_{\theta_i}^2(t)\sin^2\theta_i(t)$$

$$[\sigma_y^2(t)]_i=\sigma_{\rho_i}^2(t)\sin^2\theta_i(t)+\rho_i^2(t)\sigma_{\theta_i}^2(t)\cos^2\theta_i(t)$$

$$[\sigma_{\dot{x}}^2(t)]_i=\sigma_{\rho_i}^2(t)[\dot{\theta}_i(t)]^2\sin^2\theta_i(t)+[\dot{\rho}_i(t)\sin^2\theta_i(t)+\rho_i(t)\dot{\theta}_i(t)\cos\theta_i(t)]^2\sigma_{\theta_i}^2(t)+\cos^2(\theta_i(t))\sigma_{\dot{\rho}_i}^2(t)+\rho_i^2(t)\sigma_{\dot{\theta}_i}^2(t)\sin^2\theta_i(t)$$

$$[\sigma_{\dot{y}}^2(t)]_i=\sigma_{\rho_i}^2(t)[\dot{\theta}_i(t)]^2\cos^2\theta_i(t)+[\dot{\rho}_i(t)\cos\theta_i(t)-\rho_i(t)\dot{\theta}_i(t)\sin\theta_i(t)]^2\sigma_{\theta_i}^2(t)+\sigma_{\rho_i}^2(t)\sin^2\theta_i(t)+\sigma_{\theta_i}(t)\rho^2{}_i(t)\cos^2\theta_i(t)$$

$$[\sigma_{xy}(t)]_i=[\sigma_{\rho_i}^2(t)-\sigma_{\theta_i}^2(t)\rho^2{}_i(t)]\sin\theta_i(t)\cos\theta_i(t)$$

$$[\sigma_{\dot{x}y}(t)]_i=\sigma_{\dot{\rho}_i}^2(t)\sin\theta_i(t)\cos\theta_i(t)-[\dot{\rho}_i(t)\sin\theta_i(t)+\rho_i(t)\dot{\theta}_i(t)\cos\theta_i(t)]*[\dot{\rho}_i(t)\cos\theta_i(t)-\rho_i(t)\dot{\theta}_i(t)\sin\theta_i(t)]\sigma_{0_i}^2(t)-\sigma_{\rho_i}^2(t)\dot{\theta}_i{}^2(t)\sin\theta_i(t)\cos\theta_i(t)-\sigma_{\theta_i}^2(t)\rho^2{}_i(t)\sin\theta_i(t)\cos\theta_i(t)$$

$$[\sigma_{\dot{x}\dot{x}}^2(t)]_i=\sigma_{\rho_i}^2(t)\dot{\theta}_i(t)\cos\theta_i(t)\sin\theta_i(t)+\rho_i(t)\sigma_{\theta_i}^2(t)\sin^2\theta_i(t)[\dot{\rho}_i(t)\sin\theta_i(t)+\dot{\theta}_i(t)\rho_i(t)\cos\theta_i(t)$$

$$[\sigma_{\dot{x}y}^2(t)]_i=-\sigma_{\rho_i}^2(t)\dot{\theta}_i(t)\sin^2\theta_i(t)+\sigma_{\theta_i}^2(t)[\dot{\rho}_i(t)\sin\theta_i(t)+\dot{\theta}_i(t)\rho_i(t)\cos\theta_i(t)]$$

$$[\sigma_{\dot{x}\dot{y}}^2(t)]_i=-\sigma_{\rho_i}^2(t)\dot{\theta}_i(t)\sin^2\theta_i(t)+\sigma_{\theta_i}^2(t)[\dot{\rho}_i(t)\sin\theta_i(t)+\dot{\theta}_i(t)\rho_i(t)\cos\theta_i(t)]$$

$$[\sigma_{y\dot{y}}^2(t)]_i=\sigma_{\rho_i}^2(t)\dot{\theta}_i(t)\sin\theta_i(t)\cos\theta_i(t)-\rho_i(t)\sigma_{\theta_i}^2(t)\cos\theta_i(t)[\dot{\rho}_i(t)\cos\theta_i(t)-\dot{\theta}_i(t)\rho_i(t)\sin\theta_i(t)]$$

and $\overline{R}_i(t)$ is computed by

$$\rho_i(t)=\hat{\rho}_i(t\mid t-1)$$
$$\dot{\rho}_i(t)=\hat{\dot{\rho}}_i(t\mid t-1)$$
$$\theta_i(t)=\hat{\theta}_i(t\mid t-1)$$
$$\dot{\theta}_i(t)=\hat{\dot{\theta}}_i(t\mid t-1)$$

Hence, formulas (1)and (14) form the linear model of N radar, systems, and its observation covariance matrix is $\overline{R}_i(t)$ .And each local radar state filtering estimation is given by formulas (7a)-(7d), and it is transmitted to the central processing agent, while each local and fusion prediction estimation as well as their covariance matrices are operated by formulas (7c),(7d),(9c),(9d) in the central processing agent. Finally, the state quasi-optimal estimation is obtained by formulas (10a)-(10b).

5. Summary

An algorithm for quasi-hierarchical fusion estimation with transforming observation values is proposed in this paper. It is used to solve the state estimation of non-linear multisensor systems that the state model is of linear, observation model is of non-linear and this non-linear function has an inverse function. It is obtained that:

The quasi-hierarchical fusion equations of this

systems are given, and their properties are analyzed. Farther, the quasi-hierarchical fusion equations with better properties are obtained and above algorithm realization architecture is presented.

(1) The feasibility of this algorithm is shown by an example of multi-radar tracking systems. And a new preprocessing method is used in the target state quasi-hierarchical fusion estimation with transforming observation values in multi-radar tracking systems, i.e. it is made use of to increase one dimension (angular velocity) in observation data and reestablished observation model, then the quasi-hierarchical fusion estimation with transforming observation values is made. This preprocessing method decreases the load of general processing method that changes the position and azimuth under polar-coordinate into the component of velocity of the target state with two sensor observation values or a sensor prediction value.

In short, an algorithm for quasi-hierarchical fusion estimation with transforming observation values is feasibility for the target state estimation of non-linear multisensor systems (1)-(2).

6. References

[1] Yaakov Bar-Shalom, Editor, Multitarget-multisensor training: advanced applications, Artech House,1989.

[2] Hongyan Sun, Wenlong Hu, Pingxing Lin and Shiyi Mao, A Study on an algorithm of

[3] multisensor data fusion, NAECON'94, Dayton, Ohio.239-245, May,94.

[4] Edward Waltz and James Llinas, Multisensor Data Fusion, Artech House,1990.

[5] A.Farina and F.A.Studer, Radar Data Processing (Volume I), Research studies LTD,1985.

Session RA1
Image Fusion III
Chair: Robert S. Lynch
Naval Undersea Warfare Center, RI, USA

Visible/IR Battle Field Image Registration using Local Hausdorff Distance

Yunlong Sheng, Xiangjie Yang, Daniel McReynolds, Dept. of Physics, Laval University, Ste-Foy, Quebec, Canada G1K 7P4,
Piere Valin, Lockheed Martin,6111 Royalmount Ave. Montreal Qc Canada H4P 1K6
Leandre Sevigny, Defence Research Establishment Valcartier, 2458 Boul. Pie XI Nord, C.P. 8800, Courcette Qc. Canada G0A 1R0

Abstract: Feature inconsistency and low contract and noise in the infrared image background consist of the principle difficulty in the IR/Visible battle field image registration. Feature-based approaches are more powerful and versatile to process poor quality IR images. Multi-scale hierarchical edge detection and edge focusing and salience measure are used in the feature horizon extraction. The common features extracted from images of two modalities can be still different in detail. Therefore, the transformation space match methods with the Hausdorff distance measure is more suitable than the direct feature matching methods. We have introduced image quadtree partition technique to the Hausdorff distance matching, that dramatically reduces the size of the search space. Image registration of real world visible/IR images of battle fields is shown.

1. Introduction

Multiple imaging sensors have different electromagnetic spectral responses to capture distinguished signatures from the input scene in different spectral bands. The design of the multisensor imaging system maximizes the independence of the acquired data. This is natural, since if one sensor captures images that are similar or correlated to the images already obtained by other sensors, then this sensor provides no additional information and should be removed from the system. In principle, the images from multiple sensors should be uncorrelated and independent from each other, which implies that the features in the multisensor images are inconsistent. Some features in one image can donnot show up in another image.

Multiple sensors can act in a synergistic manner. The images to be registered in our research project are two broad band visible and infrared video sequences, as shown in Fig.1. When soldiers and a truck are hidden behind the smoke in the visible image they appear clairly as high contrast hot objects in the IR images.

In this paper we present techniques for IR/visible battle field image registration. and we implement the feature based approach. We use multi-scale hierarchical edge detection and edge focusing and the edge salience measure to extract salient edges from the low contrast and noisy IR image background. We use the Hausdorff distance measure for matching between the curves from two different modalities. We introduce the image partitioning technique in the Hausdorff distance matching, so that the affine transformation is approximated by local translations. This speeds up significantly the Hausdorff distance matching process.

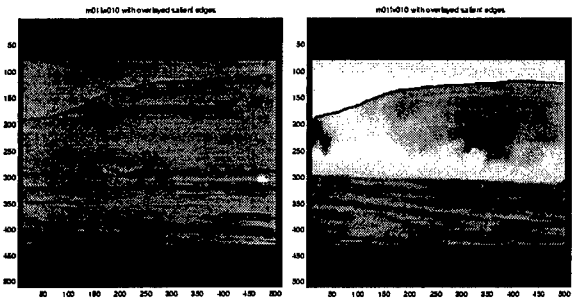


Fig.1 IR and Visible battle field images. With some edge features extracted for registration

2. Feature inconsistency

The radiometric data from IR passive sensors consist of 1) energy emitted by thermal radiation from the object bodies; 2) atmospheric emission reflected from object surfaces. In general, the gray-scale level of IR images depend on differences in body temperature, emissivity and reflectivity of the objects in the scene. The IR images of the battle field have high contrast for hot objects in the scene, which are in most cases moving objects and targets and cannot be used as landmarks for registration. Image registration should rely on the stationery objects on the background of the scene, where, unfortunately, the IR outdoor images have very low contrast, owing to the uniform temperature field on the background in the thermal equilibrium state. The background in the outdoor IR images is usually of very low contrast and noisy, or simply a dark region, that makes image feature extraction and registration more difficult.

There exist significant gray-level disparities between the IR and visible image. The thermal emitters are not necessarily good visual reflectors. A surface of high visual reflectivity (white surface) in visible band usually has low emissivity, so that the bright objects in the visible image may be dark in the thermal scene and vice versa. The sky is usually the brightest region in the visible image. It is, however, a dark region in the IR image because of the low temperature and the

lack of reflectance. This is the reversal of contrast polarity between the visible and IR images.

The gray level disparity between the IR and visible images of real-world natural out-door scene is much more complex than the simple contrast polarity reversal.

Fig. 2 shows a contrast reversed IR image compared with the visible image of the same scene. The gray level distributions in most regions are similar in the contrast reversed IR and visible images, although there are still important gray-level disparities. However, in the contrast polarity reversed IR images, the clouds are darker than the sky, whereas they are brighter than the sky in the visible images because of its higher reflectivity. The clouds are also brighter than the sky in the original IR image because of its higher reflectivity and emissivity. Hence, a simple reversal of contrast polarity can not remove all the gray level disparities. Also, shadows in the visible images are absent in the IR images.

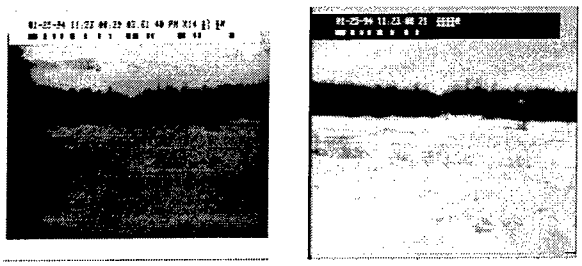


Fig.2 Contrast reversed IR image (left) and visible image (right)

3 Area and feature-based approaches

Area-based approach for image registration utilizes full image information, can be applied to any images with rich or poor structure and the cross-correlation based matched filter approach is optimal for the robustness against random noise. The area-based image registration using cross-correlation can account only for image translations.

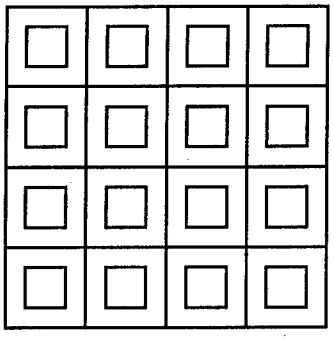


Fig.3 Block matching with partitioning of images

To fit more general image transformations such as affine transformation, one can partition images into a number of sub-images, which are located in a regular grid as shown in Fig.3., and then define a central window in each sub-image as a template and correlate those blocks with the corresponding sub-images in another image. The block matching results in a number of displacement vectors, which are evenly distributed over the image and are then useful to determine the transformation parameters for image registration. In the block matching the sub-images are only translated in a small neighborhood of the respective grid, to approximate more complicated distortions[i].

3.1 Laplacian pyramid transform

The area-based approach requires the radiometric data of two images to be similar. If these distributions are different, then area-based match will fail. To use the area-based method for multisensor image registration one has to transform the two dissimilar images into similar. The intensities of the Laplacian pyramid images are insensitive to gray-scale level disparities and polarity reversals of contrast.

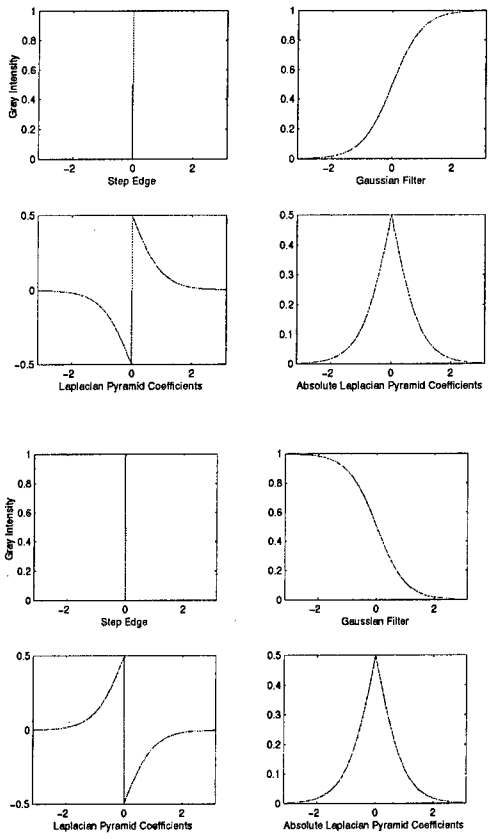


Fig.4 In the clock-wise order from the left-top corner: Step edges, smoothed by Gaussian filter, Laplacian pyramid and absolute coefficients of the Laplacian pyramid.

Figure 4 shows two step edges with opposite polarities of contrast. The edges are smoothed by the Gaussian filter. The Laplacian pyramid images are the differences between the original edge and the smoothed edge. When we take the absolute values of the Laplacian pyramid coefficients, the two Laplacian pyramid images become the same for the two contrast reversed step edges. Then, the contrast reversal is removed and the area-based image registration can be applied to the Laplacian pyramid image intensities[ii].

3.2 Phase matching

Images from different sensors have different radiometric intensity distributions due to the different spectral responses of the sensors. Those differences appear mostly as slow variations over wide regions in the image, such as sky, land and forest, which are usually represented with low spatial frequencies and are concentrated in a narrow low frequency band.

In the Fourier transform-based registration[iii], the displacement is found by cross-correlation between two images. The location of the cross-correlation peak mainly depends on the Fourier spectrum phase and is insensitive to Fourier spectrum energy. One can then whiten the Fourier spectrum and use the phase-only cross-correlation for the registration. In this approach, the low and high frequencies contribute equally to the cross-correlation. Therefore, contribution of the high frequencies is greatly highlighted, compared with the conventional cross-correlation. The location of the cross-correlation peak would not change if the image intensity variations are limited to a narrow spatial frequency band. The Fourier phase correlation registration method is then relatively independent of the sensors.

3.3 Feature-based matching

The IR/visible image grayscale level disparities can not be removed completely by a reversal of contrast polarity, as shown in Fig.2, and by the Laplacian pyramid representation or by whitening the Fourier spectrum. The residual disparities in some blocks can lead to erroneous displacement vectors, which will be the outliers for fitting to image transformation. The robust image fitting techniques against outliers, such as the least median of squares or M-estimation must be used.

We notice that both the Laplacian pyramid representation and the phase matching technique in the area-based matching benefit from the use of high spatial frequencies of the image for overcoming the feature inconsistency. The Laplacian pyramid therefore represents detailed information, namely contours, in the image. In the phase matching approach, the whitening of the Fourier spectrum highlights the high spatial frequencies. The inverse Fourier transform of the whitened spectrum is an edge and contour enhanced image. In some sense both the Laplacian pyramid representation and the phase matching share the same features with the edge matching approaches for image registration. However, the edge extraction by the Laplacian pyramid and by the whitening Fourier spectrum is not as powerful and precise as that implemented by using directly the edge detectors. Especially, for our real-world IR images, which are typically of low contrast and noisy, the Laplacian pyramid approach is not able to extract salient contours for image matching, so that feature-based image registration is adopted.

The feature-based approach requires extracting the common features from two images. If both types of images represent the same real world objects, then the objects should appear in both image types. While images may appear differently in different sensors, different objects appear always differently in the multisensor images, no matter what are the spectral responses of the imaging sensors. As a result, the boundaries between objects would be preserved, that can be used for image registration. Although we find that the edges extracted from the same real world objects in two image modalities still can have different details due to the differences in the radiometric responses of two sensors.

Advantages of the feature-based image registration are that the common image edge features are not sensitive to the spectral responses of the multiple sensors; the processing speed is independent of image displacement; any image transformations can be accounted for and the powerful and versatile edge detection, edge saliency techniques can be used.

4. Multi-scale edge detection

In the 3-D real world scene, objects are separated from the background by depth discontinuities, which are usually manifest as intensity discontinuities in the 2-D images. Those edges and boundaries represent structures in the image, that are common for multiple image types and can be used for multiple sensor image registration. Edges are defined as points where the modulus of gradient is a maximum in the gradient direction. Along an edge the image intensity can be singular in one direction while varying smoothly in the perpendicular direction. Edges can be created by occlusions, shadows, sharp changes of surface orientation, changes in reflectance properties, or illumination. In IR images of a 3-D scene, most edges represent occlusions and depth discontinuities between objects in the scene, which represent structural information in the image.

4.1 IR image edge detection

A particular difficulty arises in the edge detection for IR/visible image registration. Image registration requires to extract common features which are static in the scene background. In most cases, the

background objects in the IR images have the same thermal equilibrium temperature, so that the contrast in the IR image background is related to only the differences in the emissivities and reflectivities of the object surfaces and are therefore very low. Also, the IR images are typically noisy.

Optimal filter for step edges detection can be approximated by the first derivative of Gaussian, which is usually called Canny edge detector[iv]. After the filtering, there is a non-maximum suppression process that keeps only the pixels where the values of the output are the local maximum in the direction of the gradient. The values at the neighboring pixels are determined by the linear interpolation. The third process in the Canny detector is the edge linking, which uses a hysteresis thresholding. We first determine edge pixels, which are above a high threshold. Then, among all other local maxima, which are above a low threshold, we keep only those pixels that are located in the neighborhood of the edge pixels.

The parameters in the Canny edge detector are the width of first derivative of Gaussian filter σ and the low and high threshold values. One problem of the Canny edge detector is its sensitivity to threshold. When the response of an edge point is close to the detection threshold, a small change in edge strength or in the pixellation may cause a large change in edge topology, that makes the extracted edges suspicious, non-reliable, especially near the corners.

The sensitivity to noise is another important problem in the edge detection. The noise in IR images occur as local fluctuations of the image brightness function, which have strong derivative magnitudes, but represent unnecessary image details which are unrelated to image structure. In the IR image background of low contrast with the contrast varying cross the image, the effect of noise becomes important, so that the structural edges may be disrupted and even completely disappear in the edge maps, if a thresholding on the gradient magnitude is applied. The non-maximum suppression in the Canny detector is excessively reliant on the estimation of the gradient angle and so often fails to mark edge pixels at junctions, corners and even on some smooth curve portions where the contrast changes are too poorly defined[v]. This is the reason for broken edges.

For detecting structural edges in the IR image background, we use the Canny edge detector without thresholding on the gradient magnitude. We avoid the use of threshold on the gradient magnitude, since the contrast is a poor indicator for significance. When the strength threshold is used, of the edges with response close to the threshold, a small change in edge strength or location can cause a large change in the edge topology. We use the large Canny filter of $\sigma \geq$ 6-7, which corresponds to a filter size of 37 - 43 pixels, to obtain the structural edge as a continuous curve, which is the horizon in the scene of battle field, so that the curve length thresholding can be applied to extract the horizon from noisy edges in the edge map. With a small σ, the extracted horizon line is broken. However, with a large σ, the extracted horizon line does not follow the real contour at high curvature. The larger the filter support σ, the less broken the edges are, and, however, the more image details are filtered out by the large size filter, resulting in a loss of edge localization. Therefore, the multi-scale edge detection is used to recover the localization in the coarse edges.

4.2 Hierarchical Edge Detection

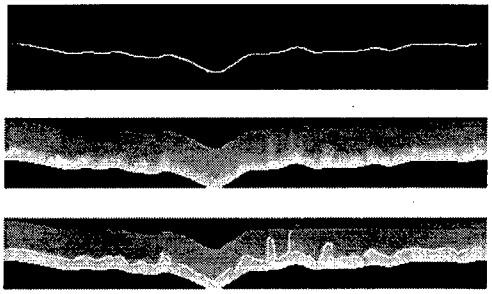
First, the horizon curve is detected at a coarse level with a large Canny edge detector which smoothes the images with a Gaussian of large support σ_0. The horizon is usually the longest curve in the image. For favoring continuity of the extracted curve, no thresholding on the gradient magnitude is applied, such that the horizon appears as a continuous curve or, at least, less broken. Then, the horizon is extracted from the noisy edge map by a curve length thresholding. In the cases where the horizon curves are still broken, we apply the edge saliency measure and combine both edge and region information in order to ensure the extraction of the horizon at the coarsest level, as explained in Section 6.

The coarse horizon is used to guide the search of edges at fine scale. We define a sub-image in the neighborhood of the coarse edge in the original image. The sub-image covers the region along the horizon with 40 pixels above and 10 pixels below each coarse horizon point. The choice of the sub-image size is according to the observation that the images of trees on the hill were cut by the smoothing at the coarse scale. To recover the top of trees we need a search in a large region above the horizon curve. We then apply the Canny edge detector with a small filter width σ within the sub-image. In the experiment, the fine Canny filter was with $\sigma = 0.7$ for visible and $\sigma = 1.5$ for IR images. The noise still exists after the Canny edge detection at the fine scale. However, this noise is within the sub-image zone and may be removed easily by a curve length thresholding, that results in a clearly defined horizon curve. A specific modification on the Canny edge detector was made to prevent the artificially defined sub-image boundaries from appearing as new edges.

The coarse horizon extracted from an IR image is shown in Fig.5a, where $\sigma_0 = 7.0$, the minimum length threshold applied was 500 pixels. The sub-image is shown in Fig.5b. Figure 5c shows the fine edges obtained by applying a fine Canny edge detector with $\sigma = 1.5$.

The hierarchical edge detection is quit reliable and fast. Since at the fine scale the edge detection is guided by the coarse level edge, the search in large

area is avoided, that reduces the computational cost. The shortcoming of the algorithm is the ad-hoc determination of sub-images.



(a)

(b)

(c)

Fig.5 Results of Hierarchical Edge Detection

4.3 Edge Focusing

Edge focusing is a coarse-to-fine edge tracking algorithm for recovering the edge points at the finest scale. The scale-space tracking is implemented in a continuous manner. With continuous scaling, the edges are gradually focused by varying the resolution continuously, and moving in the scale space with sufficiently small steps, such that the edge element do not jump farther away than one pixel between successive steps. Our implementation of edge focusing is as following:

1. Detect edge using Canny Detector with the Gaussian smoothing σ_0 sufficiently large so that horizon curve is detected;
2. Extract the horizon using a threshold on the curve length; The horizon curve is denoted as $E(i,j,\sigma_0)$. If (i,j) is an edge point, then $E(i,j,\sigma)=1$.
3. Detect edges $E(i,j,\sigma_k)$ in a window centered at each edge point $E(i,j,\sigma_{k-1})$, using the Canny edge detector of size $\sigma_k=\sigma_{k-1}-\Delta\sigma$ with $k=1,2,3...$ and $\Delta\sigma=0.5$. The window size is 7×7, when $\sigma_k>2.0$, and is 5×5 when $1.0\le\sigma_k\le2.0$, and is 3×3 when $\sigma_k<1.0$.
4. Go on step 3) until a weak Gaussian smoothing of size σ_K.

In the successive Canny edge detection, after application of the first derivative of Gaussian filter the non-maximum suppression process is applied which keeps only the local maximum in the gradient direction. There is no threshold at finer resolution. The only threshold is on the curve length applied at the coarsest scale σ_0.

Bergholm[vi] investigated the deformation of four elementary contour structures: step edge, corner, double edges and edge box. During the edge detection, those contours are generally deformed in four ways: rounding-off, expansion, transformation into circles, or merger, owing to the large Gaussian average operator which blurs the image. In each of the four cases, Bergholm showed that the displacement vector, describing the deformation of the edge contour, is normally of length within the range from 0 to $2|\Delta\sigma|$, where σ is the width of the Canny edge detector, $\Delta\sigma$ is the increment of size of the successive Canny filters. Therefore, if $|\Delta\sigma|$ = 0.5, the displacement of the edge points would be normally less than one pixels, so that corners and junctions may be recovered with a precision less than one pixel.

In real world images we have mostly ramp edges instead of ideal step edges. It is easy to show that the Gaussian blurring operating on a ramp edge always yields smaller displacement than that yielded on a step edge as affirmed by Bergholm. A ramp edge may be modeled as a step edge smoothed by a Gaussian G whose size σ_1 depends on the imaging condition and on the camera. Let $r(x,y)$ denote the step edge and $f(x,y)$ the ramp edge in gray level image, then

$$f(x,y)=r(x,y)\otimes G(\sigma_1)$$

where $\otimes$ denotes the convolution. When we use Canny Edge Detector, the image is blurred again with a Gaussian smoothing whose size σ_2 depends on the scale of the edge detector. Let $g(x,y)$ denote the blurred ramp edge before computing the first derivative, then

$$g(x,y)=f(x,y)\otimes G(\sigma_2)$$

therefore

$$g(x,y)=r(x,y)\otimes G(\sigma_1)\otimes G(\sigma_2)=$$
$$r(x,y)\otimes(G(\sigma_1)\otimes G(\sigma_2))$$

which is equal to

$$g(x,y)=r(x,y)\otimes G(\sqrt{\sigma_1^2+\sigma_2^2})$$

We implement the edge focusing algorithm with the filter size increment $\Delta\sigma=0.5$ and varying size windows. We chose to use the window size larger than the usually used, 3 x 3, so that the gradient magnitude values can be evaluated at the two neighboring pixels, because in the non-maximum suppression the determination of an edge pixel requires to compare with at least two neighboring pixels. We believe that the length of rounding-off displacement ρ can be larger than one pixel, because the real ramp edges in our IR images were noisy and do not follow the theoretical model described in the precedent.

Therefore, the length of rounding-off displacement ρ from the corner of ideal step edges to the detected corner is equal to $c\sqrt{\sigma_1^2+\sigma_2^2}$, where c is a constant. However, the displacement from the center of the

ramp corner to the detected corner would be

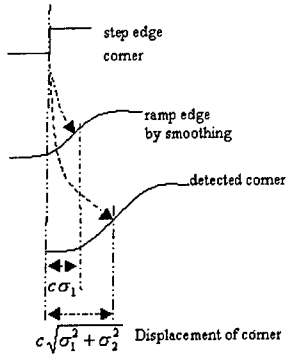
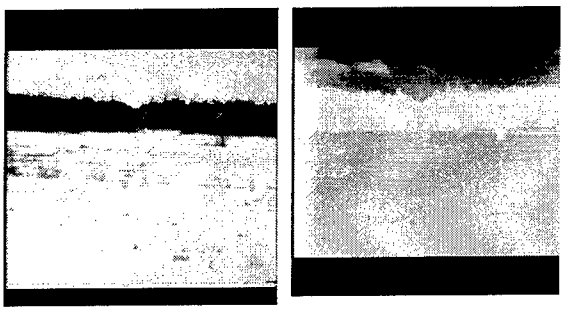


Fig.6 Rounding-off displacement for a ramp edge

proportional to $\sqrt{\sigma_1^2+\sigma_2^2}$ - σ_1, as illustrated in Fig. 6 and would be less than σ_2. Therefore, if $|\Delta\sigma_2|$ = 0.5 in the edge focusing, the displacement of the ramp edge corner would be less than one pixels.

In our IR images the ramp edges of trees can be very slow of more than 20 pixels wide, corresponding to a large σ_1 more than 10. The edges around the trees were cut completely when a Canny edge detector of



(a) (b)

Fig.7. Experiment results of Edge Focusing. (a) Visible image. (b) Infrared image.

σ_2 = 7 was applied. This is because the large displacement of the corner $\sqrt{\sigma_1^2+\sigma_2^2}$. However, using the edge focusing we were able to recover the edges and tops of the trees, which would be important for the image registration.

For images shown in Fig.7, we first detected the coarse horizon with $\sigma_0 = 4.5$ for visible image and $\sigma_0 = 7.0$ for IR image using Canny Edge Detector. Then we applied the edge focusing with the scale step $\Delta\sigma = 0.5$ and the varying size windows. The final scale was $\sigma = 0.7$ for visible image and $\sigma = 1.5$ for IR image. Figure 7 shows the extracted edges which follow nicely the silhouette of the hill with some flat tops of trees recovered in both visible image and infrared image.

5. Curve saliency

The Edge detector is basically a local operator. However, the structural edges useful for image registration are not local features, but exhibit regional and global nature in many cases. Salient structures can often be perceived in an image at a glance[vii]. They appear to attract our attention. Therefore, we use curve saliency measure to help detecting the structural edges.

When cameras are mounted on a grounded vehicle, it is reasonable to assume that the camera axis is pointing approximately horizontally. Unless the terrain is very steep (or the vehicle is driving alongside a wall) the horizon is usually visible in the image and is usually the most distant part of the scene (except for the sky). The horizon lines are common in the IR and visible images, and are independent of the grayscale level disparities and contrast polarity reversals. More importantly, it may be possible to segment the horizon line from noisy edges on the ground by a threshold of curve lengths, since the horizon line has the longest length in the image. The fact that the horizon is the most distant part in the image helps fitting the distortion transformation for image registration.

Salience measures can be region-based or curve-based. In the visible images the horizon bounds the brightest part of the image, which is usually always the sky. Duric and Rosenfeld[viii] use the horizon detection for stabilization of image sequence from a ground bright parts of the image (sky) and then estimating the boundaries of these parts. This approach uses then the regional information. We attempt to use the curve-based salience measure to detect horizon lines in IR and visible images. The curve saliency measure is defined to favour long over short curves and smooth over wiggly curves. For horizon, we define the salience measure which is estimated at each pixel along a curve *i*, as

$$\Phi_i = \frac{L_i}{N}\sum_{k=1}^{N}\sigma_k\frac{|y_{k+1}-y_k|}{|x_{k+1}-x_k|}+\alpha_k$$

where N is the total number of pixels on a segment of the curve *i*, whose horizontal extension in x is L_i

$$\alpha_k = \begin{cases} 1 & if \quad y_{k+1}-y_k = 0 \\ 2 & if \quad y_{k+1}-y_k \neq 0 \end{cases}$$

$$\sigma_k = \begin{cases} 0 & if \quad x_{k+1}-x_k = 0 \\ 1 & if \quad x_{k+1}-x_k \neq 0 \end{cases}$$

The horizon in the natural scene usually is not a horizontally straight line. This curve saliency measure favorites inclined segments rather than horizontal or vertical ones. The horizontal segments have contribution of 1 to the saliency Φ_i, vertical segments have saliency of 2, since σ_k = 0. Inclined

segments have saliency of 3. However, wiggly curves will receive small total saliency measure because of small L_i. We evaluate the salience measure for each curve in the edge map and retain 3 - 4 most salient curves. Then, we fill gaps between the salient curves and re-evaluate the saliency of the connected curves. With this approach we can detect the horizon at a coarse scale without thresholding of curve length, so that the horizon is detected even it is broken by noise into several segments and contains a number of gaps.

5. Hausdorff distance

The purpose of image alignment is to register a pair of images such that the extracted static scene features are optimally aligned. Feature based image registration requires a specification for the features, the parameter space, the image transformation, which aligns the image, and the search strategy for finding the best alignment according to some objective function. For aligning images from different modalities, edges arising from depth discontinuities can be considered as most salient. Given a set of salient edges from each image, the next step is to determine the image transformation which aligns those features considered to be a static reference for the scene. The search for the optimal image transformation can be implemented in several ways. The methods can be classified as feature matching methods or transformation space methods.

Feature matching methods determine the correspondence between the elements of the feature sets, i.e., corresponding features are projectively given by the same scene feature. The transformation space methods search the parameter space for the solution that achieves an optimal alignment of the static projected scene features. The drawback of feature matching methods is the prohibitive cost of detecting and eliminating outliers, i.e., features which do not have a match. Its advantage is that once a set of correct matches is found the image transformation is, in general, quick to compute. Transformation space methods can be prohibitively expensive because the search space is generally very large, however, outliers are easily handled by using rank order statistics. A strategy for efficiently searching the parameter space is given by Huttenlocher *et al.*[ix] In view of the large proportion of outliers in feature based multi-modal image alignment a transformation space method based on the directed Hausdorff distance was implemented. The size of the search space is reduced by partitioning the image into blocks and searching for translations that minimize the Hausdorff distance between corresponding blocks. The assumptions are that the motion can be locally approximated by simple translations of blocks, and the percentage of outliers and an error bound for the feature alignment are known approximately.

The Hausdorff distance is defined by

$$A = \{a_1, \ldots, a_m\} \text{ and } B = \{b_1, \ldots, b_n\}$$

$$H(A,B) = \max(h(A,B), h(B,A))$$

$$h(A,B) = \max_{a \in A} \min_{b \in b} \|a - b\|$$

where A and B are point sets, H is the generalized Hausdorff distance and h is the directed Hausdorff distance. In the presence of outliers the Hausdorff distance will return the greatest distance which is likely due to an outlier. To be able to compare portions of the data sets the partial directed Hausdorff distance is defined,

$$h_k(A,B) = \text{kth}_{a \in A} \min_{b \in B} \|a - b\|.$$

This expression evaluates to the k^{th} ranked distance. The alignment method using Hausdorff distances proceeds as follows for a pair of images after extraction of the salient edges

1) Compute a quadtree partition of each edge image such that no block without edge points is further subdivided. The partition with fewer blocks is retained for both images. Define a set of model edge points for the first block in image 1 from the edge points that lie within that block. Create a model image from these edge points.
2) Define a set of subimage edge points from the corresponding block in image 2 from the edge points within the block extended by a border whose dimensions correspond to the largest expected vertical and horizontal displacements. Create a target image from these edge points.
3) Compute the directed partial Hausdorff distance under a translation transformation from the model image to the target image. The translation which minimizes the kth ranked distance is retained. The search strategy in the translation parameter space is described in Huttenlocher et al.12
4) Repeat steps 2 and 3 for the remaining non-empty blocks. If at least 3 blocks provide local translation estimates from step 3 then the global affine transformation is estimated, the nonreference image is resampled according the global affine transformation and the images are fused. The image fusion is accomplished by an appropriately weighted combination of the aligned images brightness values.

Fig. 1 shows a scene taken simultaneously by a daylight and IR camera at Defense Research Establishment Valcartier. The viewpoints of the two cameras are displaced slightly and there is a slight relative rotation about the optical axis which would yield a very poor fused image if no alignment is made. The quadtree decomposition stops at the first level, i.e., there are 4 blocks. The salient edges

include the silhouette of the hill and some ground structure, which are overlayed on the images.
Finally, Fig. 8 shows the fused aligned images. The salient edges are registered in each of 3 blocks, the fourth block contains no edge points. The Hausdorff distance is used to find the optimal displacement assuming 5 percent outliers for the blocks covering the hill edge and 10 percent outliers for the edge in the lower right block. The specified search strategy finds the translation for each block such that 90 percent of the visible image edge points are no more than 5 pixels from some IR image edge point for the corresponding block. The local displacements are then used to determine the global affine transformation to register the two images. The estimated (x,y) displacements for the blocks upper left, upper right and lower right that are supplied to the global affine estimator for aligning the visible image to the IR image are (33,-3), (-11,-7) and (-8,-11) respectively.

Figure 8. Aligned and fused IR and visible images. Fusion is by weighted combination of image brightness values after alignment

The estimated affine transformation parameters that map point p in the visible image to the point p' in the IR image such that $p' = \mathbf{M}p+\mathbf{t}$ are

$$\mathbf{M} = \begin{bmatrix} 0.8239 & 0.0544 \\ -0.0179 & 0.9897 \end{bmatrix} \quad \text{and} \quad \mathbf{t} = (17.1925, -6.2471)^T.$$

Note that the image coordinate system origin is top left with positive x to the right and positive y down.

7. Conclusion

We have analyzed the problems in the real world visible/IR image registration. The area-based approaches are still feasible. However, feature extraction of structural edges as common features for registration and feature matching methods are more powerful to process with our low quality IR images. We have implemented multi-scale hierarchical edges detection and edge focusing and introduced a new salience measure for the horizon. For multisensor image registration, the common features extracted from images of two modalities can be still different in detail. Therefore, the transformation space match methods with the Hausdorff distance measures are more suitable than the direct feature matching methods. We have introduced image quadtree partition technique to the Hausdorff distance matching, that dramatically reduces the size of the search space into that of the search for translations which minimize the Hausdorff distance between corresponding blocks. We have shown image registration of visible/IR real world images of battle fields. The key point is to extract salient features from the real world images using local, regional and global information and appropriate salience measures.

References

[i] L. Sevigny, "RSG.9(panel3) Canadian test sequences", DREV internal report (1997).

[ii] Sharma R. K. and Pavel M., "Registration of video sequences from multiple sensors", Proc. Image Registration Workshop, NASA Goddard Cenetr, 361-364 (1997)

[iii] Brown L. G., "A survey of image registration techniques". ACM comput. surveys, vol.24, 325-376 (1992).

[iv] Canny J., "A computational approach to edge detection", Trans. IEEE PAMI-8, No.6, 679-698 (1986).

[v] Rothwell C., Mundy J., Hoffman B. and Nguyen V., "Driving vision by topology", Report No. 2444, INRIA, France (1994).

[vi] Bergholm F., "Edge focusing", IEEE Trans. PAMI-9, No.6, 726-741 (1987).

[vii] Ullman S. and Sha'ashua A., "Structural saliency: the detection of Globally salient structures using a locally connected network", A.I. Memo No. 1061 MIT (1988).

[viii] Duric Z. and Rosenfeld A. "Image sequence stabilization in real time", Real-Time Imaging, vol. 2, 271-284 (1996).

[ix] Huttenlocher, D.P., Klanderman, G.A., Rucklidge, W.J., "Comparing images using the Hausdorff distance," IEEE Trans. PAMI-15, No.9, 850-863 (1993).

Target Imagery Classification System (TICS)

Dr. Scott C. McGirr
SPAWAR Systems Center Code D722
53560 Hull Street
San Diego, CA 92152

Mr. Gerald Bartholomew
SPAWAR PD-18E1
4301 Pacific Highway
San Diego, CA 92110

Dr. Ronald Mahler
Lockheed Martin Tactical Defense Systems
3333 Pilot Knob Road
Eagan, MN 55121

Mr Robert Myre
Summit Research Corporation
1300 Diamond Springs Road
Virginia Beach, VA 23455

Abstract - *The Target Imagery Classification System (TICS) project is developing a state-of-the-art classification system for Inverse Synthetic Aperture Radar (ISAR) and other forms of target imagery. An open systems approach is being used for classification, which insures product outputs can be used in conjunction with others for multi-sensor data fusion. A two step process is used for target identification. First, essential features are extracted from ISAR video images. Second, features are compared to known ship feature sets to derive a classification. This process enables the construction of new message sets for reporting ISAR target features for support of information transfer and multi-sensor data fusion. The TICS system will also work on Synthetic Aperture Radar (SAR), Forward Looking InfraRed (FLIR), and Electro-Optical (E-O) Target Imagery.*

Key Words: imagery, radar, sensor, fusion, message, classification, automation, ship.

1. Introduction

A critical need exists for Automated Target Recognition (ATR) and identification through both cooperative means (e.g., IFF) and non-cooperative means. A strong need also exists on tactical platforms to perform target classification. Radar has proven effective at surveillance and classification of ships at extended ranges [1]. Three methods of radar target discrimination are High Range Resolution (HRR) radar, Inverse Synthetic Aperture Radar (ISAR), and Synthetic Aperture Radar (SAR). Classification and surveillance require different radar modes and can not be performed simultaneously.

Today, operators are trained to use specific methodologies to classify ships using imaging sensors. The classification process requires highly trained operators and is platform exposure time intensive. The need exists to develop methods to assist the operator in rapid target identification and reporting. The AN/APS-137 Radar [of the ISAR class] on P-3C aircraft and a similar system on the new SH-60R helicopters is used to classify surface ships. In contrast to the optical image of a ship, illustrated by Figure 1, for the ISAR class of radars the motion of ships due to wave action results in Doppler shift from structures in proportion to the height above the water line. This creates what appears as ship silhouettes like that of Figure 2. Individual images, or sequences of such images can be grouped and manipulated into a composite which can be compared with known ship profiles, parametric shapes and data for deriving top levels of target classification*.

---

* This technology may be the subject of one or more invention disclosures assignable to the U.S. Government. Licensing inquires may be directed to:
Harvey Fendelman
Patent Counsel
SPAWARSYSCEN Code D0012
San Diego, CA 92151-5765
(619) 553-3001

Figure. 1 Optical image.

The x-axis provides relative length; the y-axis shows proportional height.

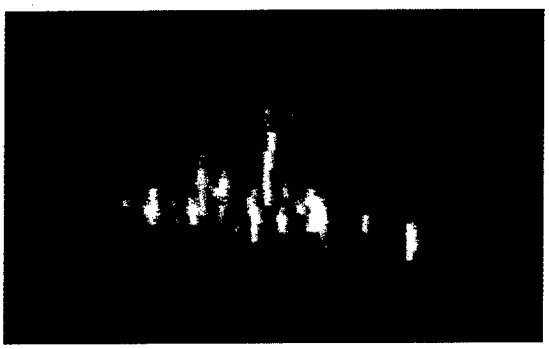


Figure 2 ISAR image.

The x-axis gives approximate ship length; the y-axis shows relative height (Doppler shift).

The vision of "network centric" warfare is the timely exchange of information over networks in the battlespace. Data is exchanged over networks by VMF messages, USMTF messages, TADIL-J Messages, and NITF imagery. The Joint Chiefs of Staff (JCS) and individual Services have developed C4ISR technical architectures for information exchange. It is important that data elements transferable by message be defined. The Navy has defined terminology and established a basic hierarchy for ship classification. This hierarchy uses Perceptual, Gross, Naval Fine and Type/Class/Unit levels of classification.

A great deal of work has been done on use of radar for target classification. Of particular interest, work has been done on ship classification by ISAR using 2D or 3D models by the Navy [2, 3, and 4]. This project, the result of a recent DOD-SPAWAR SBIR with Summit Research Corporation (SRC), with Lockheed Martin Corporation (LMC)-Eagan sub-contract support, builds on this earlier work with the goal of integrating operator and machine approaches to target classification and reporting standardization.

2. Approach

A two-step approach is taken for target identification. First, incoming imagery is enhanced and "focused" to provide an integrated, multi-frame summed target image, and then key features are extracted from sensor video imagery. Second, target features are compared to feature sets of known ship types to derive a classification. This two-step approach is important for three reasons: 1) This method is consistent with operator training. Operators are currently trained to recognize targets based on key attributes such as ship length, position of super structure, key uprights and shapes/features, etc. In this method, abbreviated data collection can provide a certain level of classification; greater levels of classification are gained with longer, "crisper" data collection. 2) This method supports information transfer. Target attributes can be transferred by message in text form with accompanying, integrated and enhanced summary target images, for reassessment at remote sites. 3) This method supports computer processing and multi-sensor data fusion. The target features can then be correlated with Link, video tape (VHS)-Imagery, Video-Relayed and/or parametrically-reported features obtained from other similar or dissimilar sensor types.

The analysis of target imagery and derivation of essential features for classification can be done by either

operator or machine. The advantage of machines is that they are consistent, do not tire during continued operation, and operate very fast. New Automatic Target Recognition (ATR) methods can support this process (Figure 3). The advantage of operator methods is that they can support selective extraction of easily understood target attributes over segments of video which is difficult to duplicate by mechanistic algorithms*. New Human Computer Interfaces (HCI) can be used to facilitate operators (Figure 4) and become a spring-board for further automated feature extraction.

In this project, an approach was taken that captures the strengths of both operator and machine. This was done by defining similar target attributes for feature extraction and classification by operator and machine in a parallel process. Incoming target images are first subjected to image declutter-noise suppression-enhancement techniques. A string of image frames are then integrated and summed to a robust 2D/3D composite image, and top level key (ship) target features are then automatically extracted [e.g. target length, bow/stern points, main (high doppler) superstructure position along the length/hull]. The ATR Classifier then performs Top-Level ship target classification at the Perceptual and Gross-levels of classification.

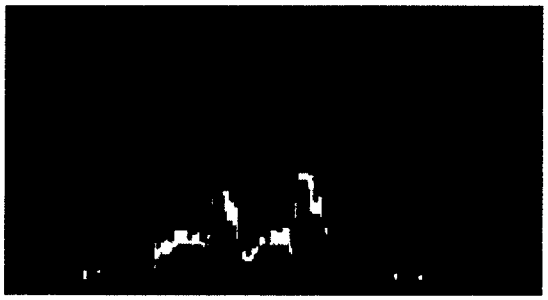


Figure 3 ATR Process.

Automated comparison of ship ISAR image, length and height [upper silhouette in light color] to ship profiles in model-base [lower silhouette in lower darker color].

%	DESCRIPTION		
16	SAM Launcher	50	Aft Superstructure
29	SAM Launcher	62	Gun
35	Director	71	Radar
42	For Superstructure	82	Superstructure
		98	Antenna

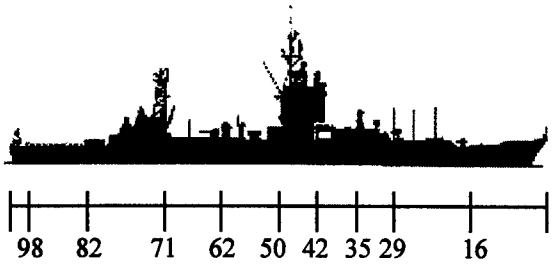


Figure 4 HCI Process.

Manual feature extraction and comparison to feature sets in electronic knowledge and model Dbases.

Resulting OTH Gold (Modified) Message Report Format:

ISTGTXXX/L1023/W22/S35/M39/M57/S 70/S85/W90//55NM

The ATR classifier accepts imagery as either raw sensor data (In-phase & Quadrature-phase, or I&Q) or video tape in standard VHS format. LMC, under SBIR subcontract to SRC, developed the ATR algorithms used. The ISAR imagery data is input into a computer where ATR algorithms are used to clean up the image and extract ship broadside outline. In Phase I tests, the ship outline was matched against a two dimensional (2D) database of models of known ship types. The result was a ranked list of likely target types from a model base, at the top, Perceptual and Gross levels only. Major issues for this type of classifier are separating targets types that are similar at the model level and classifying new targets not in the model base.

A HCI was developed to assist operators once they have extracted features from sensor displays with the subsequent

* Optimal feature vectors extracted by machine may not be easily understood by operators (5).

classification process. SRC developed the HCI used. The Parametric Classifier algorithms and routines accept features from ISAR imagery and can determine target identity by scoring against parametric knowledge-based feature models at an appropriate Naval and Fine (class/unit level) level in a SRC-designed hierarchical knowledge base. If there is insufficient parametric evidence to support a fine level classification, a partial classification at a higher level can be made. A major issue regarding this type of classifier is the development of robust automated feature estimation or extraction techniques.

A significant advantage of this Target Imagery Classification System design is that only 10-15 2D/3D "models" of composite Perceptual and Gross-level ship types need to be built. Only the "model match" has selected a Perceptual and Gross target classification category, the SRC-developed Parametric Target Imagery Knowledge Dbase takes over and classifies the target/ship to Type-Class-Unit using key ship image features extracted by HCI-operator input or through the use of developing auto-features extraction algorithms. This makes it unnecessary to 2D/3D model thousands of the worlds ships. The Parametric Target Imagery knowledge Dbase can inexpensively build individual Type-Class-Unit level parametric models very efficiently, using multiple sources of input data which already exist.

3. Results

Real world, operationally collected ISAR imagery test data sets were obtained from the AN/APS-137 Radar System. These empirical data sets included 45 video tape (VHS) targets and 26 parametric feature reports in SRC modified, Over-The-Horizon Gold message format (OTG). An overview will be given; further details can be obtained from SRC.

The baseline Target Imagery Classification System (TICS)-ATR System correctly classified all 45 video images down to the Class/Unit level. The test sets were of varying image quality but all assessed as containing essential detail for classification. Auto-focus algorithms were used for processing of ISAR I&Q or video tape images. Images were decluttered, enhanced and integrated, and selective frames captured. The computer encoded the 2D pattern as a function of Doppler and range. Pattern matching was based on use of deformable templates [6]. The advantage of this approach is that it is invariant to Doppler affects of translation, "shearing" and inversion. Target model association is then done using normalized Cross Correlation (CC) and a String Edit Metric. The features used for classification included ship length and height profile and overall shape/silhouette. Future capability exists to include flash points created by rotating radar.

The HCI System then took operator selected features and then correctly classified 25 of 26 OTH-Gold message formatted target reports. Ships were classified in a hierarchical fashion using the four levels: Perceptual, Gross, Naval Fine and Type/Class/Unit level. The target test database included 45 ships. Perceptual classes consisted of carrier (CV) with 2 ships, combatant (CBT) with 30 ships, auxiliary (AUX) with 8 ships, and small craft (SC) with 5 ships. The approach to classification was based on the modeling and hierarchical methods currently used by operators. Low resolution images support gross features used for classification such as ship length and relative position of superstructures. Higher resolution images support additional recognizable objects such as guns, antenna and missile launchers, etc.

Prototype systems were demonstrated by SRC and LMC-Eagan at SPAWAR Systems Center on 7/98 and 9/98. The two systems now operate in series. It is desired to integrate the ATR and HCI functions into one robust system. Interest has been expressed for use of the automated ISAR classification capability on P-3C aircraft

and SH-60R helicopters. Other Navy applications may include Tactical Support Centers (TSC) and Mobile Operation Command Centers (MOCC). In addition, there are applications for other military Services and commercial industry.

4. Conclusions

The Phase I SBIR and SRC/LMC IR&D Project results show that a robust classification process can be developed that incorporates the strengths of human and machine based systems. Major issues for this type of classifier are separating targets that are similar at the model level and classifying new targets not in the model-base through the use of parametric Dbase matching. The ATR approach uses machine extraction of ISAR imagery to classify based on length, height, and overall shape/silhouette probability-of-match. The ship profile model-base is processed for matches and a ranked list of likely target types is produced. The HCI approach then uses operator assisted or automated parametric inputs from ISAR imagery displays to classify based on ship length and other-key features. The hierarchical database supports classification to the appropriate level (e.g., to the Type-Class-Unit level) as supported by available data. Areas for future enhancement have been identified and discussed.

It has been shown that a robust ISAR/Image message set can be used to report relevant target attributes to the level necessary to support detailed target classification. This would be of value for reassessment of target identity reported as well as for combination with data from other sensors for multi-sensor data fusion. The vision of "network centric" warfare requires sharing of data in the battlespace. To make this vision a reality, an architecture must be defined for sharing relevant information. It is essential that data elements be defined and put in a structured form for dissemination and automated processing. The TICS project works towards this vision with the development of a prototype ISAR/Image message set and video-imagery for transfer of data over networks to higher echelons for support of "Sensor Grid" correlation and data fusion.

5. References

1. Wehner, Donald R., *High Resolution Radar (2nd Edition)*, 1995 Artech House, Inc, 685 Canton St., Norwood, MA, 02062.

2. McCune, Brian P., Robert J. Drazovich, "Radar With Sight and Knowledge", Defense Electronics, V.15, Aug 1983, p. 80-96.

3. Baras, John and Sheldon Wolk, "Hierachical Wavelet Representations of Ship Radar Returns", Technical Report NRL/FR/5750-94-9593, Feb 22, 1994.

4. Musman, Scott, D. Kerr and C. Bachman, "Automatic Recognition of ISAR Ship Images", IEEE Transactions on Aerospace Electronic Systems, Vol .23, No.4,1996.

5. Zu-Han Gu and Sing H. Lee, "Optimal Implementation of the Hotelling Trace Criterion for Image Classification", Optical Engineering, Vol. 23, No. 6, Dec 1984.

6. Jain, A. K., Yu Zhong and S. Lakshmanan, "Object Matching Using Deformable Templates", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 3, Mar 1996 p.267-78.

Fusion of color information for image segmentation based on Dempster-Shafer's Theory

P. VANNOORENBERGHE
PSI, UNIVERSITE/INSA de ROUEN
Place Emile Blondel, BP 08
76131 Mont-Saint-Aignan Cedex, France
Patrick.Vannoorenberghe@univ-rouen.fr

O. COLOT
PSI, UNIVERSITE/INSA de ROUEN
Place Emile Blondel, BP 08
76131 Mont-Saint-Aignan Cedex, France
Olivier.Colot@insa-rouen.fr

D. DE BRUCQ
PSI, UNIVERSITE/INSA de ROUEN
Place Emile Blondel, BP 08
76131 Mont-Saint-Aignan Cedex, France
Denis.Debrucq@univ-rouen.fr

Abstract : ***In this paper, we propose a color image segmentation method based on the Dempster-Shafer's theory. The tristimuli R, G and B are considered as three independent information sources which can be very limited or weak. The basic idea consists in modeling the color information in order to have the features of each region in the image. This model, obtained on training sets extracted from the gray level image, allows to reduce the classification errors concerning each pixel of the image. The proposed segmentation algorithm has been applied to synthetic and biomedical images in order to illustrate the methodology.***

Keywords: Color Image, Segmentation, Dempster-Shafer's Theory, Data fusion.

1 Introduction

In color image segmentation, color of a pixel is given as three values corresponding to the well known tristimuli R (Red), G (Green) and B (Blue). Different kinds of colors spaces have been developed by several authors [1], [2], [3], [4]. They are derived from this representation of the color using linear and nonlinear transformations. In the framework of segmentation, each color model is more or less convenient, efficient or reliable [5]. The major problem consists in choosing the adapted color model for a specific application. In our study, we choose to work only with the tristimuli (R, G and B) given by the sensor. Each color plane is considered as an information source which can be imprecise or uncertainty. The basic idea of our purpose consists in combining these three information sources using the Dempster-Shafer's theory of evidence [6]. This well known tool in classification problems [7] provides a convenient framework which allows modeling uncertainty in situations where the available evidence is limited or weak. Some works related to image processing propose to use this approach derived from the confidence measure theory [8]. The proposed method uses the formalism of belief functions to represent the color information provided by each training set. Section 2 introduces the problem we want to solve in the framework of Dermatology. We present in the section 3 the segmentation strategy and propose finally some experimental results. The segmentation algorithm has been applied to biomedical images in order to detect a form of skin cancer.

2 Color Image Processing in Dermatology

In Dermatology Science, melanoma is an increasing form of cancer. It has increased twice times for 15 years in Canada and it is now 3% of cancers in the USA. The rates of clinical diagnostic accuracy are about 65% at the very best. In particular, it is very difficult to distinguish some atypical lesions - which are benign - from melanoma because they have the same properties according to the well known ABCDE rules used by dermatologists [9]. There is a visual inspection problem for the atypical lesions class. Unnecessary excisions are often practise for these lesions. The variability of colors and shapes (see Figure 1) can lead to several interpretation by different dermatologists. However, melanoma

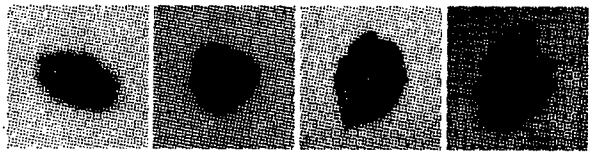


Figure 1: Original images of lesions

is well suited for color image processing because it is on the skin. Some researches [10] have shown the advantages to use image processing in dermatology. Furthermore, the image processing by computer ensures the reproductibility of the analysis. However, the essential difficulty is to design robust and relevant parameters to ensure the separation between melanoma and benign lesions, in particular the atypical lesions (benign), called naevus, which can be clinically mistaken for melanoma. At first, the lesion border is the feature to identify. It is the first step of the processing to engage in order to extract information about the lesion. So, the border extraction or identification is a critical step in computerized vision analysis in skin cancer as pointed out in [10]. Then, the segmentation step which takes place in any classification processing has to be really accurated. In the framework of our application, only two regions are considered. So, the problem is to separate lesions from the surrounding safe skin. So as to obtain geometric and colorimetric information on a lesion, it is necessary to run a segmentation process which will allow to extract the pixels belonging to the lesion from the image. Dempster-Shafer's theory of evidence [6] is also used in two different steps of the detection system. It is first used in the segmentation scheme but managing uncertainty in the classification procedure is very important. Figure 2 illustrates this both utilisation.

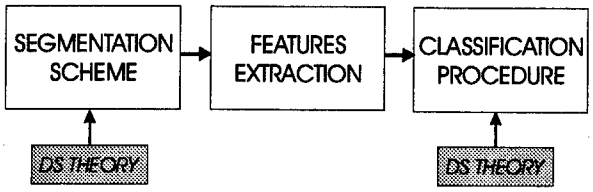


Figure 2: Dempster-Shafer's Theory

3 The segmentation scheme

A segmentation of an image I is a partition of I into disjoint nonempty subsets $\mathcal{R}_u$ for $u = 1, 2, ..., U$ such as :

$$I = \bigcup_{u=1}^{U} \mathcal{R}_u \qquad (1)$$

Under the assumption that images containing only two regions, we can compute a single threshold on the gray level image obtained by means of the Maximum Entropy Principle (MEP) [11]. This coarse segmentation gives two training sets containing pixels which belong surely to one of the considered regions. This first segmentation, based only on the use of gray level image, induces some classification errors. The proposed method is based on the color information contained in the image. It is decomposed in three steps :

- Modeling the belief on the training sets,
- Combining the M information sources with the Dempster's rule,

- Taking a decision to classify each pixel P to a region $\mathcal{R}_u$.

For our segmentation scheme, we choose to work with $M = 3$ where the different information sources are the tristimuli R, G and B.

3.1 Modeling the belief on the training sets

Let Θ represents the finite set of regions $\mathcal{R}_u$ such as :

$$\Theta = \{\mathcal{R}_u\} \text{ for } u = 1, 2, ..., U \qquad (2)$$

Each color plane is assimilated to an information source S_i for $i \in 1, ..., M$. Let us consider a basic belief assignment m_i defined as :

$$m_i : 2^\Theta \longmapsto [0, 1] \qquad (3)$$

with $m_i(\emptyset) = 0$ and $\sum_{A \subseteq \Theta} m_i(A) = 1$. Under the assumption of Gaussian distributions, basic belief functions can be written :

$$m_i(\mathcal{R}_u) = \frac{R_i}{\sigma_u \sqrt{2\pi}} \exp^{-\frac{(x_i - \mu_u)^2}{2\sigma_u^2}} \qquad (4)$$

where x_i is a realization of a M-dimensional random variable X. In our case, x_i is the value of a pixel P for one of the three color planes. The values $\mu_u = E(X)$ and $\sigma_u^2 = E(X - E(X))^2$ are respectively the mean and the variance on the region $\mathcal{R}_u$. These values are replaced by their statistical approximations computed on the training sets. The advantage of Dempster-Shafer's theory lies in representing uncertainty by means of a belief on the whole frame of discernment. This basic belief assignment allows to define $m_i(\Theta)$ with the following equation :

$$m_i(\Theta) = \frac{R_i}{\sigma_\Theta \sqrt{2\pi}} \exp^{-\frac{(x_i - \mu_\Theta)^2}{2\sigma_\Theta^2}} \qquad (5)$$

with $\mu_\Theta = (\mu_1 + \mu_2)/2$ and $\sigma_\Theta = \max(\sigma_1, \sigma_2)$. In the equations (4) and (5), the coefficient R_i is a normalization coefficient. It allows to verify the condition $\sum_{A \subseteq \Theta} m_i(A) = 1$.

3.2 Belief function attenuation

An additional aspect of the Dempster-Shafer's theory concerns the attenuation of the basic belief assignment m_i by a coefficient α_i. The attenuated belief function $m_{(\alpha,i)}$ can be written as :

$$m_{(\alpha,i)}(\mathcal{R}_u) = \alpha_i . m_i(\mathcal{R}_u) \quad \forall \mathcal{R}_u \in 2^\Theta \qquad (6)$$

$$m_{(\alpha,i)}(\Theta) = 1 - \alpha_i + \alpha_i . m_i(\Theta). \qquad (7)$$

The problem consists in evaluating for each source S_i, the coefficient α_i in order to have the more certain information to aggregate. After the learning step, the main idea is to resume the information contained in each source S_i by means of an optimum histogram computed on the set $\bigcup_{\mathcal{R}_u} \mathcal{X}_{(u;i)}$ in the sense of the maximum likelihood and of a mean square cost. This histogram will be used in order to establish the relevance of a source of information. First, we have to build an approximation of the unknown probability distribution with only the samples given in each source. That is done by means of a histogram building which is led by the use of an information criterion. We will see that different information criteria initially designed for model selection can be used [12], [13]. Once this histogram is obtained, we use the Hellinger's distance between the approximated distribution computed on the set $\mathcal{X}_{(u;i)}$ and the approximated distribution computed on the set $\mathcal{X}_{(u';i)}$. This distance gives a dissimilarity between the two probability densities that is to say the ability of the source to distinguish the two regions $\mathcal{R}_u$ and $\mathcal{R}_{u'}$.

3.2.1 Probability density Approximation

Let be $A_1 A_2 \ldots A_p \ldots A_q$ an initial partition Q of an unknown distribution λ with $q = Card(Q)$. The aim is to approximate λ with a histogram built on a subpartition $C = B_1 B_2 \ldots B_c$ of Q with c bins such as $c \leq q$. The probability distribution $\hat{\lambda}_C$ built with C is an optimum estimation of λ according to a cost function to define. C results from an information criterion called IC issued from

the basic Akaike's information criterion (AIC) [12], AIC^* or ϕ^* [13] which are respectively Hannan-Quinn's criterion and Rissanen's criterion. These criteria have the following form :

$$IC(c) = g(c) - \sum_{B \in C} \hat{\lambda}_c \ln \frac{\hat{\lambda}_c(B)}{\nu_c(B)} \qquad (8)$$

where $g(c)$ is a penalty which differs from one criterion to another one. Let us note ε a random process of a probability distribution λ supposed absolutely continuous to an *a priori* given probability distribution ν. Let ω be the set of all values taken by ε. The probability density f of λ is given by the Radon-Nycodim's derivative such as :

$$\forall \epsilon \in \omega \qquad f(\lambda, \epsilon) = \frac{d\lambda}{d\nu}(\epsilon). \qquad (9)$$

The probability density f is approximated from N samples (ϵ_k) of ε by means of a histogram with c bins obtained with these N values. An optimum histogram to approximate the unknown probability distribution λ is obtained in two steps. The first one consists in merging two contiguous bins in a histogram with c bins among the $(c-1)$ possible fusions of two bins. This is made by minimizing the IC criterion. The second one consists in finding the "best" histogram with c bins. The optimum histogram with $c = c_{opt}$ bins is the one which minimizes IC.

3.2.2 Maximum likelihood estimator for a partition Q

Let Q be a partition with q bins and let $\epsilon_1 \ldots \epsilon_N$ be a N-observation sample and let be λ_Q the probability distribution according to Q. The maximum likelihood estimator $\hat{\lambda}_Q$ of λ_Q is given by the following equation :

$$\forall p \in \omega \qquad \hat{\lambda}_Q(A_p) = \frac{1}{N} \sum_{\epsilon_k \in A_p} \epsilon_k \qquad (10)$$

where A_p is a bin of the partition Q. This result derives from the density expression of λ_Q :

$$\forall \epsilon \in \omega \qquad f(\lambda_Q, \epsilon) = \sum_{A \in Q} \frac{\hat{\lambda}_Q(A)}{\nu(A)} 1_A(\epsilon) \quad (11)$$

with $1_A(\epsilon) = 1$ if $\epsilon \in A$ and 0 otherwise.

3.2.3 Selection of the bin number of a histogram

The obtaining of the optimum histogram is based on the use of an information criterion IC which gives the number of bins optimal thanks to a cost function based on the Kullback's contrast or the Hellinger's distance. We define the cost to take $\hat{\lambda}$ when λ is the true probability density by :

$$W(\lambda, \hat{\lambda}) = E_\lambda \left(\psi \left[\frac{f(\hat{\lambda}, \epsilon)}{f(\lambda, \epsilon)} \right] \right) \qquad (12)$$

where E_λ is the mathematical expectation according to λ and ψ is a convex function. According to the expression of ψ the cost function leads to different information criteria to choose the histogram with c bins. So, if ψ is the Hellinger's distance we get :

$$AIC(c) = g(c) - 2 \sum_{B \in C} \hat{\lambda}_c(B) \ln \frac{\hat{\lambda}(B)}{\nu(B)}. \quad (13)$$

with $g(c)$ represents the penalty term defined as :

$$g(c) = \frac{2c-1}{N}. \qquad (14)$$

It can be seen that it is identical to the classical Akaike's information criterion. If the cost function $W(\lambda, \hat{\lambda})$ is expressed according to the KullBack's contrast, we obtain two new criteria $\phi^*(c)$ and $AIC^*(c)$ with different penalty terms $g(c)$ defined respectively as :

$$g(c) = \frac{c(1 + \ln(\ln N))}{N} \qquad (15)$$

$$g(c) = \frac{c(1 + \ln N)}{N} \qquad (16)$$

These criteria can be used to select the optimum histogram with c bins to approximate the unknown probability density of a N-sample. Detailed demonstrations are available in [12] or [13].

3.2.4 Optimum histogram building process

At first, an initial histogram with $q = Card(Q) = 2 \times In[\sqrt{N} - 1]$ bins is built giving the partition Q, where $In[\ ]$ denotes the integer part. Then, a partition with $(q-1)$ bins is considered. For each possible fusion of two contiguous bins among $(q-1)$ the criterion $IC(q-1)$ is computed. The choice of the best fusion is made according to the minimization of $IC(q-1)$. When it is done, we look for the best partition with $(q-2)$ bins according to the same rule. Finally, the histogram with c bins such as $IC(c)$ for $c \in \{1, \ldots, q\}$ is retained. An initial histogram is built with a N-sample ($N = 90$) randomly generated according to a gaussian distribution with mean equal to 0 and with a variance equal to 1. Figure 3 gives the behaviour of the three criteria. It can be seen that AIC^* and ϕ^* give the same final histogram. AIC gives a final histogram with an upper bin number. This difference is linked to the type of convergence for each information criterion [13]. The optimum histogram is

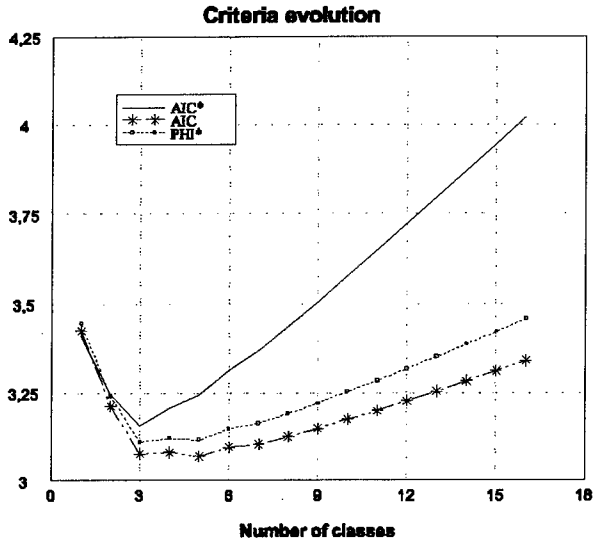


Figure 3: Criteria evolutions

computed on the set $\bigcup_{\mathcal{R}_u} \mathcal{X}_{(u;i)}$. Once this histogram is obtained, we use the Hellinger's distance between the approximated distribution $\hat{\lambda}_C$ computed on the set $\mathcal{X}_{(u;i)}$ and the approximated distribution $\hat{\lambda}'_C$ computed on the set $\mathcal{X}_{(u';i)}$. This distance gives a dissimilarity between the two probability densities that is to say the ability of the source to distinguish the two hypotheses $\mathcal{R}_u$ and $\mathcal{R}_{u'}$.

3.3 Fusion of several sources

The Dempster-Shafer's theory allows the fusion of several sources using the Dempster's combination operator. It is defined like the orthogonal sum (commutative and associative) following the equation :

$$m(\mathcal{R}_u) = m_1(\mathcal{R}_u) \oplus \cdots \oplus m_M(\mathcal{R}_u) \qquad (17)$$

For two sources S_i and $S_{i'}$, the data fusion can be written as :

$$m(\mathcal{R}_u) = \frac{1}{\mathcal{K}} \sum_{\mathcal{R}_v \cap \mathcal{R}_w = \mathcal{R}_u} m_i(\mathcal{R}_v).m_{i'}(\mathcal{R}_w). \qquad (18)$$

where $\mathcal{K}$ is defined by :

$$\mathcal{K} = 1 - \sum_{\mathcal{R}_v \cap \mathcal{R}_w = \emptyset} m_i(\mathcal{R}_v).m_{i'}(\mathcal{R}_w). \qquad (19)$$

The normalization coefficient $\mathcal{K}$ evaluates the conflict between two sources. $\mathcal{K} = 0$ corresponds to the case where the sources are totally in conflict.

3.4 Decision rule

The credibility Bel and the plausibility Pl can be computed from the basic belief assignment using following equations :

$$Bel(\mathcal{R}_u) = \sum_{\mathcal{R}_v \subseteq \mathcal{R}_u} m(\mathcal{R}_v) \qquad (20)$$

$$Pl(\mathcal{R}_u) = \sum_{\mathcal{R}_u \cap \mathcal{R}_v \neq \emptyset} m(\mathcal{R}_v). \qquad (21)$$

Finally, the decision is made by assigning a pixel P to a region $\mathcal{R}_u$ with the maximum credibility or with the maximum plausibility. The first one corresponds to a pessimistic decision rule and the second one to an optimistic decision rule. Some results are presented in the

section 4. We attenuate each belief function m_i according to the equation (7) where α_i is the Hellinger's distance. The information sources S_i are then aggregated using the Dempster's combination rule. The decision rule is based on the decision function δ which assignes a pixel P to the region $\mathcal{R}_u$ following :

$$\delta(P, \mathcal{R}_u) = u \text{ iif } \mathcal{R}_u = \arg \max_{\mathcal{R}_v \in 2^\Theta} (Bel(\mathcal{R}_v)) \tag{22}$$

or

$$\delta(P, \mathcal{R}_u) = u \text{ iif } \mathcal{R}_u = \arg \max_{\mathcal{R}_v \in 2^\Theta} (Pl(\mathcal{R}_v)) \tag{23}$$

4 Experimental Results

This section is devoted to present some results concerning the segmentation scheme in order to evaluate the methodology. The proposed segmentation algorithm has been applied to biomedical images in order to illustrate the strategy. Some images, in the context of dermatology, are presented in the figure 4. First

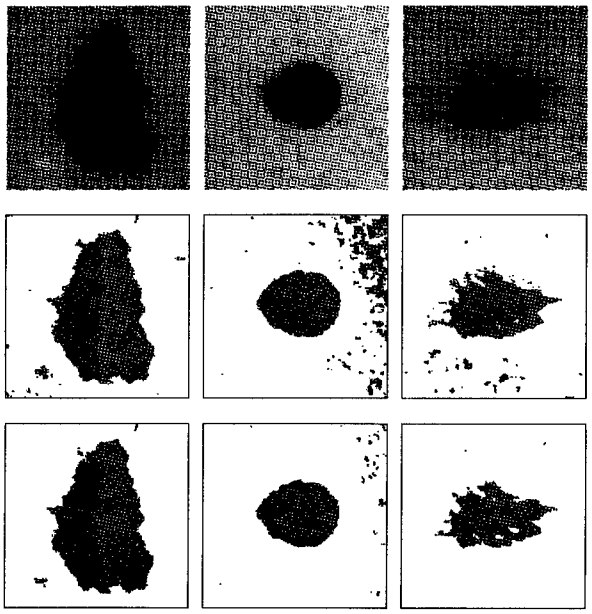


Figure 4: Results images

row corresponds to the original color images. The respective two other rows represent the segmentation scheme results with the decision concerning the maximum of credibility (second row) and the maximum of plausibility (third row). We can note that the lesion (red color) is correctly extracted from the safe skin (white color). The blue color corresponds to pixels which cannot be classify either to the safe skin or to the lesion. We present in the figure 5 some images containing edges superimposed on the original images of lesions.

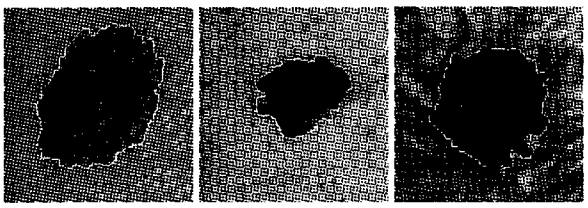


Figure 5: Edges detection

5 Conclusion

In this paper, we have presented an original color image segmentation procedure using both information criteria and Dempster-Shafer's theory. The proposed methodology consists in intializing the belief functions with probability densities obtained by learning. By means of information criteria, we determine the attenuation of the belief assignment based on the dissimilarity between probability distributions. This framework allows to use the whole information contained in the image as better as possible. The proposed methodology has been applied in order to classify lesions in the framework of a kind of skin cancer frequently meet in Dermatology science. Future work is concerned with analysis of several decision rules using uncertainty measures proposed by Klir [14, 15].

6 Acknowledgements

The authors would like to acknowledge the support of the Normandy regional council. The co-operation of P. Joly of the clinic of dermatology of Charles Nicolle's hospital of Rouen in

supplying diagnosis on melanoma and acquisition of image lesions, is also gratefully acknowledged.

References

[1] R. W. G. Hunt. *Measuring Color.* Applied science and industrial technology. Ellis Horwood, 2 edition, 1991.

[2] M. Miyahara and Y. Yoshida. Mathematical transform of (r,g,b) color data to munsell (h,s,v) color data. In *SPIE Proceedings : Visual Communications and Image Processing*, volume 1001, pages 650–657, San-Jose, 1988. SPIE.

[3] N. Ohta. Correspondance between cielab and cieluv color differences. 2(4):178–182, 1977.

[4] G. Wyszecki and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae.* John Wiley & sons, second edition, 1982.

[5] Y.I. Ohta, T. Kanade, and T. Sakai. Color information for region segmentation. *Comput. Graph. and Imag. Proc.*, 13:222–241., 1980.

[6] G. Shafer. *A Mathematical Theory of Evidence.* Princeton University Press, 1976.

[7] T. Denoeux. Analysis of evidence-theory decision rules for pattern classification. *Pattern Recognition*, 30(7), 1998.

[8] I. Bloch and H. Maître. Fusion de données en traitement d'images: modèles d'informations et décisions. *Traitement du Signal*, Numéro Spécial: Fusion de données(11):435–446, 1994.

[9] Friedman R. J. Early detection of melanoma: The role of physician examination and self examination of the skin. *CA*, 35:130–151, 1985.

[10] Golston J. E. et Al. Boundary detection in skin tumor images: An overall approach and radial search algorithm. *Pattern Recognition*, 23(11):1235–1247, 1990.

[11] Kapur J. N. A new method for grey-level picture thresholding using entropy of the histogram. *CVGIP*, 29:273–285, 1985.

[12] O. Colot. *Apprentissage et détection automatique de changements de modèles - Application aux signaux électro-encéphalographiques.* PhD thesis, Université de Rouen, 1993.

[13] O. Colot and al. Information criteria and abrupt changes in probability laws. In M. Holt, C. Cowan, P. Grant, and W. Sandham, editors, *Signal Processing VII: Theories and Applications*, pages 1855–1858. EUSIPCO'94, September 1994.

[14] G.J. Klir and T.A. Folger. *Fuzzy Sets, Uncertainty, and Information.* Prentice Hall P T R, Englewood Cliffs, New Jersey 07632, 1988.

[15] Z. Wang and G.J. Klir. *Fuzzy Measure Theory.* Plenum Publishing Corporation, 233, Spring Street New York, 1992.

Testbed for Fusion of Imaging and Non-Imaging Sensor Attributes in Airborne Surveillance Missions

Alexandre Jouan and Pierre Valin
Lockheed Martin Canada,
6111 Royalmount Ave., Montréal, Québec,
H4P 1K6, Canada
Email: alexandre.jouan@lmco.com , pierre.valin@lmco.com

Eloi Bossé
Defense Research Establishment Valcartier,
2459 Blvd Pie XI Nord, Val Bélair, Québec,
G3J 1X5, Canada
Email: eloi.bosse@drev.dnd.ca

Abstract - *An Adaptable Data Fusion Testbed (ADFT) has been constructed that can analyze simulated and/or real data with the help of modular algorithms for each of the main fusion functions and image interpretation algorithms. The results obtained from data fusion of information coming from an imaging Synthetic Aperture Radar (SAR) and non-imaging sensors (ESM, IFF, 2-D radar) on-board an airborne maritime surveillance platform are presented for two typical scenarios of Maritime Air Area Operations and Direct Fleet Support. An Image Support Module (ISM) has been designed and implemented that consists of a four-step hierarchical SAR classifier that can extract attributes such as ship length, ship category, ship type and (in the future) ship class. The SAR classifier can distinguish between merchant and combatant categories and can select amongst 5 combatant types and it estimates confidence levels for each sensor declaration that it produces, for example through the use of properly trained neural nets. A truncated Dempster-Shafer evidential reasoning scheme is used that proves robust under countermeasures and deals efficiently with uncertain, incomplete or poor quality information. Since the Dempster-Shafer method reasons over an exhaustive list of all possible platforms, an extensive set of realistic databases has been created that contains over 140 platforms, carrying over 170 emitters and representing targets from 24 countries.*

1. Introduction

This paper describes an existing Adaptable Data Fusion Testbed (ADFT) which is based on a Knowledge-Based System (KBS) BlackBoard (BB) architecture to perform data fusion of imaging and non-imaging sensors present on-board the CP-140 Canadian maritime patrol aircraft. Much additional material is presented here compared to [1] and only a quick summary of relevant data from [1] is presented for clarity.

The ADFT architecture must process the data coming from radar, Electronic Support Measures (ESM), Identification Friend of Foe (IFF) and datalink information both for the *planned* Aurora Modernization Program (AMP) and the Maritime Helicopter Project (MHP) which will replace the ageing Sea Kings. The new sensors that are exclusively present on the airborne platforms are of the imaging type, namely the Forward Looking Infra-Red (FLIR) and Synthetic Aperture Radar (SAR) which can operate in Strip Map, RDP and Spotlight modes (Adaptive or Non-Adaptive). The attribute data that these sensors can provide is important in determining the identification of target platforms, particularly the long range features that the Spotlight SAR can furnish.

2. ADFT Architecture

The real-time KBS BB shell developed by Lockheed Martin (LM) Canada and Defence Research Establishment Valcartier (DREV) is the basis of the ADFT infrastructure. This system is totally generic, and could be used to implement any system comprised of components which can be numeric or AI based. It has been implemented in *C++* rather than in a higher-level language (such as LISP, Smalltalk, ...) to satisfy the real-time requirement.

The testbed is designed to accommodate modular interchangeable algorithm implementation and performance evaluation of:

1. Fusion of positional data from imaging and non-imaging sensors;
2. Fusion of attribute information obtained from imaging and non-imaging sensors and other sources such as communication systems, satellites, etc., and
3. Object Recognition (OR) in imaging data.

The algorithms incorporate state-of-the-art tracking in clutter and evidential reasoning for target identification. The end result offers the user a flexible and modular environment providing capability for:

1. addition of user defined sensor simulation models and fusion algorithms;
2. integration with existing models and algorithms, and

3. evaluation of performance to derive requirement specifications and help in the design phase towards fielding a real Data Fusion (DF) system.

3. Fusion Function Implementation

Any generic DF application must contain the following set of sequential functions to act on real or simulated data:

1. *registration* to first perform spatial and temporal alignment of the simulated sensor data,

2. an *association* mechanism to then correlate the new incoming data with possible existing tracks found in the BB database and to send associated positional data to positional fusion and associated attribute data (e.g. image features of a given target) to information fusion,

3. *positional estimation* to then update the tracks in the time domain with the associated new data and write this positional information to the BB database, possibly extracting attribute data such as speed, acceleration and sending to information fusion, and

4. *identification estimation* (or information fusion) to then fuse all attribute data through evidential reasoning, whether they originate from imaging (through image understanding and feature extraction) or non-imaging sensors and consequently update the dynamic BB track database.

The control flow for the fusion of information is data driven directly from the simulators. The algorithms used within the DF function include: Jonker-Vogenant-Castanon (JVC) algorithm which is an optimal single-scan associator for the *association* function, Kalman filters for the *positional estimation* function, and LM Canada developed truncated Dempster-Shafer (DS) algorithm for the *identification (ID) estimation* function. The positional estimation function uses radar, IFF, ESM and Link-11 data and ID estimation uses IFF, ESM, Link-11 and imaging features.

4. Database Attributes For Identification

For ID estimation to be properly achieved, all possible attributes that can be measured by all of the sensors must be listed in the Platform DataBase (PDB). The attributes which we have catalogued in the PDB split into 3 groups (more explanations can be found in [1]):

1. *Kinematic attributes* which can be estimated by tracking by positional estimation, IFF and Link-11: the maximum acceleration ACC, the maximum platform speed V_MAXI and the minimum platform speed V_MINI all serve as bounds to discriminate between possible air target identifications. ALT_MAXIM is the maximum altitude that a platform may reach, which serves as a bound for altitude reported by the IFF.

2. *Geometrical attributes* which can be estimated by algorithms within the FLIR and the SAR classifiers: in addition to the three geometrical dimensions of height, width and length, one also needs the variables RCS_FOR, RCS_SID, RCS_TOP corresponding respectively to radar cross-section (RCS) of the platform seen from the front, the side and the top. The RCS values are empirically much larger than the geometrical cross-section obtained by the product of the two relevant dimensions (HEI, WID, LEN) since metallic objects offer strong radar backscatter when compared to the geometrical cross-section.

3. *Identification attributes* which can be directly given by the ESM, or as outputs of the FLIR and SAR ISM. ACRO is the acronym of the country name indicated in the GPL and used also to refer to the country that owns the platform in the PDB. In the PDB, ACRO is used by the attribute fusion function to link the PDB platform with the country allegiance indicated in the GPL. The variable EMITTER_LIST is an exhaustive list (labelled by number) of all the emitters that are carried by the platform. The variable PLATYPE forms the first level of platform classification used in this PDB. This variable is closely related to the category descriptor given by the ISM and reflects its platform military utilization.

Some sensors measure attributes quite directly. For example the ESM will provide an emitter list with some confidence level about the accuracy of the list that reflects the confidence in its electromagnetic spectral fit. However an IFF response can lead to an identification of a friendly or commercial target but the lack of a response does not necessarily imply that the interrogated platform is hostile. One has to distribute the lack of a response between at least two declarations: the most probable foe declaration and a less probable friendly or neutral declaration that allows for an IFF equipment that is not working or absent.

Similar complications arise when dealing with kinematic parameters reported occasionally by the tracker in *positional estimation.* Firstly, each physical quantity has a different dimension (speed, acceleration) and an accurate determination is not necessarily needed for fusion. Indeed it is convenient

to bin the attribute "speed" into fuzzy classes like "very fast", "fast", "average", "slow" and "very slow" (separately for air and surface targets). The same can be done for length in bins of 40 meter width for example. Membership in each class is a measure of how well the measured value fits into the descriptor as described below. Further, speed reports must be fused only if they involve a significant change from past historical behaviour in that track. The reason is two-fold: firstly no single sensor must attempt to repeatedly fuse identical ID declarations otherwise the hypothesis that sensor reports are statistically independent is violated, and secondly the benefits of the fusion of multiple sensors is lost when one sensor dominates the reports. Furthermore, a measured value of speed only indicates that the target is capable of that speed, not that it corresponds to either the maximum or minimum speeds listed in the PDB. It is a reasonable working hypothesis to fuzzify the value reported by the tracker into adjacent "bins" to account for the target being at, say only 80% of its optimal speed (a "very fast" target can occasionally travel "fast"), or travelling with a strong tailwind (a "fast" target can occasionally appear as "very fast"). Finally the concept of binning can be generalized to continuous membership functions of a fuzzy set.

The PDB presently used contains over 140 platforms, carrying over 170 emitters (which are listed in an Emitter Name List) and representing targets from 24 countries (which are enumerated in an Geo-Political Name List that serves to determine allegiance on any given mission).

5. Identification Estimation

A truncated Dempster-Shafer evidential reasoning scheme is used that proves robust under countermeasures and deals efficiently with uncertain, incomplete or poor quality information. The evidential reasoning scheme can yield both single ID with an associated confidence level and more generic propositions of interest to the Commanding Officer. Our approach of reasoning over attributes provided by the imagery will allow the ADFT to process in the next phase (currently under way) both FLIR imagery and SAR imagery in different modes (Spot Adaptive and RDP for naval targets, Strip Map and Spotlight Non-Adaptive for land targets).

The DS theory of evidence offers a powerful approach to manage the uncertainties within the problem of target identity. Every sensor declaration about the M possible "values" of an attribute assigns a Basic Probability Assignment or Mass (BPM) value m_i (i=1...M) to that attribute (present in the database) and generates M propositions which are just the numerical list of platforms in the PDB that can attain the said value for the attribute. For a PDB containing N platforms, the numerical list of platforms which forms a proposition is represented in the current implementation by a string of zeroes and ones in the location of a string of N bits. This is done to speed up calculations by bit manipulations for ensemble operations such as union and intersection, which are needed in DS theory. For physical quantities like speed, length, RCS and image classification attributes like category or class, M is usually greater than 1. This is due either to the fuzzification of the physical quantity or to the inherently complex nature of the algorithmic determination of the attribute (e.g. by NN outputs). DS theory is particularly suited for our application because it requires no a priori information, can resolve conflicts (present in hostile environments due to countermeasures), and can assign a mathematical meaning to ignorance (which is the result of some of the chosen algorithms).

As various evidences are combined over time, DS combination rules will have a tendency to generate more and more propositions which in turn will have to be combined with new input evidences. Since this problem increases exponentially, the number of retained solutions must be limited. Our truncated version of DS theory of evidence performs the conventional combination rules of DS theory but retains the final solution proposition according to the following criteria [2]:

1. All combined propositions which have BPM > BPM_MAX are retained (presently chosen as 0.05).
2. All combined propositions which have BPM < BPM_MIN are eliminated (presently chosen as 0.001).
3. If the number of retained propositions in step 1 is smaller than MAX_NUM, the subroutine will retain, by decreasing BPM, the propositions consisting of one element (singleton) until MAX_NUM is reached. If MAX_NUM is not reached, one retains, by decreasing BPM, the propositions consisting of two elements. The process is repeated until MAX_NUM is reached (presently chosen as 8). This step takes into consideration that the platform's commanding officer favours propositions of the singleton type.

6. Image Support Modules (ISM)

The ISM for either the SAR or the FLIR can also generate a nearly infinite set of declarations from a single given image. Care must be taken to preserve as much independence between the declarations and certainly prevent any conflict. Such an independence can be achieved to a reasonable extent if different features are extracted from the image in different steps or if totally different mathematical algorithms are used in each step. The ISM which LM Canada has designed for image interpretation of SAR data is the 2-D equivalent of the ESM's 1-D signal interpretation. The present ISM design involves the four steps described in Figure 1, of which the first three have been implemented and tested [2]. The design logic shown in Figure 1 involves a hierarchical decision tree for ship features extraction and ship classification which has considerably evolved since the description in [1].

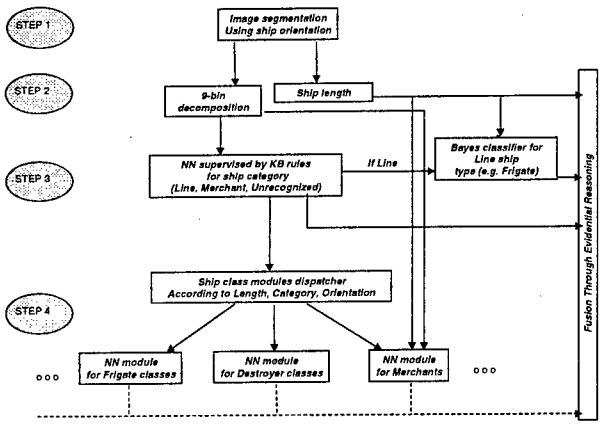


Figure 1 - SAR ISM hierarchical design

The SAR ISM thus preferentially extracts target features at long range feature, namely

1. ship length,
2. ship category: combatant (line), merchant or unrecognized,
3. ship type, e.g. if line, then either frigate, destroyer, cruiser, battleship or aircraft carrier, and
4. ship class, e.g. if frigate, then Halifax class or MacKenzie class.

Given the image acquisition parameters and the navigation data, the first step checks if proper ship orientation is achieved (e.g. the image is sufficiently elongated), and, if so, an image segmentation process detects a target whose image is simply connected. In a second step, a Hough transform then permits an estimation of the ship length, which is immediately sent to MSDF for the ID estimation process.

In the third step, Artificial Intelligence rules based on the relative position and number of main scatterers (as identified by pixel intensities being above a certain threshold) allow the determination of ship category into "line" or "merchant" categories by locating its superstructure. The presently implemented method is a Neural Net (NN) trained on 37 production rules based on the location of the main radar scatterers in 9 different regions along the length of the ship. The possible outputs of the NN are "line", "merchant" or "unrecognized". It should be noted that these categories are only a subset of the NATO STANAG where "line" is only a subset of combatant ships and "merchant" is a subset of so-called "non-naval" entities. They are however the main categories relevant for the Aurora missions mentioned earlier. An "unrecognized" declaration from the NN indicates that it could not reach an ID and consequently that declaration is assigned to the ignorance in the DS algorithm for evidential reasoning.

The third step also performs an attempt at identifying ship class if the NN declaration for "line" is sufficiently large (say >50%). This is due to the correlation [1] between ship length and ship class observed from a survey of about 100 classes of ships in Jane's Fighting Ships (no such correlation exists however for merchant ships). The line types which are generated in this fashion can discriminate between frigates, destroyers, cruisers, battleships and carriers (as identified in the PDB). An indication of the fuzziness of the declaration is given by the relative overlap between classes for a given measured length.

Finally, in the fourth step, specialized NNs trained on subsets of the database of ship images (artificially created from a simulator for various aspect and depression angles), that span a given length interval, refine the ID declaration to ship class (e.g. frigate of Halifax class, destroyer of Spruance class). The outputs of the neural net for each possible class are again numbers between 0 and 1 which are interpreted as the level of confidence in obtaining the correct class ID. The neural net also provides an "unrecognized" class which again reflects its inability to reach a conclusion about ship class. This is then attributed to the ignorance in the DS sense, as in step 3. The merchant ship classifier has been implemented as a 2 hidden layer NN trained on over 200 merchants and tested on a restricted set of simulated SAR imagery.

For the FLIR classifier, a similar two hidden layer neural net design is presently being studied and trained

on more than two hundred merchant ships. Since merchants, unlike combatants, cannot readily be identified through the radar emitters, the FLIR performance will be crucial in determining their type: cargo, RoRo, ferry, oiler/tanker, or passenger. Results will be presented elsewhere, once validation has been quantified on real FLIR imagery.

7. Test Scenarios

Two representative scenarios were run on the ADFT, one for Maritime Air Area Operations (MAAO), the other for Direct Fleet Support (DFS). Three Russian line ships (Mirka II frigate, Udaloy II destroyer and Kara-Azov cruiser) are imaged by the SAR in MAAO as indicated in Figure 2 (the dashed vertical line is the flight pattern of the CP-140) while three American line ships (Coontz destroyer, Ticonderoga cruiser and Virginia cruiser) are imaged in DFS. In MAAO, one has in addition, countermeasures for the Udaloy, namely emitters are detected from the Udaloy which do not correspond to the entries in the PDB.

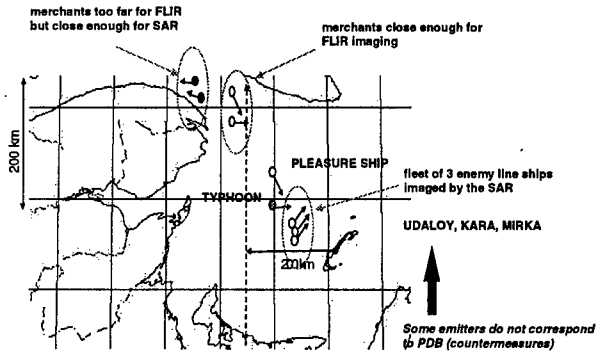


Figure 2 – Maritime Air Area Operations scenario

8. SAR ISM Results

Figure 3 shows the raw SAR imagery in reverse video and histogram equalized (on top), the segmented image with its extracted centerline by the Hough transform and the thresholded major scatterers for the Udaloy destroyer, the Kara cruiser and the Mirka frigate (respectively from left to right). The images are not necessarily to scale. According to the scenario, the SAR acquisition parameters are: an aircraft altitude of 3 km, a range to target of 100 km, an aircraft speed of 0.15 km/sec (300 knots), a SAR wavelength of 0.03 m, common ship heading of 45 degrees, slant-range resolution of 0.75 m and cross-range resolution of 2.0 m (intentionally unclassified numbers).

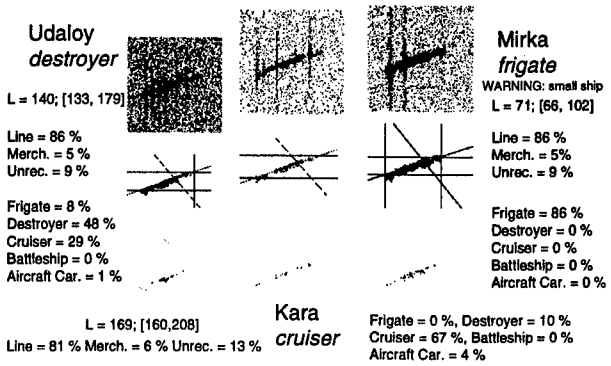


Figure 3 - SAR images of Russian fleet in MAAO scenario with ISM declarations

For each of the 3 imaged ships, the ISM's hierarchical classifier generates successively 3 attributes, each of which leads to several identity declarations (with associated BPMs in the DS sense) for line ships. First the length obtained after centerline detection, which is further fuzzified into bins corresponding to length increments of 40 m (an interval for length is shown in the figure). Next the line category with its confidence level is obtained by keeping the top 10% of the strongest pixels and confidence levels are given for line, merchant and unrecognized categories. Finally the line type, from a choice of 5 line types: frigate, destroyer, cruiser, battleship or aircraft carrier (identification are again in percentages).

Note that all ships are correctly identified by the SAR ISM in the MAAO scenario. The correct ISM declaration for the Udaloy will offset the incorrect ESM reports. In the case of the Mirka, its small length is flagged to the operator since the algorithm is not certain of correct ID. In this case, the operator should fuse the ISM result, but in other scenarios that were run (such as Counter Drug Operations,) the operator should decide against fusion.

The results are quite different in the case of the DFS scenario, which consists of a Canadian and an American fleet. The Canadian fleet of 4 ships is overflown so closely that only FLIR acquisition is possible (to be analyzed at a later date when the FLIR ISM is mature) while the American fleet of 6 ships is sufficiently distant that only SAR image acquisition of a selected subset of ships is possible. In the DFS case, one of the American ships is incorrectly identified by the SAR ISM (the Virginia cruiser is an atypically small cruiser such that the ISM Bayes length classifier identifies it primarily as a destroyer) but the ESM

functions properly. Because the type declaration consists of the 3 most probable types, which includes some confidence in the Virginia being a cruiser, the ISM does not contradict completely accumulated ESM information and the small amount of conflict is correctly handled by the truncated DS scheme. Figure 4 below shows the SAR image for the 3 American ships and the ISM declarations.

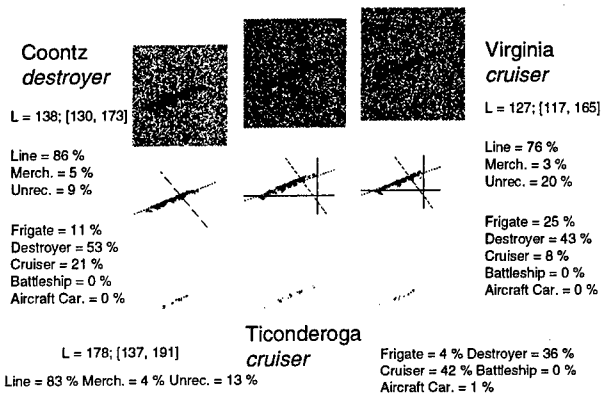


Figure 4 - SAR images of American fleet in DFS scenario with ISM declarations

9. Identification Results

The DF algorithms have been tested on complex scenarios representative of the main Aurora missions, namely MAAO, DFS, Counter Drug Operations and Maritime Sovereignty Patrols. This paper deals only with the first two and concentrates on line ships rather than merchants.

In the MAAO case the emitters carried by the platforms have many common elements as shown in Table 1 below, where the Udaloy's false emitters are listed in bold. Reporting emitters are selected at random in the emitter list of the corresponding platform in the PDB.

Name	List of emitters
Udaloy II	63 65 69 71 **77** 91 93 97 **129**
Kara II	45 46 62 64 68 78 84 85 92 93 103 104
Mirka	44 47 55 56 103 109

Table 1 – Emitter list for the Russian Fleet

Figure 5 shows the ID evolution for the Mirka frigate.

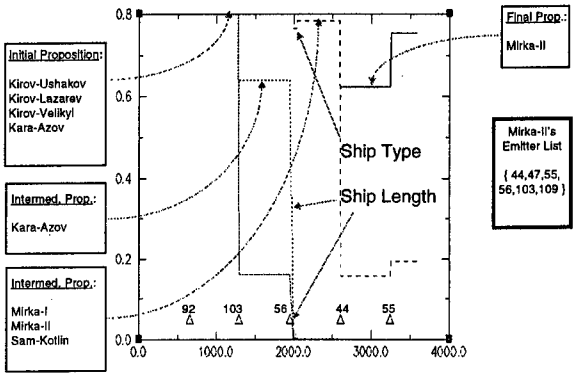


Figure 5 - ID evolution for the Mirka II

Five triangles at the bottom of the figure represent the time at which an ESM report has been fused. After the first 10 minutes (t=656s), the Kara-Azov and the Mirka are not properly resolved (within an angle of 1°). The emitter #92 belonging to the Kara-Azov and other platforms is detected, initiating a proposition in which the Mirka-II is absent. Then, at t=1293s, the emitter #103 is detected which belong to the Mirka-II and to the Kara-Azov (the ground-truth shows that it is emitted by the Mirka-II but the Kara-Azov proposition already existed). At t=1950 s, the emitter #56 is detected which does not belong to the kara-Azov but to the Mirka-II. A SAR image is acquired and analyzed at time t=1980 s. The fusion of the Ship-Length attribute confirms the Mirka ID since the Kara is a cruiser two times longer than the Mirka-II. The fusion of the Ship Type attribute at time t=2040 s provides further reinforcement. Then, at time t=2606 s and 3243 s, two emitters (#44, #55) belonging only to the Mirka-II create the final correct ID.

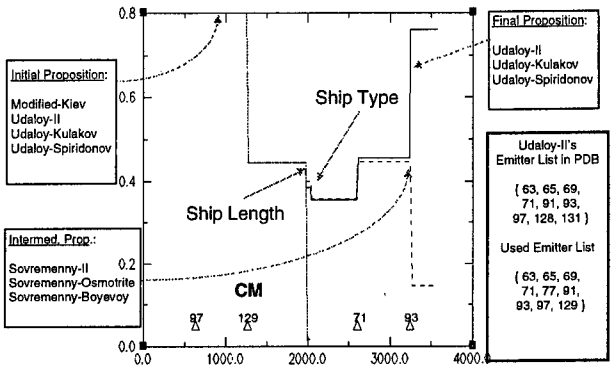


Figure 6 - ID evolution for the Udaloy II

Figure 6 above shows, for the Udaloy-II, the same type of information shown in Figure 5.

At t=637s, the emitter #97 is detected which belong to the ships of class Udaloy and to the modified-Kiev. Then, at t=1275s, emitter #129 is fused which does not normally belong to the Udaloy-II but has been placed intentionally in its list of emitters to simulate a countermeasure; as a result a false ID list is generated. A SAR image is acquired an analysed at t=1980s. The fusion of the Ship-Length attribute split amongst two propositions (see also figure Figure 7) has the effect of decreasing the false ID while creating from the initial proposition an ID containing ships of class Udaloy. The fusion of the Ship Type helps in decreasing the BPM associated to the false identity. At time t=2606 s, emitter #71 is detected which unfortunately does not help in discarding the false ID since this emitter belongs to the ships of both classes Udaloy and Sovremenny. The correct decision is made at time t=3243 s, when emitter #93 belonging only to the ships of the class Udaloy is detected and fused.

Let us now consider an electromagnetically silent version of the same scenario, i.e. one where only the SAR ISM can provide ID estimation.

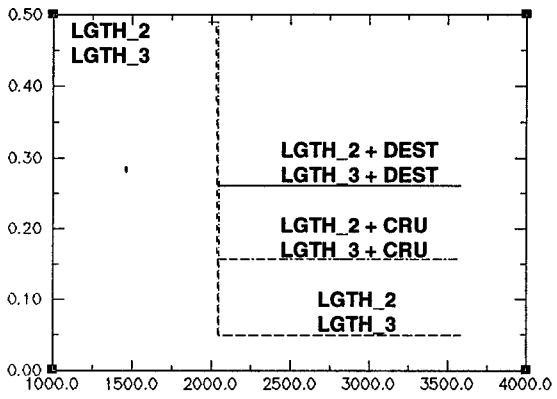


Figure 7 - Time evolution of the generic identification for the Udaloy-II from the SAR ISM

In this case, Figure 7 shows that, at first, the Udaloy length declaration is equally distributed amongst two propositions consisting of the fuzzified length classes LGTH_2 and LGTH_3. These are then fused with the declaration consisting of 3 non-null propositions coming from the Bayes classifier namely: Destroyer 48%; Cruiser 29%; and Frigate 4%. The resulting fusion results in the correct proposition "LGTH_2+ DEST" together with "LGTH_3+DEST" having the largest mass, followed by "LGTH_2+CRU" together with "LGTH_3+CRU" and finally the joint identification "LGTH_2" together with "LGTH_3" by itself. This demonstrates again that the ISM alone can provide an adequate (if more complex) platform identification in an electromagnetically silent environment.

In the case of the DFS scenario with SAR imaging of the American fleet (at the same distance and aspect angles as in the case of the MAAO scenario), one obtains the results for the ID of the American fleet shown in the following figures. In order to follow the ID evolution, Table 2 below shows the emitter list for the American fleet. Note that only the first 3 are imaged, but all ships are so closely separated that their emitters are occasionally associated to other ships depending on the line-of-sight and the ESM bearing accuracy used (representative of a classified number). Again the emitters carried by the platforms have many common elements and emitters are selected at random from run to run.

Name	List of emitters
Coontz	7 8 13 16 18 33 34 35 57
Ticonderoga	7 8 13 32 53 54 57 110 112
Virginia	7 8 13 15 16 31 32 53 54 57
Spruance	8 14 18 31 32 43 53 57 114 115 119 121
Sacramento	7 13 18 33 42 121 130
Nimitz	7 8 16 17 54 57 115 117 121 122 124 125 126 127

Table 2 - – Emitter list for the American Fleet

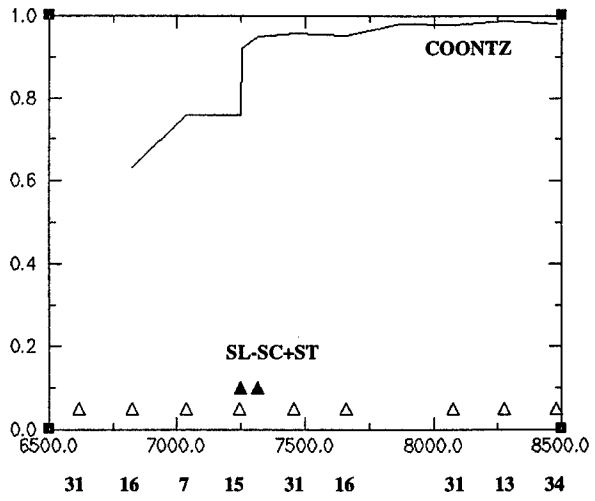


Figure 8 – ID Time evolution for the Coontz

Figure 8 shows that the ESM reports already prefer the Coontz destroyer identification since emitter #16 was identified (common also on the Virginia) and that the Ship Length (SL) (since the Coontz is smalle rthan

the Virginia) and, at a later time, the Ship Category (SC) and Ship Type (ST) only reaffirm this correct ID, despite emitter #15 having been fused at similar times.

The situation for the Ticonderoga is slightly more complicated. Figure 9 shows that, at first, due to association of the closely separated American fleet at such a large distance (relative to unclassified sensor accuracy used), both the Virginia and the Ticonderoga are possible. The simultaneous fusion of emitter #110 (solely on the Ticonderoga), without the help of the SL and SC+ST ISM declarations, resolves the ambiguous ID.

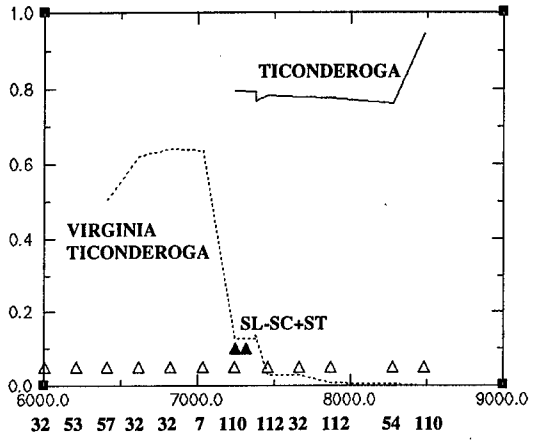


Figure 9 – ID Time evolution for the Ticonderoga

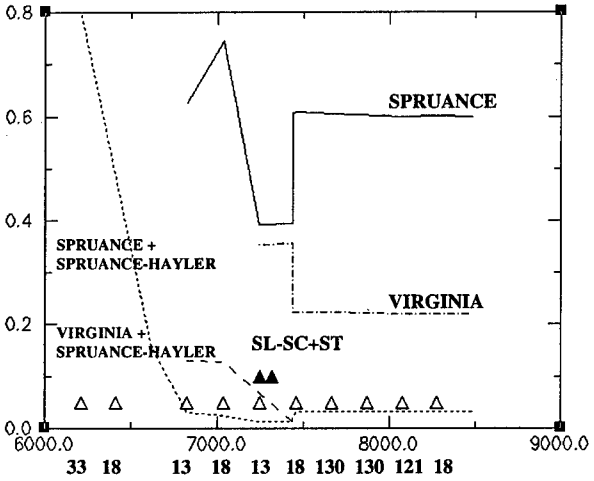


Figure 10 – ID Time evolution for the Virginia

As for the Virginia cruiser (which is incorrectly identified by the SAR ISM as a destroyer like the Spruance), Figure 10 below shows that emitter #13 (present on the Virginia but not the Spruance) decreases the belief in the Spruance and increases the one for the Virginia. However, after fusion of emitter #18 (present on the Spruance but not the Virginia) and the incorrect ISM declaration, rapidly followed by yet another emitter #18 declaration, the final identification favors the Spruance until one is well past the American fleet. In this case, the closeness of the two platforms has caused too many "false" ESM declarations to be associated to the Virginia to reverse the incorrect ISM declaration.

10. Conclusions

A KBS BB-based architecture has been chosen for the airborne fusion testbed at LM Canada. The KBS BB environment allows incremental implementation of any MSDF function in a context-dependent way. It has been tested on many scenarios relevant to missions of the CP-140 Aurora and with the Aurora's non-imaging and imaging sensors. Analysis of SAR imagery proceeds through a hierarchical classifier that extracts long range attributes from Spotlight SAR imagery such as ship length, category, type and class. Image interpretation results (coming from a SAR imagery simulator and CAD models of ships) and platform identification results for Maritime Air Area Operations and Direct Fleet Support scenarios were presented. Correct final ID is achieved in all cases where the targets are sufficiently well separated and most of the time in highly dense environments. Through a proper interpretation of the non-imaging sensor reports and an appropriate understanding of features extracted from images, sensor declarations can be generated which consist of sets of propositions with an associated confidence level. These propositions consist in a list of platforms that realize the attribute "value" and are mathematically treated using a truncated Dempster-Shafer evidential reasoning scheme. A special effort has been made during the generation of the PDB to enumerate all possible attributes that the sensor inputs can provide.

11. References

[1] Jouan A., Gagnon L., Shahbazian E., Valin P., "Fusion of Imagery Attributes with Non Imaging Sensors Reports by Truncated Dempster-Shafer Evidential Reasoning", in Proc. FUSION'98, July 1998, Vol. II, pp. 549-556, and references herein.

[2] Gagnon L., Klepko R., "Hierarchical Classifier Design for Airborne SAR Images of Ships, in Proc SPIE Aerosense'98, Vol. 3371, April 1998.

Comparison of two Integration Methods of Contextual Information in Pixel Fusion

Sophie Fabre
ONERA/DOTA
BP 4025
31055 Toulouse cedex, France
Email : sfabre@onecert.fr

Alain Appriou
ONERA/DTIM
BP 72
92322 Chatillon cedex, France

Xavier Briottet
ONERA/DOTA
BP 4025
31055 Toulouse cedex, France

Abstract: *Pixel fusion is used to elaborate a classification method at pixel level and optimize target detection. It must take into account the more accurate as possible information and take advantage of the statistical learning of the previous measurements acquired by sensors. The classical probabilistic fusion methods lack of performance when the previous learning is not representative of the real sensors measurements. The Dempster-Shafer theory is then introduced to face this disadvantage by integrating a further information which is the context of the sensors acquisitions. In this paper, we propose a formalism of the sensor reliability modelization that leads to two methods of integration when all the hypotheses, associated to objects of the scene acquired by sensors, are previously learnt. Afterwards, we are interested in the evolvement of these two methods in the case where the previous learning is unavailable for an object of the scene and a global method of contextual information integration can be deduced.*

Keywords: pixel fusion, Dempster-Shafer theory, contextual information, degree of trust, mass set.

1. Introduction

During these last years, the number of image sensors has drastically increased. Then a large set of images simultaneously acquired on the same landscape but in different spectral bands is often available. As the information associated to an object depends on the spectral band, the multi-sensors data fusion aims at combining the information from the different spectral bands in order to significantly increase scene perception. Our pixel fusion method is used to elaborate a new classification method at pixel level and also to optimize target detection. It must take into account the more accurate as possible information and take advantage of the statistical learning of previous measurements acquired by sensors. The classical probabilistic fusion methods lack of performance when the previous learning is not representative of the real sensors measurements due to varying environmental conditions for example.

Consequently, we propose a fusion method based on the Dempster-Shafer theory which allows to easily integrate the context of the sensors measurements in order to take the more accurate as possible information.

The fusion methods need the determination of an "a priori" database made up of probability density laws defined for a given context [1][2]. The acquired measurements are related to the surface properties and the context. Furthermore the probability density laws of the acquired measurements can be different from the laws that are previously learnt to construct the database. Some disturbing parameters must be considered in order to justify this difference. These parameters are either atmospheric disturbances or surface variations (as temporal evolution) [3][4]. These disturbing contributions define the context and are called contextual variables.

The sensor reliability to the context depends on the contextual variables values and must be considered by fusion method [5][6]. We propose a formalism modelization of sensor reliability to the context that leads to two methods of integration when all the hypotheses, associated to objects of the scene acquired by sensors, are previously learnt : the first one amounts to integrate this further information in the fusion rule as degrees of trust and the second models it as mass set (§ 2). These two methods are based on the theory of fuzzy events.

Afterwards, we are interested in the evolvement of these two methods in the case where the previous learning is unavailable for a hypothesis associated to an object of the scene and compare these two methods in order to deduce a global method of contextual information integration in the fusion process (§ 3).

2. Fusion and contextual information modelization methods

The multi-sensors system is composed of M sensors S_j (j=1,..., M) that provide measurements L_j. This system is used to recognize an object among N ones. An exclusive hypothesis H_i is associated to every object i.

The frame of discernment E is defined by the hypotheses H_i : $E = \{H_1, H_2, ..., H_N\}$. $\overline{H_i}$ is the complement of H_i in E.

Some disturbing parameters must be considered in order to evaluate the sensor reliability. These parameters define the context and are called contextual variables. A particular context $z = \{z_1, z_2, \ldots, z_P\}$ is then defined by P contextual variable values. Moreover, the vector z^m represents the context measurements : $z^m = \{z_1^m, z_2^m, \ldots, z_P^m\}$.

The fusion method needs the construction of an "a priori" database made up of the probability density laws defined for a given context. The "a priori" probability density $p(L_j / H_i, z^S)$ of every sensor S_j under hypothesis H_i is previously modeled for the standard contextual variable values $z^S = \{z^S_1, z^S_2, \ldots, z^S_P\}$. These variables allow us to consider all reliable sensors measurements. We suppose that the "a priori" probability densities are known for all the hypotheses of the frame of discernment.

The fusion method is based on the Dempster-Shafer theory which allows an easier integration of the context in the decision rule formalism (§ 2.3.2 and 2.3.3). The construction of the basis mass sets $m_{ij}^b(.)$, each representative of the evidence assigned to the frame discernment E, thanks to the measurements provided by the sensor S_j and the learning on H_i, is based on the works led by Appriou [1][2] (§ 2.3.1). These mass sets will be considered afterwards as elementary sources.

The contextual information is modeled in the form of mass sets representative of the sensors availability for the considered context and can be introduced at two levels :

- At the elementary level of each sources : the mass set, noticed $m_{Cij}(.)$, is equivalent to degrees of trust d_{ij} introduced as weakening factors [1][2][7]. It takes into account the validity of each separate sources.
- At the global level of the association of many sources : the mass set $m_C(.)$ amounts to introduce the competitively validity of all the possible associations of sources.

The estimation of these mass sets is based on the fuzzy sets theory.

We propose an original combination rule called CC (Contextual Combination) rule that allows to combine the contextual information with the "a priori" information (§ 2.2). It can be applied to two different levels : before or after fusion operation.

Therefore two different methods of fusion and contextual information integration are proposed and each one uses an unlike representation of the contextual information. The first one uses the set $m_{Cij}(.)$ and is called CDT (Contextual Degree of Trust) method (§ 2.4). It introduces the CC rule before the fusion operation. The second method uses the set $m_C(.)$ and consists in applying the CC rule after the fusion process (§ 2.5). This method is called CMS (Contextual Mass Set) method.

2.1. Notations and definitions

The notations and definitions, used in the next paragraphs, are the following :

- The P-dimensional space where the context is represented is called Z.
- C_{ij} represents the inclusive validity domain or fuzzy subset of contexts for which the assessment of the hypothesis H_i provided by the sensor S_j is valid ($C_{ij} \subset Z$), without knowledge on the validity of any other sensor for any hypothesis and the validity of the sensor S_j for all the hypotheses H_k different from H_i (Figure 1).
- The index V represents a subset of indexes $\{ij\}$ included in the set stemming from the Cartesian product $\{1,\ldots,N\} \times \{1,\ldots,M\}$:
  $V \subseteq \{1,\ldots,N\} \times \{1,\ldots,M\}$
- c_V is the exclusive validity domain or subset of contexts for which every sensor S_j of the association set represented by V is valid for the discrimination of the hypothesis H_i ($\{ij\} \in V$), and all the others associations of sensor and hypothesis no represented in V are excluded :

$$c_V = \bigcap_{ij \in V} C_{ij} \bigcap_{ij \in \overline{V}} \overline{C}_{ij} \qquad (2.1)$$

with $V \subseteq \{1,\ldots,N\} \times \{1,\ldots,M\}$

and

$$c_\varnothing = \bigcap_{ij \in \{1,\ldots,N\} \times \{1,\ldots,M\}} \overline{C}_{ij} \qquad (2.2)$$

The Figure 1 illustrates the notions of exclusive and inclusive validity domains for the case of two sensors and two hypotheses. The validity domain C_{11} is the subset of contexts for which the sensor S_1 allows to discriminate the hypothesis H_1 without knowledge on the validity of any other sensor for any hypothesis : S_2 and H_1, S_2 and H_2; and the validity of S_1 for H_2. The validity domain $c_{\{11\}}$ is the subset of contexts where the only valid association corresponds to the hypothesis H_1 and the sensor S_1.

According to the equation (2.1), this domain is expressed as :

$$c_{\{11\}} = C_{11} \cap \overline{C}_{12} \cap \overline{C}_{21} \cap \overline{C}_{22}$$

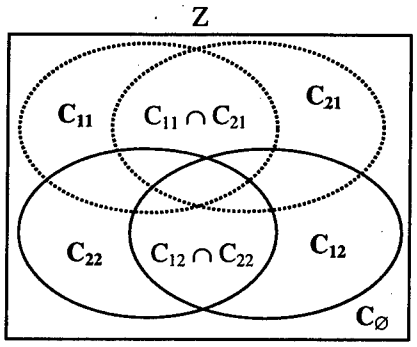


Figure 1 : Representation of the context space Z

2.2. CC rule

We aim at finding a global mass set $m(.)$ on the frame of discernment E, set of the hypotheses of discrimination H_i (i=1,..., N). This global set $m(.)$ is obtained according to the mass sets $m_V(.)$ provided on E by sources or combination of sources represented by V and a bayesian mass set $m_C(.)$ on $E_{SH}=\{c_V\}$ by using the suitable operations of conditioning, coarsening and refining [7][9].

The term c_V represents :

- The inclusive validity domain, with V = {ij} and {ij}∈{1,..., N}×{1,..., M},
- The exclusive validity domain, with V⊆{1,...,N}×{1,..., M}.

According to the demonstration led by Fabre [7], the global mass set $m(.)$ is explained as :

$$m(A) = m_C(c_\varnothing)\cdot m_\varnothing(A) + \sum_{\substack{V \\ V\neq\varnothing}} m_C(c_V)\cdot m_V(A) \qquad (2.3)$$

2.3. Fusion method and decision rule

2.3.1. Expression of the basic mass set

Appriou suggests an approach that consists in introducing each "a priori" probability density $p(L_j/H_i, z^S)$ among an appropriate mass set $m_{ij}^b(.)$ [1][2]. Two models of mass set are then defined by an axiomatic approach on the frame of discernment $\{H_i, \overline{H}_i, E\}$. We select among these models the less specific one. This mass set, called "basic mass set", is explained by considering all the degrees of trust d_{ij} equal to 1 [1][2] :

$$m_{ij}^b(H_i) = 0 \qquad (2.4)$$
$$m_{ij}^b(\overline{H}_i) = 1 - R_j \cdot p(L_j/H_i, z^S)$$
$$m_{ij}^b(E) = R_j \cdot p(L_j/H_i, z^S)$$

where the term R_j represents a normalization factor defined as :

$$R_j \in \left[0, \left\{\max_{i\in[1,N]}\left[\sup_{L_j}\left(p(L_j/H_i, z^S)\right)\right]\right\}^{-1}\right] \qquad (2.5)$$

2.3.2. Fusion rule

The global mass set $m(.)$ results from the combination of the M mass sets $m_j(.)$ associated to the sensor S_j. The combination rule is the orthogonal Dempster-Shafer's rule [9] :

$$m(.) = \bigoplus_{j\in[1,M]} m_j(.) \qquad (2.6)$$

The mass set $m_j(.)$ is obtained by the combination of the N elementary mass sets $m_{ij}(.)$:

$$m_j(.) = \bigoplus_{i\in[1,N]} m_{ij}(.) \qquad (2.7)$$

The mass set $m_{ij}(.)$ represents the "a priori" information, modeled by the basis mass set $m_{ij}^b(.)$, and the contextual information at once. When the contextual information is unavailable, the mass set $m_{ij}(.)$ is similar to the set $m_{ij}^b(.)$. The focal elements associated to $m_{ij}(.)$ are H_i, $\overline{H}_i$ and E.

2.3.3. Decision rule

The objective is to choose one decision d_i among a finite set D of Q possible decisions owing to the assessment provided by the mass set $m(.)$ on E. The decision d_i corresponds to the assignment of the observation $\vec{L}$ to the set F_i made up by one or many hypotheses of E. The choice of taking a decision d_i, when the observation $\vec{L}$ belongs to H_k, generates a cost $\lambda(d_i/H_k)$.

The more consensual decision is provided by minimizing the risk function $R(d_i/\vec{L})$ on the set of all the possible decisions [8] :

$$R(d_i/\vec{L}) = \sum_{B\subseteq E}\left[m(B)\cdot \min_{H_k\in B}\{\lambda(d_i/H_k)\}\right] \qquad (2.8)$$

The costs $\lambda(d_i / H_k)$ are defined for $i \in [1, Q]$ and $k \in [1, N]$ by considering the two following propositions :

- The cost to declare that the observation $\bar{L}$ is assigned to the set F_i of hypotheses associated to d_i when it really belongs to the class H_k , is maximum :

$$\lambda(d_i / H_k) = 1 \quad if \quad H_k \notin F_i \qquad (2.9)$$

- The cost to make a good decision is such that :

$$\lambda(d_i / H_k) = \lambda_i \quad if \quad H_k \in F_i \qquad (2.10)$$

By integrating the expressions (2.9), (2.10) in the relation (2.8) and using the definition of the plausibility [9], the risk function becomes :

$$R(d_i / \bar{L}) = 1 - [1 - \lambda_i] \cdot Pl\left(\bigcup_{H_k \in F_i} H_k \right) \qquad (2.11)$$

Consequently, the decision rule is obtained by minimizing the risk function (Equation (2.11)) and can be explained as follows :

$$\max_{i \in [1,Q]} \left\{ [1 - \lambda_i] \cdot Pl\left(\bigcup_{H_k \in F_i} H_k \right) \right\} \qquad (2.12)$$

In this particular approach, only singletons of hypothesis are taken into account. Therefore, the set D is composed of as many decisions d_i as hypotheses H_i in the frame of discernment E. Moreover, we consider that the cost λ_i for a good classification is equal to 0. Consequently the most likely hypothesis has to justify a maximum plausibility criterion. The decision rule (Equation (2.12)) is rewritten as :

$$\max_{i \in [1,N]} \{ Pl(H_i) \} \qquad (2.13)$$

This decision rule is coherent with the criterion introduced in the works led by Appriou [1][2].

As the hypotheses H_i are singletons of E, the expressions of plausibility and communality are the same [1][2][9]. Consequently, the plausibility is explained as follows :

$$Pl(H_i) = \prod_{j \in [1,M]} \{ Pl_j(H_i) \} \qquad (2.14)$$

By using the plausibility definition [9], the mass sets $m_j(.)$ (Equation (2.7)) and $m_{ij}(.)$ (Equation (2.4)), the plausibility $Pl_j(.)$ can be explained as follows :

$$Pl_j(H_i) = K_{fj} \cdot \frac{m_{ij}(H_i) + m_{ij}(E)}{m_{ij}(\overline{H}_i) + m_{ij}(E)} \qquad (2.15)$$

The term K_{fj} is a normalization factor independent of the hypothesis H_i.

2.4. CDT Method

In this case, the sensor reliability is represented by the mass sets $m_{Cij}(.)$ (§ 2.4.1). These mass sets are combined by the CC rule with the mass sets, associated to the previous learning and obtained owing to the basis mass set $m_{ij}^b(.)$, in order to obtain the elementary mass set $m_{ij}(.)$ (§ 2.4.3). The operation of fusion is then applied on these elementary mass sets in order to obtain a global mass set $m(.)$ introduced in the decision rule (§ 2.4.4). The architecture of the CDT method is described on Figure 2.

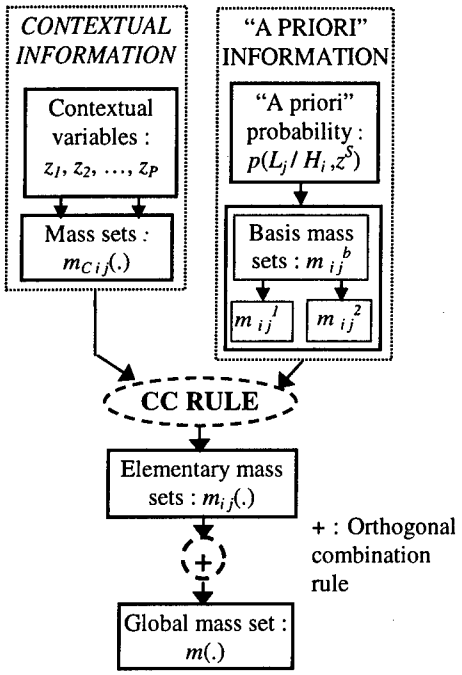


Figure 2 : CDT method architecture

2.4.1. Contextual information representation

The reliability of the source $\{ij\}$, defined by the association of the sensor S_j and the hypothesis H_i, to the context is represented by a bayesian mass set $m_{Cij}(.)$ established on the frame of discernment $E_{Cij} = \{C_{ij}, \overline{C}_{ij}\}$. The set C_{ij} is defined in § 2.1.

The estimation of the mass set $m_{Cij}(.)$ is performed in several stages :

- *Stage 1* : The probability of the context

Let $z=\{z_1,\ldots,z_P\}$ be as a random vector of probability density $p(z/z^m)$ where $z^m=\{z_1^m,\ldots, z_P^m\}$ is the vector

associated to the variable z measurements. We have to model this probability density law $p(z/z^m)$.

- *Stage 2* : Validity domain of every sensor

The fuzzy sets theory is used to define the validity domain of every sensor.

$\mu_{iju}(z_u)$ is the elementary fuzzy membership function associated to the contextual variable z_u, the sensors S_j and the hypothesis H_i.

The fuzzy membership function $\mu_{ij}(z)$ characterizes the sensor availability for the context z and is expressed with the elementary fuzzy functions :

$$\mu_{ij}(z) = \bigwedge_{u\in[1,P]} \mu_{iju}(z_u) = \mu_{ij1}(z_1) \wedge \mu_{ij2}(z_2) \wedge ... \wedge \mu_{ijP}(z_P) \quad (2.16)$$

The operator $\wedge$ represents the operator of minimum conjunction [5][6].

- *Stage 3* : Probability of sensor validity

$P(S_j/H_i, z^m)$ is the probability that the sensor S_j is reliable according to the context value z^m and the hypothesis H_i.

This probability is explained by using the definition of the probability measures of fuzzy events [10][11]:

$$P(S_j/H_i, z^m) = \int \mu_{ij}(z) \cdot p(z/z^m) \cdot dz \quad (2.17)$$

When the variable value is certain, the probability density $p(z/z^m)$ is replaced by the Dirac function $\delta(z-z^m)$ and the equation (2.17) becomes :

$$P(S_j/H_i, z^m) = \mu_{ij}(z^m) \quad (2.18)$$

- *Stage 4*: Mass set $m_{Cij}(.)$

The probability (2.17) can be explained as a bayesian mass set $m_{Cij}(.)$ such that :

$$m_{C_{ij}}(C_{ij}) = P(S_j/H_i, z^m) \quad (2.19)$$
$$m_{C_{ij}}(\overline{C}_{ij}) = 1 - P(S_j/H_i, z^m)$$
$$m_{C_{ij}}(C_{ij} \cup \overline{C}_{ij}) = 0$$

2.4.2. "A priori" information representation

Two mass sets $m_{ij}^{w}(.)$ (w = 1,2) are introduced to model the "a priori" information : one mass set uses the measurements as if there were completely reliable (w = 1) and the other is representative of the total uncertainty (w = 2). These mass sets are defined $\{H_i, \overline{H}_i, E\}$ owing to the basis mass set $m_{ij}^{b}(.)$ (Equation (2.4)) :

$$m_{ij}^{1}(.) = m_{ij}^{b}(.) \quad (2.20)$$
$$m_{ij}^{2}(E) = 1$$

2.4.3. Combination of the mass sets

The expression of the elementary mass sets $m_{ij}(.)$ is provided by applying the CC rule (Equation (2.3)) on the mass sets $m_V(.) = m_{ij}^{W}(.)$ (w = 1, 2) and on the bayesian mass set $m_C(.) = m_{Cij}(.)$.

This expression $m_{ij}(.)$ is the similar to the one resulting from a weakening operation applied on the basis mass set $m_{ij}^{b}(.)$ in the case where the weakening factor d_{ij} is such that:

$$d_{ij} = P(S_j/H_i, z^m) \quad (2.21)$$

The CDT method is then the same as the method improved by Appriou based on the introduction of degrees of trust as weakening terms [1][2][9]. Consequently, the elementary mass set $m_{ij}(.)$ can be explained as :

$$m_{ij}(H_i) = 0$$
$$m_{ij}(\overline{H}_i) = d_{ij} \cdot \left[1 - R_j \cdot p(L_j/H_i, z^s)\right] \quad (2.22)$$
$$m_{ij}(E) = 1 - d_{ij} + d_{ij} \cdot R_j \cdot p(L_j/H_i, z^s)$$

Notes : When the degrees of trust d_{ij} are equal to 1, the Dempster-Shafer theory represents the probabilistic approach of the maximum likelihood which supposes that the probability density $p(L_j/H_i, z^S)$ is perfectly representative of real probability density.

2.4.4. Fusion and decision rule

The elementary mass sets $m_{ij}(.)$ are fused according to the fusion rule (Equations (2.6) and (2.7)) in order to deduce a global mass set $m(.)$.

By using the equations (2.14), (2.15) and (2.22), the decision rule (2.13) becomes :

$$\max_{i\in[1,N]} \left\{ \prod_{j=1}^{M} [1 - d_{ij} + d_{ij} \cdot R_j \cdot p(L_j/H_i, z^S)] \right\} \quad (2.23)$$

2.5. CMS Method

In a first time, the expression of the mass set $m_C(.)$ related to the contextual information is established. In a second time, the mass set $m_V(.)$, representative of the weight of evidence assigned to an association of sources and obtained by the fusion of these sources, is explained. Lastly, the combination of these mass sets

$m_C(.)$ and $m_V(.)$ is realized in order to obtain the global mass set $m(.)$ introduced in the decision rule. The principle of the CMS method is described on Figure 3.

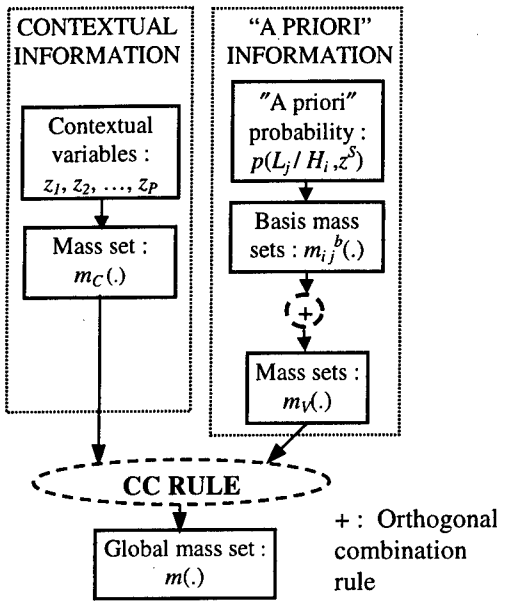


Figure 3 : CMS method architecture

2.5.1. Contextual information representation

The construction of the mass set representing the reliability of one or many associations of sensor and hypothesis is performed in several stages :

- *Stage 1* : Validity domain of every source

The fuzzy membership function allows to define this validity domain. The fuzzy function is defined as in the stage 2 of the § 2.4.1. Consequently $\mu_{ij}(z)$ is explained according to the equation (2.16).

- *Stage 2* : Probability of validity of one or many sources associations

It is the probability of conjunction of the fuzzy subsets corresponding to each source for given contextual variable values z^m [10] :

$$P\Big(\bigcap_{\substack{rq\in V \\ V\subseteq\{1,\ldots,N\}\times\{1,\ldots,M\}}} C_{rq} / z^m\Big) = \int_Z \Big[\bigwedge_{rq\in V} \mu_{rq}(z)\Big]\cdot p(z / z^m)\cdot dz \qquad (2.24)$$

It is to note that when only one association of sensor S_j and hypothesis H_i is considered, the probability $P(C_{ij} / z^m)$ is then explained as the probability $P(S_j / H_i, z^m)$ (Equation (2.17)).

- *Stage 3* : Expression of the mass set m_C

The probability of validity of one or many sources associations (Equation (2.24)) is used to explain the exclusive probability of validity of one or many sources associations $P(c_V)$ [5][6]. $P(c_V)$ can be explained as a bayesian mass set $m_C(.)$ constructed on the set $\{c_V\}$ ($V \subseteq \{1,\ldots,N\}\times\{1,\ldots,M\}$).

Then, the mass set expression $m_C(.)$ may be defined as follows [5][6] :

$$m_C(c_V) = P(c_V) = \sum_{\substack{W\subseteq\{1,\ldots,N\}\times\{1,\ldots,M\} \\ V\subseteq W \\ V\neq\{\varnothing\}}} \Big[(-1)^{|W-V|}\cdot P\Big(\bigcap_{rq\in W} C_{rq} / z^m\Big)\Big] \qquad (2.25)$$

$$m_C(c_\varnothing) = P(c_\varnothing) = P\Big(\bigcap_{ij\in\{1,\ldots,N\}\times\{1,\ldots,M\}} \overline{C}_{ij} / z^m\Big)$$

$|W-V|$ represents the cardinal of the subset $W-V$.

2.5.2. "A priori" information representation

A mass set $m_V(.)$ is constructed on the frame of discernment E and supposed that all the associations of sources, represented by the subset V of indexes $\{ij\}$, are valid. These mass sets result from the orthogonal sum of the basis mass $m^b_{ij}(.)$ (Equation (2.4)) where $\{ij\} \in V$:

$$m_V(.) = \bigoplus_{\substack{\{ij\}\in V \\ V\subseteq\{1,\ldots,N\}\times\{1,\ldots,M\}}} \Big[m^b_{ij}(.)\Big] \qquad (2.26)$$

2.5.3. Expression of the global mass set

In this case, the CC rule (Equation (2.3)) is applied on the mass sets $m_V(.)$ on E and the bayesian mass set $m_C(.)$ on $\{c_V\}$ in order to obtain the global mass set $m(.)$ on E. The global mass set $m(.)$ is explained as :

$$m(A) = m_c(c_\varnothing)\cdot m_\varnothing(A) + \sum_{\substack{V\subseteq\{1,\ldots,N\}\times\{1,\ldots,M\} \\ V\neq\varnothing}} m_c(c_V)\cdot m_V(A) \qquad (2.27)$$

$m_\varnothing(A) = 0$ if $A \neq E$ and $m_\varnothing(A) = 1$ if $A = E$

with $A \subset E$

The decision rule is given by the maximum of plausibility explained with the mass $m(.)$ (Equation (2.13)).

3. Evolution of the CDT and CMS methods for a non-learnt class

The principle is the same as the one described in the § 2. The only change arises from the classes used to construct the frame of discernment. In fact, we admit the existence of a further hypothesis for which the previous learning is not available. Consequently, the "a priori" probability associated to this further hypothesis is unknown. The introduction of this further class is quite legitimate. In fact, a class like backgrounds can be made up of several objects and the previous learning of one or many objects of this class can be unavailable. Consequently, the initial class is divided in two classes : the first one amalgamates the objects for which the previous learning is known and the second one regroups all the objects for which the "a priori" probability is unavailable.

The new frame of discernment E^A consists then of the N hypotheses of the frame E and a further hypothesis H_{N+1} allocated to the non-learnt class : $E^A = \{H_1, \ldots, H_{N+1}\}$. The frame E^A is then refined in comparison with the frame E.

The expression of the basis mass sets related to the "a priori" information is established according to the work led in the § 2.3.1. For the hypothesis H_{N+1} associated to the non-learnt class, the corresponding mass sets are constructed by considering the fact that no information is available. Consequently, all the mass of evidence is assigned to E^A.

The fusion rule is based on the orthogonal combination rule of Dempster-Shafer introduced in the § 2.3.2. Moreover, the decision rule is the same as the one introduced in the § 2.3.3 (Equation (2.12)). In this case, it is benefit to choose the costs λ_i associated to a good decision different from 0, contrary to the values introduced in the § 2.3.2, in order to integrate the fact that no information is available on H_{N+1} and only on it.

In the case where a non-learnt class is added to the frame of discernment, the CDT and CMS methods become respectively the refined CDT and CMS methods. We have shown that the refined CDT and CMS methods may be considered as two implementations of the same global method called "Global Refined Method" [7].

In this general method, the contextual information is taken into account as a mass set in order to realize a fusion process based on the CC rule. This mass can be explained by two different ways :

- The mass set $m_C^d(.)$ depends on the probabilities of validity of each source and is explained owing to degrees of trust on the frame of discernment $\{c_V\}$ [7]:

$$m_c^d(c_V) = \prod_{\{ij\} \in V} d_{ij} \cdot \prod_{\substack{i \notin V \\ or \\ j \notin V}} (1 - d_{ij}) \qquad (3.1)$$

$$\text{with } V \subseteq \{1, \ldots, M\} \times \{1, \ldots, N+1\}$$

$$m_c^d(c_\varnothing) = \prod_{\{ij\} \in \{1,\ldots,N+1\} \times \{1,\ldots,M\}} (1 - d_{ij})$$

- The mass set $m_C(.)$ is directly explained by the probabilities of validity of the different combinations of sources. The construction of this set is inspired by the process described in the § 2.5.1.

These mass sets are combined with the mass sets $m_V(.)$, stemming from the fusion of the basis mass sets $m_{ij}^b(.)$, according to the CC rule (Equation (2.3)). The global mass set deduced from this operation is introduced in the expression of the plausibility in order to explain the decision rule.

We deal with the problem of one non-learnt class only. However the generalization to the case where several classes are not previously learnt is evident.

4. Conclusions

Pixel fusion aims at combining the images from several sensors in order to increase scene perception. The Dempster-Shafer's theory is used to realize pixel fusion and needs a minimum of "a priori" knowledge like previous learning of measurements provided by sensors.

Moreover, the Dempster-Shafer's theory allows to integrate further information such as the sensors reliability to the context. Their reliability depends on the context modeled by the contextual variables.

In the case where all the hypotheses introduced in the frame of discernment are learnt, two methods, that integrate the sensors reliability at different levels, are developed : the CDT and CMS methods.

In the CDT method, the mass sets stemming from contextual information and previous learning are combined before the fusion operation. Practically, this leads to elaborate a degree of trust assigned to each source corresponding to the association of a sensor and a hypothesis.

In the CMS method, a mass set integrates the validity of the different associations of these sources. Consequently in the later, the sensor combination uses all the possible associations of one (or many) sensor(s) and hypothesis. The combination of mass sets representative of "a priori" and contextual information is realized after the fusion operation. For these two methods, the ponderation terms are constructed owing to the theory of fuzzy events. These methods are

described in the case where the validity domain of each sensor depends on the hypothesis. However, these methods are still valid in the case where the validity domain of each sensor is not related to the hypothesis and depends only on the sensor properties since that the combination is applied to the mass sets associated to sensors.

In the case where a non-learnt hypothesis is added to the frame of discernment, the two methods of contextual information integration and fusion evolve. By comparison of these two refined methods, we deduce a global fusion method based on the integration of sensors reliability as a mass set where this one can be explained by two different ways. For the first method, the mass set related to the sensors reliability is expressed as a function of degrees of trust. The global refined method can be only used in the case where the validity domain of each sensor depend on the hypothesis.

The CDT and CMS methods have been successfully implemented for several typical cases and have provided encouraging results.

In the next future, we will work in order to compare the CDT and CMS methods and define their respective validity domains.

Moreover, we will compare the two ways of mass set expression, used by the global refined method, in order to obtain their respective validity domains. This global method can be extended to other expressions of the mass set that is representative of sensors reliability. In particular, the degrees of trust can be computed by different ways.

5. References

[1] A. Appriou, "Probabilités et incertitude en fusion de données multi-senseurs", Revue Scientifique et Technique de la Défense, 11 : 27-40, 1991.

[2] A. Appriou, "Uncertain data aggregation in classification and tracking processes", In "Aggregation and Fusion of Imperfect Information", B. Bouchon-Meunier, ed., Physica Verlag, 1998.

[3] S. Fabre, A. Appriou, X. Briottet, P. Marthon, "Introduction of contextual information in pixel fusion", FUSION'98, Las Vegas, Nevada, USA, July 6-9, 1998.

[4] S. Fabre, A. Appriou, X. Briottet, P. Marthon, "Pixel fusion : Contribution of contextual physical data for the a priori database construction", PSIP'99, Versailles, 17-21 January, 1999.

[5] V. Nimier, "Introduction d'informations contextuelles dans des algorithmes de fusion multicapteur", Journal Traitement du Signal, numéro spécial fusion, Vol. 14, n° 5, 1998.

[6] V. Nimier, "Supervised Multisensor Tracking Algorithm by Context Analysis", FUSION 98, Las Vegas, Nevada, USA, July 6-9, 1998.

[7] S. Fabre, A. Appriou, X. Briottet, "Sensor fusion integrating contextual information", International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, B. Bouchon-Meunier, Ed. World Scientific (submitted the 2 February 1999).

[8] T. Denoeux, "Application du modèle des croyances transférables en reconnaissance des formes", Traitement du Signal, Vol. 24 n° 5 Spécial Fusion (1997) pp. 443-451.

[9] G. Shafer., "A mathematical Theory of evidence", Princeton University Press, 1976.

[10] P. Smets, "Probability of fuzzy event : An axiomatic approach", Fuzzy Sets and Systems 7 (1982) 153-164.

[11] L. A. Zadeh, "Probability Measures of Fuzzy Event", Journal of Mathematical Analysis and Applications, Vol. 23 (1968) pp. 421-427.

Session RA2
Fusion Architecture and Management II
Chair: Ivan Kadar
Interlink & Northrup Grumman Corp. NY, USA

A Study on CORBA-based Distributed Observation System of Seismic Electromagnetic Wave

Xuejun TIAN, Hiroshi YASUKAWA, Tetsuo IDEGUCHI,
Takashi OKUDA, Seiji ADACHI and Masayasu HATA
Faculty of Information Science and Technology
Aichi Prefectural University, Aichi, JAPAN

Abstract

In research of earthquake, hypothesis testing or phenomenon detection is an iterative, successive-refinement process. To verify the relation between an earthquake and detected anomalies, data and applications of other information sources are needed. Not a few earthquake activity observation systems have been set up in Japan independently by different organizations, such as universities, institutes, the earthquake observation association and the Meteorological Agency. These systems, which are managed and maintained by these organizations, contain heterogeneous computing systems. With the CORBA(Common Object Request Broker Architecture) technology which facilitates communication between local and remote objects in a heterogeneous computing environment, it is possible to set up an integrated system like an expert system, for sharing informations and research results in distributed computing systems on high level.

In this paper, an integrating earthquake observation computing system, using CORBA for data exchange, analysis, and information providing, is presented. This system provides information of comparatively high probability about activity of crust, earthquake and other environment changes on the earth.

With this system, we can facilitate communications between local and remote objects, and share applications and data in a distributed computing system conveniently without awaring the low-level infrastructure concerns.

Keywords: CORBA(Common Object Request Broker Architecture), Distributed computing system, Enviromental ElectroMegnetic Wave (EEMW), SEMR(Seismogenic Electromagnetic Radiation)

1 Introduction

Seismo-EM net began thirteen years ago for observing Seismic Electromagnetic Radiation(SEMR) at ELF band[1]. Now, the number of the observation stations, which are consist of observation parts and computing systems, are more than forty over Japan.

Electromagnetic(EM) radiation resulting from local crust activities is an important principle of observation. When enormous energy stored in the crust is released, it is reasonable to think not only mechanical vibration but also electromagnetic wave, light etc, are radiated. Observing SEMR is proved to be a valid method for investigating earthquake, but the it is influenced by other factors, because the EEMW existing in nature is from many different sources. SEMR needs to be extracted from the EEMW. So, It is necessary and valid to develop a system to facilitate the use of multi information[2],[3].

Accompanying increase of observation station and new application development, how to share the data and applications distributed over those stations is becoming serious. To alleviate this problem, CORBA which facilitates communication between local and remote objects for heterogeneous computing systems, adopted by OMG(the Object Management Group) will give us a valid method, which facilitates to develop computer system to use multi information efficiently[4],[5]. In this paper, we will discuss a example that CORBA-based Distributed Earthquake Obser-

vation System.

2 Observation of SEMR

Our system Seismo-EM Net is a system of SEMR observation and data process for research of earthquake and crust activities by observing EEMW.

There are about 40 observing stations set where earthquakes often occur over Japan. EEMW at ELF(Extremely Low Frequency) band are observed and the untreated data are kept in station computer system. Those data are then sent to center to be processed.

For SEMR observation, a problem is to distinguish the local anomalies from that of global scale anomalies. A factor that leads to global scale anomalies is enfluence of the sun. Since such change of global scale is almost same in certain area, the influence can be removed with comparing the data of different stations.

The local anomalies is influenced by many local event, such as thunder, artificial noise. To removed such influences, the local information need to be collected and confirmed though some factor can be distincted by signal pattern.

The earthquake is a complicated natural phenomenon, which is related to series physical and chemical changes. As precursor of earthquake, some patterns of SEMR were confirmed by some successful case, but the probability is not satisfied. In some cases, anomalies were detected but there was no earthquake occurred. In the other hand, the patterns of SEMR of earthquake swarm and local earthquake is different, and the SEMR of same earthquake is found different at different observing stations. Considering such phenomenon, anomalous phenomenon resulting from earthquake are related to local situation of stratum, geographical environment and other factors. It is necessary to refer the data obtained in different fields to confirm the pattern of electromagnetic anomalies of a certain natural phenomenon, which is important and valid to draw out SEMR from EEMW.

3 Description of system

CORBA-based Seismo-EM Net is a try to use multi information efficiently to research earthquake, which is based Seismo-EM Net and CORBA technology.

3.1 Overview of CORBA

CORBA is a middle standard that is based on the concept of the Object Request Broker , currently serving as the basis for application in a vide area such as in telecommunications, finance, and manufacturing.

CORBA defines the followings,

1. IDL(Interface Definition Language), which is used for defining the common objects over CORBA.

2. Language Mapping, which determine how IDL features are mapped to applications that are developed in program language such as c, c++, smalltalk.

3. Interface and Services for creating and requesting objects.

4. Protocol, which are used for communication between Objects of CORBA.

As shown in Fig.1, it behaves as a mediator between clients and application objects, which arrange for those objects to access each other across networks at run times.

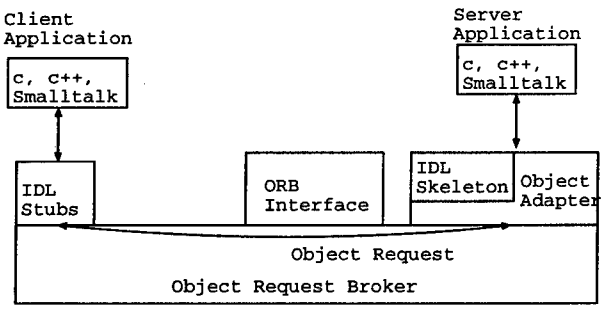


Figure 1: Common object request broker architecture

Clients can create object proxy of the server object in local address space, and operate on the proxy object to change or get the state of objects on the server. The low communication part can be hidden to developer.

CORBA give a platform for communication between objects of distributed object computing system and sharing applications distributed over heterogenous subsystems.

3.2 Available information sources

Data of other fields can be referred in two ways. When anomalous electromagnetic wave is detected the local information of observing station is used to distinct if it is SEMR. In the other hand, when nonseismogenic events occur, data of electromagnetic wave are checked for obtaining the pattern of a certain event.

1. Hi-net(High Sensitivity Seismograph Network, belongs to NIED, National Research Institute for Earth Science and Disaster Prevention Science and Technology Agency.) This system gives information of earthquake forecast and analysis results of oscillation of occurred earthquake.

2. Research Center for Earthquake Prediction(belongs to Kyodo University Earthquake Center).

Various information related earthquake is obtained and offered in this center as follows:

*Three-dimensional global and regional structures of the Earth are investigated using measurements of the travel time, dispersion and attenuation of body and surface waves.

*Investigations of the slab penetration, mantle convection, the driving force of plate tectonics, the chemical and mineral compositions of the crust and mantle, physical properties of the earth's interior including anisotropy by the analysis of seismological and tectonic data.

*Monitoring of crust movements for earthquake prediction with the use of modern space technology of the GPS in addition to conventional techniques such as extensometers and water-tube tiltmeters.

*Studies of active faults for earthquake prediction and hazard mitigation.

3. National Network of Earthquake Data(belongs to Earthquake Information Center). Earthquake records are stored in this system in detail.

4. Earthquake and Tsunami Watching note(belongs to the Meteorological Agency). From this note, information about Tsunami and weather can be obtained.

5. Land Surveying Center note(belongs to the Geographical Survey Institute).

This note offers the displacement data of 1 Month or 1 year for 10 districts and the whole country can be shown. Distance change for any pair of observation stations are also offered as data or shown by map.

3.3 System structure

As shown in Fig.2, the system includes two parts. One part is Seismo-EM Net, which consists of observing stations, data process center(NIT, Nagoya Institute of Technology), simulation(and secondary server) center(APU, Aichi Prefectural University). The observing stations are connected with CORBA. Observed data of EEMW are saved in observing stations with in a given period. As a CORBA object, observing station system can process the date of station as requested, and reply the results. For observing station system, request can be from process center or other station system of Seismo-EM net.

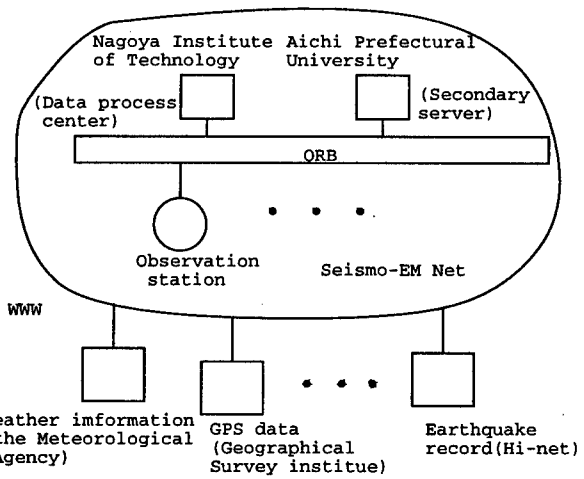


Figure 2: Structure of system

The other part of this system is data gathering, which offers related information out of Seismo-EM net. In the case of WWW, a data gathering object is needed.

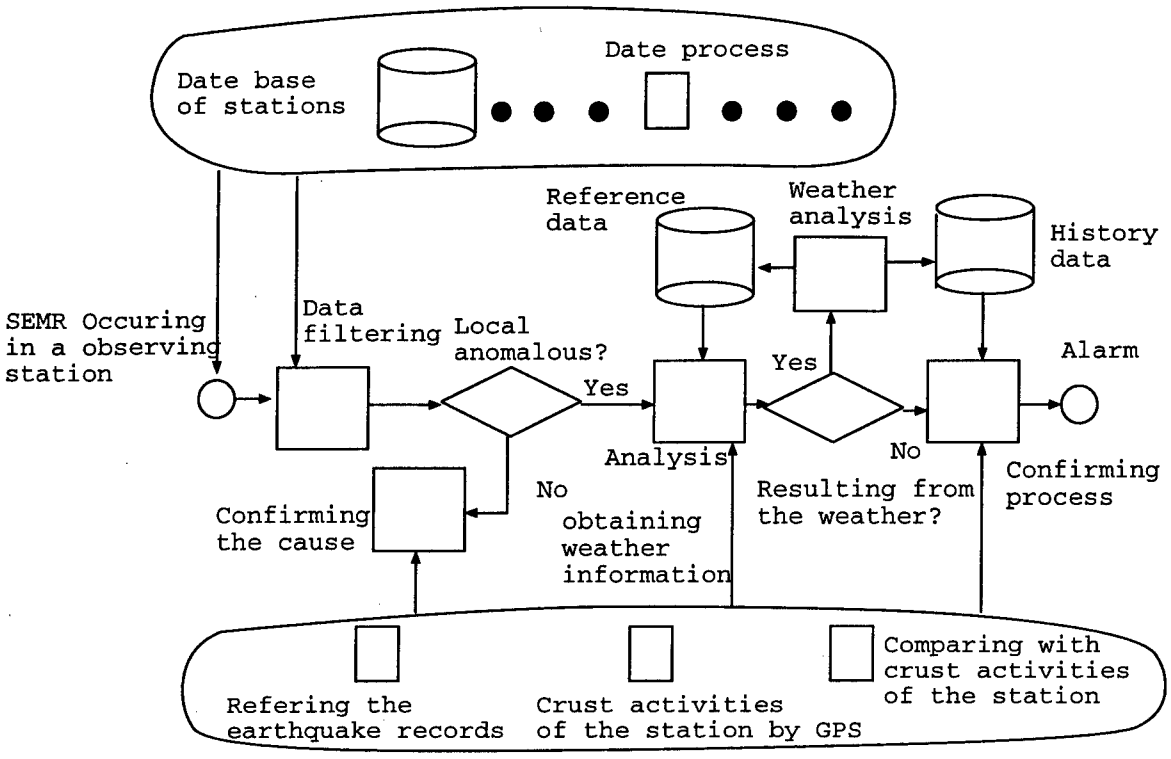


Figure 3: An alarm system

Applications operating atop this distributed computing substrate include alarm system, decision support and visible multiplex map.

3.4 Data process and management

The data of EEMW are observed by observing stations distributed in Japan. At first, data are processed in local station, and compared with standard electromegnetic radiation level of local station which is obtained from history data observed.

The center accesses the data base of observing station according to time and space of event and processes the data with Fourier Transform, Wavelet Transform and other applications. Meaning data will be carried to center(NIT) and saved by the event type, date, and location.

Some related information such as earthquake record, weather situation can be obtained through WWW(World Wide Web), and saved with related anomalous electromagnetic radiation as dataset. Datasets are modeled as collection of objects whose storage formats and location are hidden from application accessing the objects' contents.

4 Main applications

4.1 An alarm system

One application operating on the distributed computing system is an alarm system as shown in Fig.3. When anomalies are found by any one observing station, date of the other stations will be checked to confirm if those are local anomalies in Seismo-EM net based CORBA. The data of every stations are read and handled by appointed period with operations of data object performed. If anomalous change of eletromagnetic radiation is found in the other stations in same level, anomaly is properly resulted from event on a large scale such as universe radiation. This can be confirmed by related information from WWW or by human. In case of local anomalies, nonseismogenic information related local stations from WWW

will be referred. If the factor of weather can be excluded, then information of crust activity, history earthquake record, and related information will be looked up. Those information are analized and compared with anomalous electromagnetic wave in different places and times. The validity of analysis of multi information depends on understanding to the relevance of those information. In most cases, information can not be obtained as wanted, such as crust activity of where sensors are not set. So, the relevance of two specified phenomenon, for example, the relation between anomalies of specified station and crust activity near the station, need to be analized. Those works depend on history data accumulation. After history data referred, alarm will be decided if it should be given.

Certainly, if event information from WWW come, such as earthquake, weather change, crust activity, data of EM radiation will be also checked and confirmed anomalies are saved.

Rules used for dealing with events are obtained from experiences and history data. Since some of them have been not understood completely, results given by the system need to be checked and rules need to be changed ceaselessly. Therefore, a visible process Graphical user interface is given for checking the process of results and changing the algorithm on line, which makes the system to be of high practicality.

4.2 Decision support

One of our goals is to facilitate the location and retrieval of related information, that perhaps stored locally or located at remote data repositories. Decision support is a tool for experts to obtain necessary information regardless of data object's location or the implementation. With this tool, evidence for analizing anomaly can be found efficiently. As shown in Fig.4, there are three information sources, that the local, observing stations and related sources of WWW.

Upon receipt of a request, CGI(Common Gateway Interface) conducts the requests and

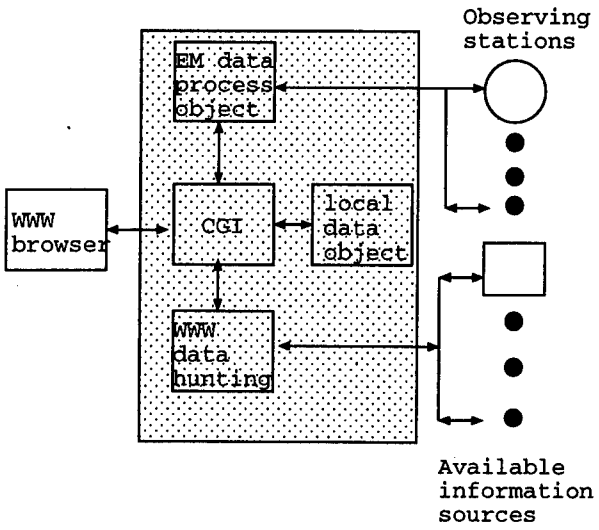


Figure 4: Decision support

parameters to appropriate object. EEMW data process object gets the data of the observing stations through CORBA, and processes the data returned, then return the results to browser. Local data object is obliged to handle the data stored locally, and Data hunting object to search information requested over WWW.

Requests can be query of a event and related information, or data query with specified space and time. The results can be shown as figure or table, and stored if necessary.

4.3 Visible multiplex map

Visible multiplex map can be regarded as the equivalent of the map like GIS, which contains a lot of information related to earthquake. It is meaningful for grasping a sort of information and analizing several sorts of information in the same map such as those of EEMW anomalies, crust activities from GPS and earthquake records. A series method of data process can be selected to process the data.

Besides showing distribution of an index, such as intensity of EM radiation, in space, animation gives a method to show the change of various indexs in a specified period.

5 Conclusions

For research a complicated natural phenomenon like earthquake, referring multi information is proved to be important and valid. About observation of SEMR by Seismo-EM net, data of multi stations and multi related information are referred to analize observation results and remove noise caused from other events. Distributed computing system facilitates this work more efficiently. CORBA technology enables programmer to develop and deploy complex applications rapidly and robustly over heterogeneous, computing and networking elements, regardless low-level infrastructure concerns.

Object hierarchy to serve as a commonfoundation has a advantage to maintain and develop new application. Specially to a application which can not be completed in a short period, a part can be completed and used firstly.

EEMW are observed at about forty observing stations continuously. And processing and storing obtained data is a troublesome affair. The CORBA-Base distributed computing system can leave necessary data process to observing station system, and only meaning data are stored in center as history data. Parallel execution facilitates complex event analysis.

This system provides a method to efficiently process complex queries on earthquakes involving computationally expensive calculations on distributing data sets. Visualizing features of factors related earthquake from a large data set lets us easily analize observational data to gain a better understanding of the earthquake.

Relations of phenomenon resulting from earthquake has not been understood completely. For alarm system, visualized event process and rule change leads to high practicality.

For research of complex natural phenomenon such as earthquake, CORBA technology give a powerful means to develop integrated multi information system in form of distributed computing system regardless of computer types of various information system. With the integrated system which facilitates use of multi information, complex natural phenomenon earthquake and related phenomenons will be understood more exactly.

References

[1] X.Tian and M.Hata. Analysis of seismogenic radiation and transmission mechanisms. *Journal of atmosphere electricity*, Vol.16, No.3:227-235, 1996.

[2] Edmond Mesrobian et al. Mining geophysical data for knowledge. *IEEE expert*, Vol.11, No.5:34–44, Oct 1996.

[3] Jane Grimson. A CORBA-Based Integration of Distributed Electronic Healthcare Records, Using the Synapses Approach *IEEE transactions on information tecnology in biomedicine*, Vol. 2, No.3:124–138, September 1998.

[4] Steve Vinoski. CORBA:Integrating Diverse Applications Within Distributed Heterogenous Environments. *IEEE communications magazine*, 46–55, February 1997.

[5] Silvano Maffeis, Olsen and Associates Douglas C. Schmidt. Constructung reliable distributed communication systems with CORBA. *IEEE communications magazine*, 56–60, February 1997.

Distributed Coordination of Data Fusion

Tiranee Achalakul, Kyung-suk Lhee and Stephen Taylor

Department of Electrical Engineering and Computer Science,

Syracuse University, Syracuse, NY 13244, USA

Email: {tiranee, klhee, steve}@scp.syr.edu

Abstract

This paper provides a brief status report on research aimed at developing a distributed data fusion architecture. The architecture has a variety of applications that range from hospital pathology to battlefield management. It is intended to provide multiple analysts with the ability to cooperatively examine multi-spectral images using concurrent algorithms. The image streams arrive in real-time from multiple sensors distributed throughout a heterogeneous network. An application in cervical cancer cell analysis is presented to illustrate the general concepts.

I. Collaborative Environment

The data fusion architecture couples the Computer Supported Cooperative Work (CSCW) concept [1, 5] with concurrent algorithms and programming concepts. The architecture leverages JAVA-based graphical user interfaces and web-based interaction to provide interactive data analysis. These tools can be used collaboratively to examine the results generated from concurrent image fusion algorithms. In this paper, one such algorithm, the principal component transformation (PCT), is used to illustrate the ideas [2, 3, 4, 8, 9, 10].

Figure 1 shows the architectural concept. A heterogeneous collection of networked PC's, workstations, and shared memory multiprocessors (SMP's), provides the computational resources for high-performance concurrent fusion algorithms. These algorithms are implemented using a concurrent tensor algebra library. This library is, in turn, built upon a heterogeneous concurrent programming library, SCPlib [6, 7], that provides load balancing and granularity control. Multiple sensors may be connected at arbitrary points in the network and interrogated through computation. Multiple analysts may connect to the running computation using a standard web-browser at arbitrary points in the network. Each analyst may utilize a broad collection of JAVA-based data analysis tools, for example, 2/3-D plots, image filters, and multi-spectral viewers. The analysts may collaborate via chat-like interfaces to discuss the results, coordinate the computation, and control sensors. The status of the three central components in this architecture is described in the sections that follow.

II. Analysis Tools and Interfaces

Figure 2 shows the working environment of the collaborative system. A designated analyst (the first connection made to the system) controls access to a concurrent computation that directly manipulates sensor input. Multiple analysts may subsequently connect to a computation from remote computers and are provided with read access to the sensor inputs and computation. Coordination and discussion between analysts is carried through a chat-like sessions. The designated analyst is provided with the privilege to control the interaction modes and computation. The

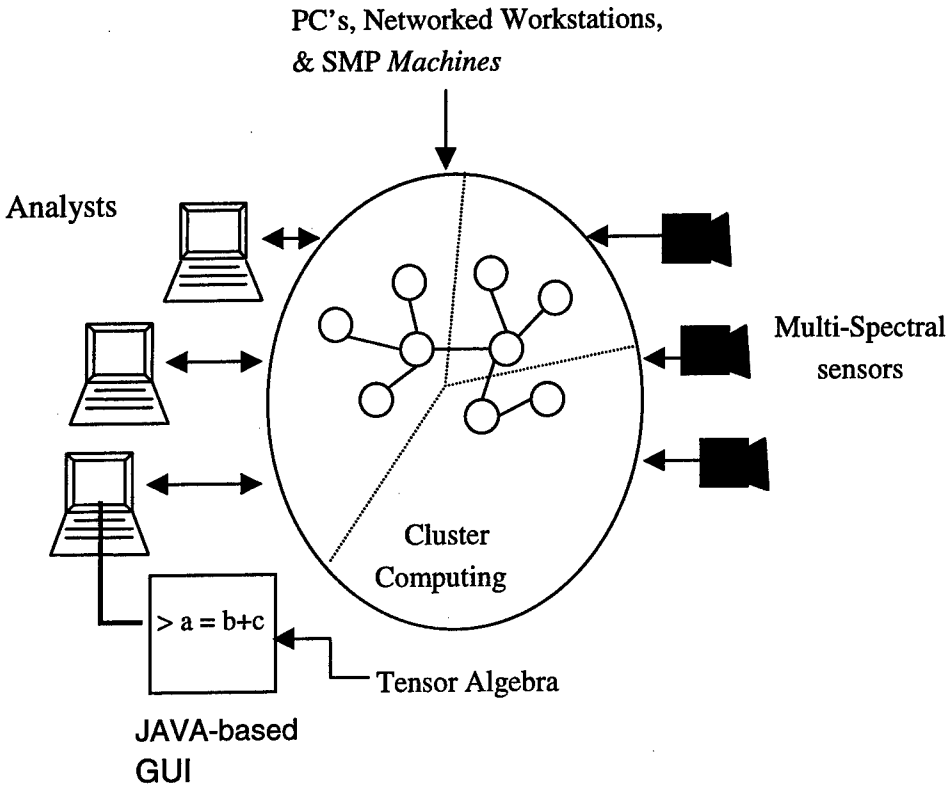


Figure 1: Collaborative Environment

privilege can be handed off to other participants as needed. The same view of computation is available to all the analysts. Each analyst is able to manipulate sensor data through JAVA-based tools individually and share results. The tools provided include 2/3-D plotting, image filtering, and multi-spectral data analysis tools.

III. Concurrent Tensor Algebra

To represent multi-spectral images, the architecture generalizes sequential matrix algebra to concurrent tensor algebra. The following hierarchy explains the relevance of this concept to multi-spectral image analysis:

Tensor Dimension	Conventional Interpretation	Alternative Interpretation
0	Scalar	Real or Complex Number
1	Vector	List of Values
2	Matrix	Table of values or Image
3	Tensor	RGB-Image
N indexed by wave length	Tensor	Multi-and Hyper-spectral Image
N indexed by time	Tensor	Video Stream

For example, using the alternative interpretation, a multi-spectral image of cervical cancel cells collected from multi-spectral sensors, can be represented directly as a third order tensor as shown in Figure 3.

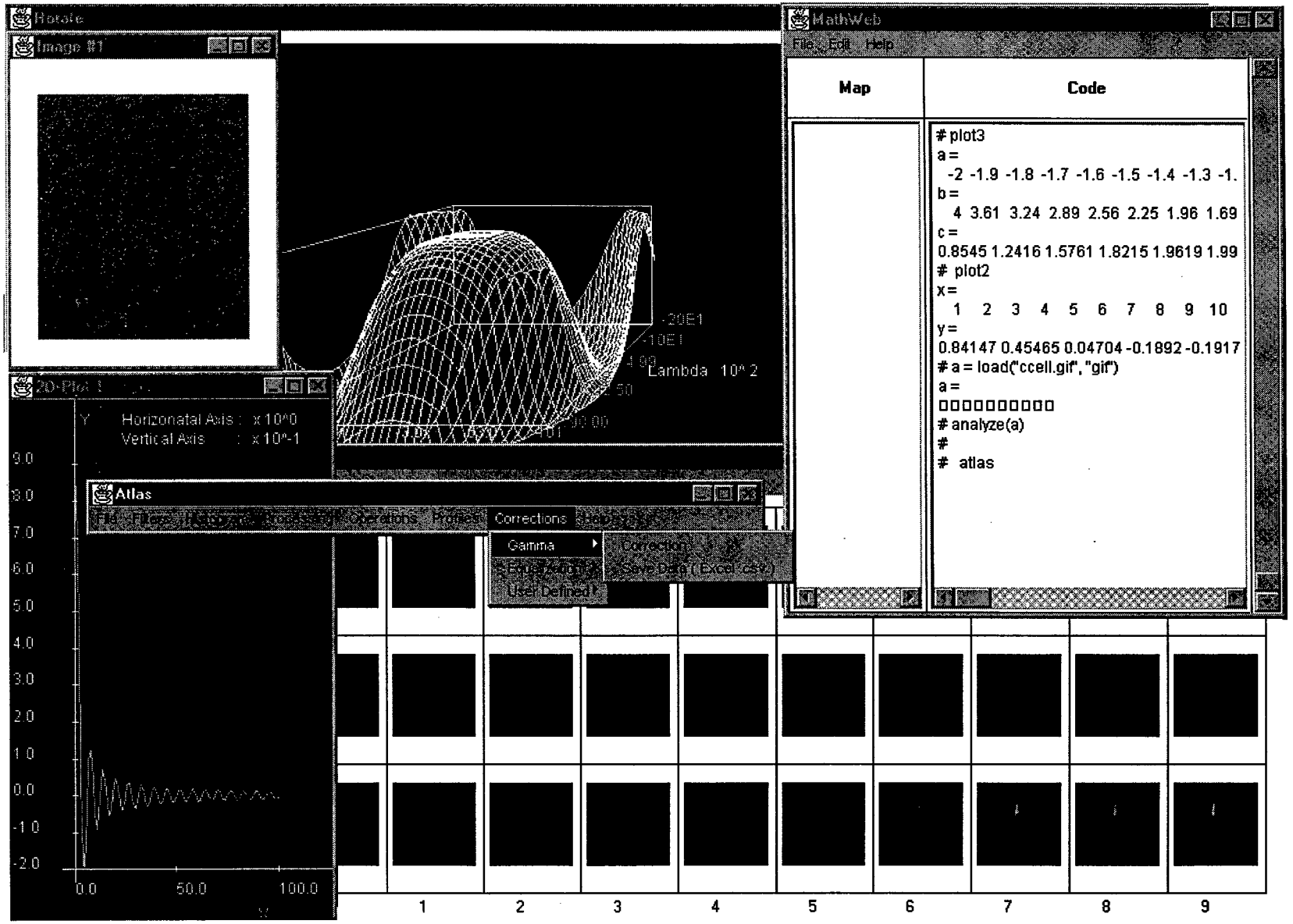


Figure 2: Graphical User Interface

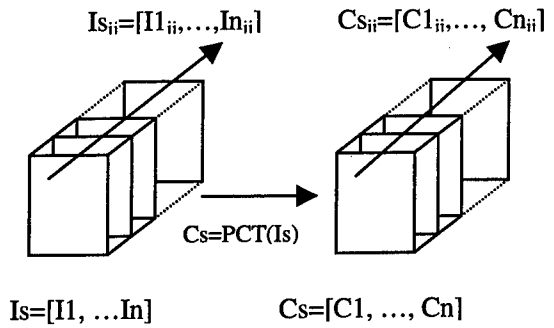


Figure 3: PCT on Multi-spectral Image

IV. Concurrent Computation

The concurrent PCT transforms a multi-spectral image, Is, into the tensor, Cs, using the transformation equation: $Cs = A(Is - m)$ where A is a transformation matrix and m is a mean vector. The computation can be divided into two parts that calculate the transformation matrix A, and subsequently transform the data as follows:

1. *Mean vector*: Each component of the mean vector, m, is the average of the pixel values of an image in each spectral band and can be computed independently as follows.

   for all i = 1 to n concurrently

   $$m[i] = \frac{1}{K}\sum_{k=1}^{K} I_k[i]$$

   where K = number of pixels in an image.

2. *Covariance sum*: To calculate the covariance matrix, pixels at a specific position in all spectra are related while neighboring pixels in the same image are not. Therefore, the pixels in a multi-spectral image are taken as a sequence from the top left to the bottom right. The sequence is divided into P parts using integer division. Each part is allocated to a thread as follows.

   for all p = 1 to P concurrently
   $sum_p = 0$
   for all pixels (i, j) in p

   $$C_{ij} = Is_{ij}Is_{ij}^T - mm^T$$

   $$sum_p = sum_p + C_{ij}$$

   where P = number of parts and sum_p is the matrix sum of the covariance in each part, p.

3. *Covariance matrix*: The covariance matrix is the average of all the matrices calculated in step 2, and is calculated sequentially since its complexity is related only to the number of threads rather than the image size.

4. *Transformation mtarix*: The eigenvectors of the covariance matrix are calculated and sorted according to their corresponding eigenvalues which provide a measure of their variances. As a result, the high spectral contents are forced into the front components. Since the degree of data dependency of the calculation is high, but its complexity is related to the number of spectral bands rather than the image size, this step is done sequentially.

5. *Transformation of the data*: Each pixel vector, Is_{ij}, can be transformed independently. Therefore, once again, pixel vectors in the multi-spectral image are taken as a sequence and divided into p parts.

   for all p = 1 to P concurrently
   for all pixels (i, j) in p

   $$Cs_{ij} = A(Is_{ij} - m)$$

   where P = number of parts.

The concurrent algorithm currently operates only on shared memory architecture, but the tensor data structures have been designed to cope with

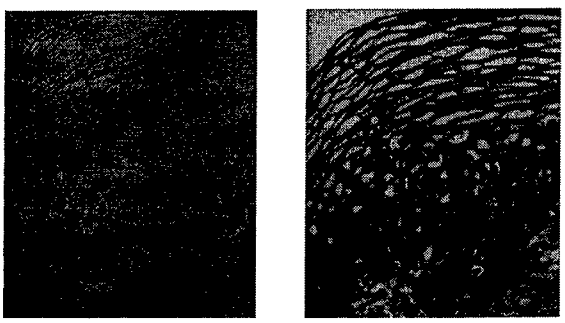


Figure 4: false and true color mapping

network architectures. In this experiment, pathology images are collected at 640x480 resolution, in the visible spectrum, from 410nm to 780nm at 10nm steps. Concurrent PCT is performed on the image cube. True and false color images are created using the first three components of the resulting image cube as shown in figure 4. These images are useful in helping pathologists identify potential cancerous areas in cell samples.

The experimental speed up is plotted against the linear speed up in figure 5. The speed up gained is close to ideal with less than 20% drop of performance. The speed degradation was caused by the thread overhead and the sequential code in steps 3 and 4. When a large number of spectral bands are used, the speed degradation reduces tremendously.

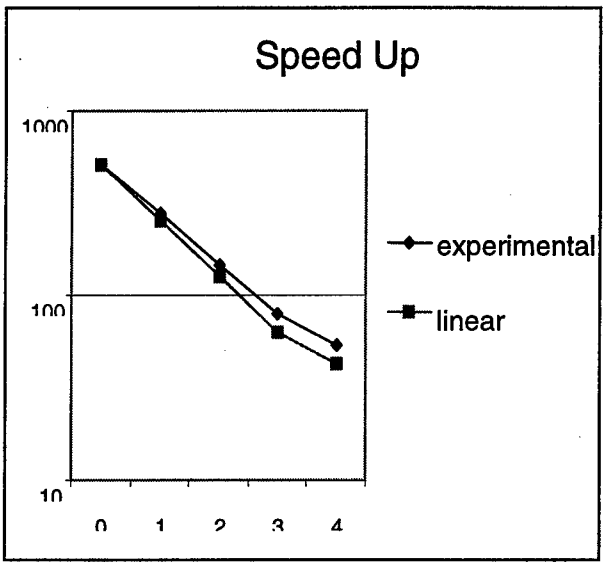


Figure 5. Speed up plot

VI. Conclusion

This paper has presented a data fusion architecture based on distributed systems. The technologies allow multiple analysts to conduct research collaboratively using Web-based programming technologies. Image fusion and analysis can be achieved using the concurrent algorithms on clusters of computers.

V. Acknowledgements

We thank Steve Landas, MD. for providing us the data sets, and the clinical expertise in understanding and analyzing the cervical cancer samples. This work is funded by the Ballistic Missile Defense Organization and managed by the Space and Naval Warfare Systems Center, San Diego under contract N66001-98-18910. We would like to thank Mike Lovern and Greg Adams for their help and support on this contract.

References

1. Annette Adler, "Software Architecture for Cooperative Systems", SIGOIS Bulletin, ACM Press, New York, April 1995
2. Bernhard Flury, *Common Principal Components and Related Multivariate Models*, John Wiley & Sons, Inc., 1988.
3. Ernest L. Hall, *Computer Image Processing and Recognition*, Academic Press, 1979.
4. J. Edward Jackson, *A User's Guide to Principal Components*, John Wiley & Sons, Inc., 1991.
5. J. Quemada, T. de Miguel, A. Azcorra, S. Pavon, "ISABEL: A CSCW Application for the Distribution of Events", Multimedia Telecommunications and Applications, November 1996.
6. J. Watts, S. Taylor, M. Rieffel, and Palmer, "The Concurrent Graph: Basic Technology for Irregular Problems, " IEEE Parallel and Distributed Technology, 4(2):15-25, Summer 1996.
7. J. Watts, S. Taylor, "A Practical Approach to Dynamic Load Balancing," IEEE Transactions on Parallel and Distributed System, Vol. 9, pp. 235-248, 1998.

8. Rafael C. Gonzalez, Paul Wintz, *Digital Image Processing*, Addison-Wesley Publishing Company, Inc., 1977.
9. Rafael C. Gonzalez, Richard E. Woods, *Digital Image Processing*, Addison-Wesley Publishing Company, Inc., 1993.
10. Tiranee Achalakul, Peter D. Haaland, Stephen Taylor, "MathWeb: A Concurrent Image Analysis Tool Suite for Multi-spectral Data Fusion," Sensor Fusion: Architectures, Algorithms, and Applications III, Volumne 3719, pp 351-358, April 1999.

An information fusion system for object classification and decision support using multiple heterogeneous data sources

Erland Jungert
FOA (Swedish Defence Research Establishment)
Box 1165, SE 581 11, Linköping, Sweden
tel +46 13 378 000, fax +46 13 378 058
jungert@lin.foa.se

Abstract: An information system for classification of vehicles and for situation analysis with heterogeneous input data from multiple sources will be proposed. Input data will generally be available from different sensors. The information system will be split into three subsystems, which will be discussed in this paper. The first one will include means for querying and reasoning about spatial-temporal information. For this reason a spatial query language called ΣQL is under development. The second sub-subsystem will be concerned with information acquired from sensors to create a synthetic environment to support situation analysis and to allow terrain feature oriented queries. Finally the third sub-system is concerned with the aspects of man/machine interaction in the system.

Keywords: information fusion, information fusion system, object classification, qualitative spatial reasoning, spatial query language.

1. Introduction

Systems for classification of objects registered by various types of sensors are becoming more complicated as the number of input data sources, i.e. mainly sensors, are growing. Another aspect that complicates the design of systems of this type is that the data from the different sensor types are heterogeneous. Therefore, systems designed to support users in automatic classification of objects collected from multiple sensor data sources must include means for decision support as well as means for visualization of the registered objects, their attributes and the surrounding environment. The decision support tools may include facilities for application of queries directed towards the sensor data; for the storage of sensor data but also for the storage of symbolic information. Clearly, the end-users cannot perceive and analyze this sensor information because of the enormous volumes of data measured in a very short time. For these reasons, a system for object classification using input data from multiple sensors and intended to support the end-users is proposed. The system is concerned with the problems of how the observed objects can be classified and how their positions, orientations, and other attributes can be determined and visualized in a realistic way. Clearly, a system of this kind must also include means for visualization of the terrain in which the classified objects are operating. The latter problem will also be addressed subsequently. Other aspects that need be addressed concerns such aspects as how to simplify the interaction with a system that with necessity will become very complex and how to keep track of the information that will be available in the system. These aspects can, for instance, be handled by means of an interactive intelligent agent, which may be specialized to deal with the problems discussed here.

The system that will be discussed in this paper is mainly intended for automatic target recognition and for integration into a military control and command system but other types of applications can be thought of as well, e.g. applications for environmental surveillance, traffic control etc. Targets of concern are primarily ground vehicles observed either from a top down position or from a slant position. In both cases the objects may be observed from short to medium long distances. Sensor data fusion is another aspect that must be dealt with in a system of the proposed type. A consequence of this is that uncertainties and other limitations of data must be focused.

A few approaches of systems similar to the one proposed here can be found in the literature. Among these can [3] by Shahbazian et al. be mentioned. In their system application programs are automatically linked together through rules available on a blackboard. Such an approach will clearly work in a flexible way. Another and somewhat related method to target recognition is suggested by Nifle et al. [2]

2. The multi-sensor system

The structure of the proposed system can be seen in figure 1. The system can be split into three main parts, i.e.:

- The query language subsystem using data from multiple sources.

- The visualization subsystem with a qualitative terrain feature query system.
- The user interaction subsystem including support for spatial/temporal reasoning.

These tree subsystems can be split further into more specialized modules which will be discussed further subsequently.

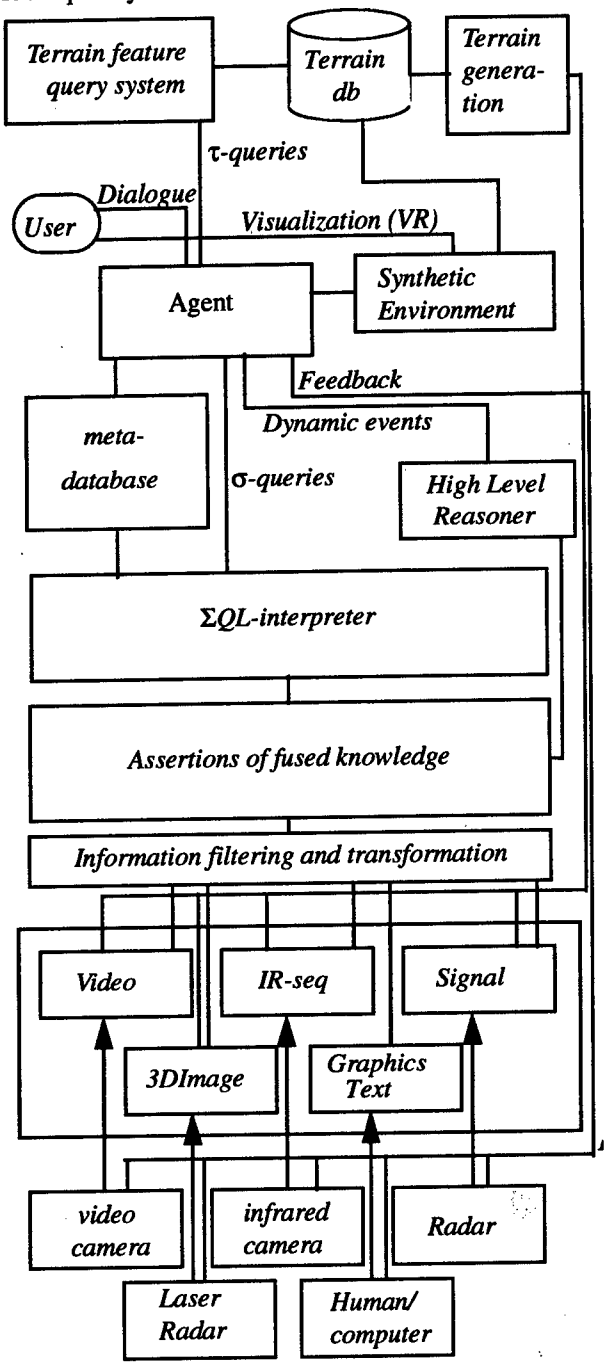


Figure 1. The basic structure of the system.

An important aspect of the system is the control loop, i.e. the feedback loop to acquire more specific information from the sensors as a result of the interaction between the user and the system but also as a consequence of the conclusions drawn by the reasoning system. A further reason is that there may be a lack of available information and therefore there is a need to control the sensors to acquire further information. As a consequence, limitations and incompleteness in the information acquired may be present in the available information and for this reason further information must be acquired from the sensors. This feed back should consequently be a result of a decision made by the user. The users may not only require object information but also information concerning the areas surrounding the object, in other words, reliable terrain information. For this reason we can look at the feedback loop as a means for deciding whether to collect further terrain information or whether to acquire more object information from the sensors. These decisions should primarily, be taken by the users. Therefore the system should be designed so that it will be able to supply them with the necessary information as a result of the dialogue between user and system.

3. ΣQL

A generalized tool to support fusion of the information from the sensors is necessary in a system of the type discussed here. For this reason efforts to support the development of such a tool are going on. The approach taken here is to develop a spatial query language that uses heterogeneous input data, from various types of sensors and transform the spatial/temporal information into a structure that can be used for reasoning on a high abstraction level. The work is part of an ongoing project for development of a query language called ΣQL, see e.g. [4] or [5].

ΣQL uses information from sensors that generally provide continuous streams of data. Such data need be transformed into abstract (symbolic) information of spatial/temporal/logic type. Operations for consistency analysis and information fusion must be available, so that the symbolic information can be used as input to the queries. Queries in ΣQL are basically made up by sequences of spatial/temporal operators, called σ-operators, which easily can be translated into ΣQL-syntax. ΣQL is a natural extension of SQL and allows the specification of spatial/temporal queries and by allowing the use of data from multiple heterogeneous data sources the need to write different queries for each data source is eliminated.

Symbolic Projection [1] and some further qualitative

structures are used by ΣQL as the basic symbolic structure. This structure was originally proposed by Chang et al. [6] for iconic indexing. In this method, space is represented by a set of strings. Each string is a one-dimensional formal description of space, including all existing objects, and their relative positions viewed along the corresponding coordinate axis in a symbolic form. This representation is qualitative as it corresponds to sequences of projected objects and their relative relations. A simple example of the *Fundamental Symbolic Projection* is given in figure 2a, where the U-string corresponds to the projections along the x-axis and the V-string corresponds to the projections along the y-axis. Object A is to the left of the objects B and C, which both have the same x coordinates. The U-string thus becomes A < C = B, and the V-string is obtained in a similar fashion. Figure 2b illustrates an alternative projection method called the *Interval Projection* method [7], where the end-points of the objects are encoded in the projection strings. Thus, the pair (U, V) of projection strings symbolically describes a given image with respect to the identified objects, including the relative positions of the objects and their interrelationships. The different variations of Symbolic Projections are more completely described by Chang and Jungert [1].

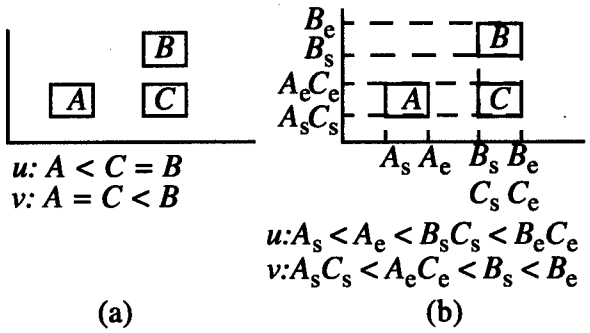


Figure 2. The original approach to Symbolic Projection including the resulting projection strings (a) and the same scene applied to interval projections (b).

Symbolic Projection is used in ΣQL as a means for expressing the spatial relations extracted by a spatial query. A σ-query, on the other hand, is made up by a sequence of σ-operators that can be translated into a ΣQL-query [5]. A σ-operator, when applied to a data source, simply corresponds to a *select* function whose result corresponds to an arbitrary projection string. The σ–operator indicates that the selection is made according to the information-lossless default clustering mechanism. To select the x-axis, $\sigma_x = \sigma_x(x_1,..., x_n)$ is generated and with this notation it is simple to show that (σ_x, σ_y) is exactly the same as the pair of symbolic projection strings in (U,V). Basically, ΣQL is intended for queries generating results of the following types:

- object classification,
- object attributes,
- locations/positions of objects,
- events (when did a certain event occur),
- moving patterns (change in position, paths etc.),
- object relations,
- object orientations.

To accomplish this, the space need to be split up into customary sub-spaces, which the query language can deal with. Such a sub-space is called a cluster. A cluster is consequently a subset of the space spanned up by the various dimensions of the information universe present in the system.

For the image (im_1) in Figure 2a, where the projection method corresponds to the Fundamental Symbolic Projection the result of the application of the σ_x-operator becomes:

$$\sigma_x im_1 = \sigma_x\langle x_1,x_2\rangle im_1 = \langle u: A < B=C\rangle$$

and its corresponding projections in the y-direction becomes:

$$\sigma_y im_1 = \sigma_y\langle y_1,y_2\rangle im_1 = \langle u: A = C < B\rangle$$

It is possible to bring the σ-query further by applying one or further specialized σ-operators as well; thus creating a sequence of σ-operators corresponding to a σ-query. In such a query it is, for instance, possible to ask for object relations such as "which is the direction between object A and B", which yields the result that "B is at the upper right of A" or "B is to the north-east of A". The σ-query corresponding to this type will look like:

$$\sigma_{direction}(N\_E)(\sigma_x, \sigma_y)im_1$$

In terms of ΣQL the above query may be expressed as:

```
SELECT direction
CLUSTER (* ALIAS D(ANY_0, ANY_1))
FROM SELECT x, y
        CLUSTER *
        FROM im1
WHERE (N-E ANY_0 ANY_1)
```

where the WHERE-clause determines the actual relationship, D(ANY_0, ANY_1), that is the pattern indicating a relation of binary type or more specifically the 'direction' between pairs of objects. The * corresponds to a default clustering along the x and y coordinates.

Observe that the existence of other object pairs with the same relationship will be determined as well.

Both input and output to and from the σ-operators are represented in single or multiple strings. This is also true for the object relation queries defined in [4], e.g. for directions, as shown above.

A more complex example concerns a situation where more than one sensor is at hand. Such a query can be formulated as follows. Given that the input information is coming from a laser-radar and a video camera, see figures 3 and 4. Then the query can be based on the observation that it is simpler to determine whether a vehicle in a video frame is moving once it is known whether there is a vehicle present. In this particular case, vehicles can be found in almost real time in laser-radar images, this has been shown by Jungert et al. [6]. However, it cannot be determined whether vehicles found in a laser radar image are moving or not. Thus, once a vehicle has been found in a laser-radar image, it is quite simple to determine if it is moving by just analyzing a small set of video frames from the same time interval. This is possible since the location of the vehicle at a certain time is known from the laser-radar information, which is illustrated in the figures 3 and 4.

Subquery1: Are there any vehicles in the laser radar image in $[t_1, t_2]$?

$$Q_2 = \sigma_{type}(\text{vehicle})\ \sigma_{xyz,interval\_cutting}(*)$$
$$\sigma_t(T^o)_{\ T>t1 \text{ and } T<t2}$$
$$\sigma_{media\_sources}(laser\_radar^o)\ media\_sources$$

Subquery2: Are there any moving objects in the video sequence in $[t_1, t_2]$?

$$Q_1 = \sigma_{motion}(moving)\sigma_{type}(vehicle)\,\sigma_{xy,interval\_cutting}(*)$$
$$\sigma_t(T^o)_{T \bmod 10 = 0 \text{ and } T>t1 \text{ and } T<t2}$$
$$\sigma_{media\_sources}\ (video^o)media\_sources$$

The information from these two sub-queries need to be fused to determine whether a specific vehicle is in motion. Therefore, a fusion operator, $\phi^{merge\text{-}and}$, has to be introduced, which yields the following complete σ-query:

$$\phi_{xyt}{}^{merge\text{-}and}(*)$$
$$(\sigma_{motion}(moving)\sigma_{type}(vehicle)$$
$$\sigma_{xy,interval\_cutting}(*)$$
$$\sigma_t(T^o)_{T \bmod 10 = 0 \text{ and } T>t1 \text{ and } T<t2}$$
$$\sigma_{media\_sources}\ (video^o)media\_sources$$
$$\sigma_{type}(\text{vehicle})\ \sigma_{xyz,interval\_cutting}(*)$$
$$\sigma_t(T^o)_{\ T>t1 \text{ and } T<t2}$$
$$\sigma_{media\_sources}(laser\_radar^o)\ media\_sources)$$

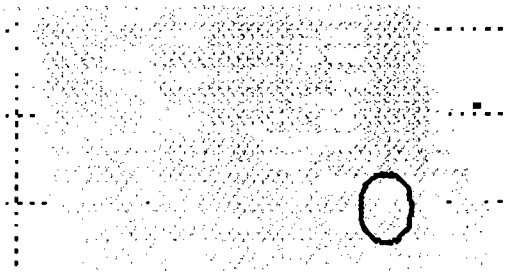


Figure 3. A laser radar image of a parking lot with a moving car (encircled).

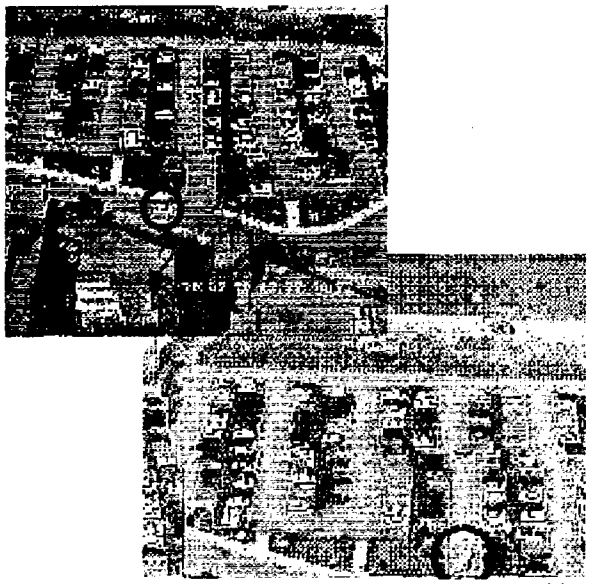


Figure 4. Two video frames showing a moving white vehicle (encircled) while entering a parking lot

Each of the images are transformed into the corresponding projection strings for each sub-query. The reasoning is then carried out by the $\phi^{merge\text{-}and}$-operator to determine whether any of the found vehicles are moving. Translation of the σ–query into ΣQL-syntax is simple and straight forward and the result of this translation is:

```
MERGE-AND x,y,t
CLUSTER *,*,[t1,t2]
FROM (SELECT type
  CLUSTER vehicle
  FROM SELECT x,y,z
    CLUSTER OPEN (* ALIAS T)
    FROM SELECT media_sources
      CLUSTER OPEN laser_radar
      FROM media_sources
      WHERE T > t1 AND T < t2,
  SELECT motion
  CLUSTER moving
```

```
FROM SELECT type
  CLUSTER vehicle
  FROM SELECT x,y
    CLUSTER interval *
    FROM SELECT t
      CLUSTER OPEN (* ALIAS T)
      FROM SELECT media_sources
        CLUSTER OPEN video
        FROM media_sources
        WHERE T mod 10 = 0
                  AND T>t1 AND T<t2)
```

Among the above operators the MERGE-AND operator obviously is the most complex and clearly it need to include means for investigation and solution of the association problem but also for handling uncertain information obtained from the sensor data. Such techniques are needed to make this operation general. For this reason, further research is needed.

4. The visualization subsystem

The visualization subsystem will be designed to present what is going on in the environment assessed by the sensors, but also to support the users in their efforts in determining the existence of various types of objects and determine their behavior and activities with respect to the geographical surroundings. Most information made available by the sensors should be possible to be acquired by the users whenever necessary. Hence, the visualization subsystem must include a spatial/visual query system for determination of information about geographical objects that eventually may be visualized in the synthetic environment thus allowing the user to follow the on-going activities of concern. This is, in other words, a way of enhancing the synthesized environment with respect to activities otherwise hidden from the users.

4.1. The Synthetic Environment

The information that will be used to create the synthetic environment will generally be coming from the sensors. This information need be classified with respect to their geographical types. Of special interest is to allow the generation of a high resolution terrain model in 3D. There are various ways of doing this and a number of sensors exist that can support this. In this work a laser-radar called TopEye, which is a civilian product from Saab Survey of Sweden, has been used. This sensor uses a helicopter as a platform. The laser-radar contains a vertical scanning direct detection laser and the successive pulse emission and repetition do not overlap. The overall accuracy of the measured point co-ordinates is approximately 0.1 m in all three co-ordinate directions.

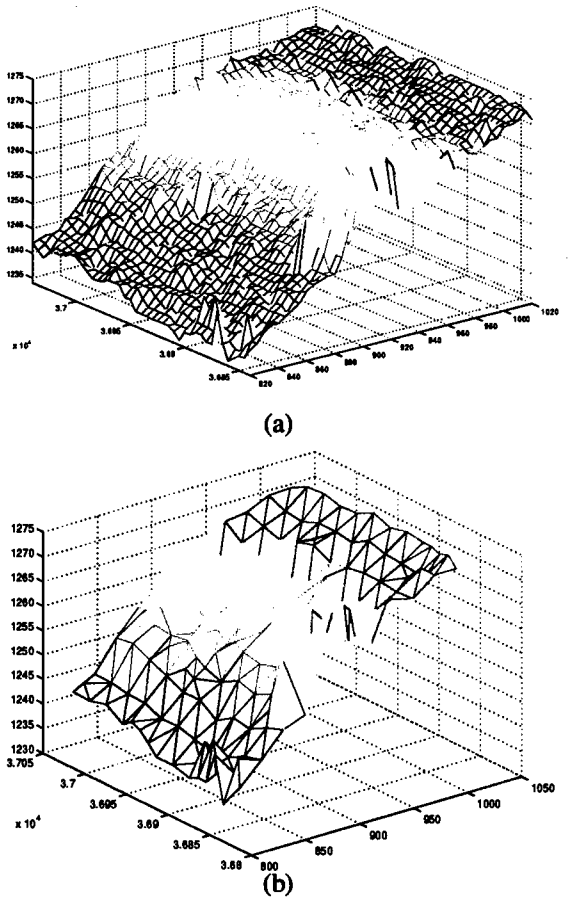


Figure 5. A terrain model in a high resolution grid structure (a), and the reduced model after the wavelet transformation (b).

The technique used here for generation of terrain elevation models from laser-radar images is discussed in [9] and [10]. A critical step in this process depends on how accurate separation between ground and forest information can be performed. A laser-radar image maps the terrain from a top view, which includes reflections from both the ground and the vegetation. In order to determine the information corresponding to the ground surface forest and vegetation information from the laser-radar images must be removed. In cases where the forests are very dense this problem has no simple solution. The main problem is concerned with the loss of direct reflections from the ground and how to interpret and fill in the areas hidden underneath the trees with information without to much loss of accuracy. This problem is further discussed in [10].

Once ground and forest information has been separated the process becomes more straight forward. In this approach a regular grid with a point distance of 0.5 m is

first created, see figure 5 a. From this grid structure an irregular sparse structure is determined. This structure is made up by a regular grid, with a grid point distance of 2 m, and a set of irregularly distributed characteristic points that eventually should be used for triangulation, see figure 5b. To accomplish this, it is necessary to substantially reduce the number of internal data points while still keeping the most significant, i.e. significant data points belonging to such terrain structures as hills and ditches etc. This is motivated by the fact that storage and access requirements require a structure where unnecessary points are eliminated. A square size of 2 m was chosen since it was considered as a suitable compromise between the goal of reducing data while keeping the elevation errors below 0.5 m in total. The irregularly distributed points are determined by a somewhat modified version of the linear wavelet transform [11]. The 2 m squares and the points that occur inside the squares constitute the basis for determination of a qualitative code that support describing the terrain.The different categories that can be identified through this technique will eventually be stored in the terrain database.

Figure 6. A terrain model with a simple skin on top, illustrating a round-about with four incoming roads.

The terrain data model will have a dual use. The triangulated grid structure including it irregularly distributed points will be used for visualization purpose only, while the categorized structure will be used for querying purposes. The latter will be discussed further in section 4.2. A triangulated terrain model using laser-radar data can be seen in figure 6. The model has been covered with simple skin and shows a round-about with four incoming roads.

The result of the terrain generation process gives a total reduction of data of about 88% compared to the original data volumes. The average error is, as determined from the maximal error in each square, 0.128 m and with a standard deviation of 0.065 m. In conclusion a reduction of data close to 90% have been possible and with a total maximum error in each square less than 0.5 m in 95% of all cases when adding also the measurement error of the sensor which is around 0.1 m.

4.2. The terrain feature query system

The queries applied to the terrain feature query systems are called τ-queries. The τ-queries have similarities to the σ-queries, since they also correspond to a kind of spatial low-level queries. The τ-queries are, however, more shape and geometric oriented as they are built up by patterns of the squares determined from the irregular grid structure in the terrain model. A τ-query is simply a matching process on a symbolic level between a given pattern and the pattern of the existing terrain model where both are represented in terms of a qualitative tile code.The purpose of the terrain feature query system is therefore to answer queries made up by the patterns of categories of grid tiles as discussed in section 4.1. The basic idea is consequently, to describe the features of a terrain object in terms of combinations of grid tile categories, and match them against the tile codes given in the terrain database. This kind of matching can be applied to determine, for instance, ditches.

About a hundred qualitative grid tiles have been identified. The simple features in the grid tiles can be combined such that more complicated types can be identified as well. Figure 7 shows some examples of the different types and their features. The grid tiles in figure 7 correspond, from left to right, to a flat area, an area with a top, a ridge, a ridge close to the lower left corner and finally a structure of two combined features e.g. a ridge at left and a valley at right. The tiles must be rotation invariant with respect to their features and it is simple to see that they can be individually triangulated for visualization purposes. Furthermore, since the tile features are qualitative a special feature like the third tile in figure 7 can have any arbitrary direction as long as it stretches out from top to bottom inside the tile.

Figure 7. Some example of grid tiles with some simple feature types.

A query may not correspond to a single tile but to a number of tiles, that in combination must be matched against the terrain model. Sets of set of tiles may be permitted as well. The matching procedure goes on as a

filtering technique between the combinations of sequences of tile codes until the specified area has been completely covered by the query. The work on development of a τ-query system is currently under way.

A further aspect when applying the τ-queries concerns the size of the objects. The 2 m grid works well for small objects represented in a high resolution. Larger objects, however, require a lower resolution, because otherwise the combinations of tiles will become too complex. Therefore, a type of resolution pyramid is being developed to make it possible to apply grid tiles of lower resolution in the matching. Useful grid sizes would be 4, 8 and 16 m, which can be generated directly from the basic terrain database. This can be done in a first step by extracting the actual grid size from the terrain database and in a second step by identifying the remaining points by means of the vawelet transformation as performed on the 2 m level.

Finally, another class of τ-queries that eventually must be focused on concerns the identification of buildings and various types of features of buildings. However, buildings require a different kind of description than the terrain models; mainly because it is simpler to describe a building more exactly. However, this requires further studies.

5. The interactive control and reasoning subsystem

In the interactive control and reasoning subsystem user related aspects are mainly in focus. Aspects as how to interact with the system, how to keep track of available and incoming information and how to draw further and more extensive conclusions from the knowledge made available by the query systems, will be in focus. So far, the studies concerning this sub-system are preliminary and further work is required before the details of this subsystem are identified and can be implemented.

The main modules in this subsystem will generally be made up by the interaction module, the high level reasoner and the meta-database.

5.1. The interaction module

The interaction module will be a combination of an intelligent software agent and a web-browser. The former is intended to support the users in a collaborative working mode such as described in [12]. The main purpose of this agent is to support decision making trough a user dialogue. The agent should guide the user towards more correct decisions, while at the same time the work-load of the users should be reduced thus eliminating the complex interactions that otherwise would require a high-level user competence. Interaction with SQL will be based on the web-browser technique, which will be a convenient way of interaction since most users are familiar with this technique. Finally, an important aspect to consider is to develop a way of dealing with the problem of how to eliminate information overload of the user.

5.2. The high level reasoner

The purpose of the high level reasoner is to support the second level of data fusion, i.e. the situation analysis as described in [8]. This module should be designed to support higher levels of information fusion and the input should be fed into the module as assertions from the depository of fused knowledge mainly coming from ΣQL. Information from the terrain feature query language will be needed by the reasoner as well. As a consequence, the high level reasoner will mainly be concerned with the objects determined by the target recognition process. The information that relates to the synthetic environment will be used to infer high level knowledge of compound type. This subsystem will be tightly coupled to the interaction module and its intelligent software agent.

5.3. The meta-database

Meta-data must be available to the user as well. Without this kind of information it will be immensely difficult to determine whether a query can be answered properly. For this reason, a meta-database must be available to determine whether data needed for a particular query is available in the system. However, the metadatabase need not be explicitly available to the users but can instead be indirectly available through the intelligent agent or connected to ΣQL. The metadatabase should include information generated by the sensors. This information does not only descriptions of raw-data but also information generated through transformations performed by the system. Of importance here is such information that tells where and when the information was acquired and, of course, information about the various types of available information.

6. Adaptation

Whenever there is not sufficient information available for a certain task it should be possible to control the positions of the sensors to get further information from a certain activity at a certain location. This information should have a better quality be more reliable etc. and thus give ΣQL a better chance to fulfill its assignments.

However, adaptation of the sensors should not only be determined by the query systems but primarily by the users in their dialogues with the agent.

7. Conclusions and future research

The system proposed here is at this time not a running system but rather at a planning stage where certain parts have been subjects to research while other parts are subject to on-going research activities. The prime goal is to build up a demonstrator of the proposed system. This also requires activities concerned with collection of relevant sensor data. Some sensor data, which has been used for earlier research, such as laser-radar data, have, however, been available for quite some time and as a consequence they have had a strong impact on the design of the system.

The system as illustrated here just shows a single user system, which however, should be possible to integrate in a multi-user system where the communication network should be transparent to the users. The system should in other words correspond to a CSCW system. This should, however, be subject to future research.

References

[1] S.-K. Chang and E. Jungert, "*Symbolic Projection for Image Information Retrieval and Spatial Reasoning*", Academic Press, London, 1996.

[2] Nifle, A., Bedworth, M. and Reynaud, R., "*Multi layout identification*", proceedings of Eurofusion, Malvern, England, October 6-7, 1998, pp 177-184.

[3] E. Shahbazian, J.-R. Duquet, M. Macieszczak, P. Valin, "*A Generic Expert System Infrastructure for Fusion and imaging Decision Aids*", Proceedings of Eurofusion98, October 6-7, 1998, Great Malvern, UK, pp 167-174.

[4] S.-K. Chang and E. Jungert, "*A spatial/temporal query language for multiple data sources in a heterogeneous information system environment*", the Int. J. of Cooperative Information Systems (IJCIS), vol. 7, Nos. 2 & 3, 1998, pp 167-186.

[5] S.-K. Chang, G. Costagliola and E. Jungert, "*Querying Multimedia Data Sources and Databases*", Proceedings of third international conference on Visual Information Systems (Visual'99), Amsterdam, The Netherlands, June 2-4, 1999.

[6] Chang, S.-K., Shi, O. Y. and Yan, C. W., "*Iconic indexing by 2D strings*", IEEE transactions on Pattern Analysis and Machine Intelligence (PAMI) 9 (3), 413-428 (1987).

[7] Jungert, E., Carlsson, C., and Leuhusen, C.,"*A Qualitative Matching Technique for Handling Uncertainties in Laser Radar Images*", Proceeding of the SPIE conference on Automatic Target Recognition VIII, Orlando, Florida, April 13-17, 1998, pp 62-71.

[8] E. Waltz and J. Llinas, "*Multisensor data fusion*", Artect House, Boston, 1990.

[9] E. Jungert, F. Lantz and G. Neider, "*Generation of High Resolution Terrain Elevation Models from Laser Radar Images*", Proceedings of the 4th EU-GIS workshop, Budapest, Hungary, July 24-26, 1998.

[10] E. Jungert, U. Söderman, S. Ahlberg, P. Hörling, F. Lantz, G. Neider, "*Generation of high resolution terrain elevation models for synthetic environments using laser-radar data*", Proceeding of the SPIE AeroSense conference, Orlando, Florida, April 5-9, 1999.

[11] G. Strang and T. Nguyen, "*Wavelet and Filter Banks*", Wellesley-Cambridge Press, Wellesley, MA, 1996.

[12] E. A. Edmonds, L. Candy, R. Jones and B. Soufi, "*Support for Collaborative Design*", Comm. of the ACM, July 1994, Vol. 37, No 7, pp 41-47.

An Architecture for Multi-Sensor Fusion in Mobile Environments

Datong Chen, Albrecht Schmidt, Hans-Werner Gellersen

TecO (Telecooperation Office), University of Karlsruhe
Vincenz-Prießnitz-Str.1, 76131 Karlruhe, Germany
{charles, albrecht, hwg}@teco.uni-karlsruhe.de

Abstract - This paper introduces a layered architecture for multi-sensor fusion, applied for environment awareness of personal mobile devices. The working environment of personal mobile devices changes dynamically depending on their user's activities. Equipped with sensors, mobile devices can obtain an awareness of their mobile working environment, to improve their performance with respect to usability. The mobility of the device presents two problems for building an awareness system. First, the contexts to be covered by an awareness system depend on the users, their tasks and activities, and also on the data that can be obtained from different sensors. Second, the power consumption and the size of the mobile device limit the processing capability of an awareness system. The solution presented here is to design a low cost sensor-based fusion system, which can be reconfigured by the user, to enable individualized awareness of environments. The software architecture presented in this paper is designed with four different layers, which can support reconfigurations in mobile environments.

Keywords: mobile environments, multisensor fusion, context-awareness, fusion architecture

I. Introduction

Personal mobile devices, such as laptop, GSM and PDA, break the traditional desktop paradigm and bring people the powers of the computing and electronic communication anywhere and anytime. Our investigation focuses on improving the function and interface of these personal mobile devices through awareness of the user's activities and the current social environment. Different from the desktop, mobile devices are portable and accompany their users from one place to another. This kind of mobility puts the device into a changing environment, which is more complex to be processed than in fixed cases, while it also offers them more opportunities to know more about their users and their own situations with certain awareness techniques. For example, a PDA may track the locations of its user from the home to the office and adjust the items in the "to do list" from home-related issues to the work-related issues. It may also recognize that the user starts to walk after a calmly sitting and then change its display to the large font automatically to ease reading. Many investigations have already been done on applying the desktop-based awareness to improve the interaction between human being and the computing device [1, 2]. Based on these former works, a multi-sensor fusion architecture to enable awareness for the mobile devices is presented in this paper.

To enable the awareness of mobile devices, a small multi-sensor device is developed by the European Commission funded research project Technology for Enabling Awareness (TEA, [3]). This multi-sensor device can be connected to a mobile device as an additional part and offers useful context information to the host. Aiming not to destroy the portability of the mobile device, the multi-sensor device is designed to employ only low cost sensors and rely on fusion techniques to extract useful contexts from the data obtained from these low cost sensors. "Low cost" means that:

First, the size of the sensors should be small enough to keep the multi-sensor device much smaller than the size of the host device. Second, the sensors should consume low power and the signals they produced can be processed with little processing power. Finally, the price of the sensors is also a factor that should be regarded.

Investigating how to enable awareness in mobile environments, two kinds of adaptation are necessary when the working environment of the mobile device is dynamically changing with situation and location. One is that in different situations certain sensors are more useful than others. For example, the air pressure sensor may be useful when the user is on a flying plane, but can not offer much useful information when the user is siting in the office room. Operations to adjust sensors, such as switch on/off, affect the related fusion algorithm to produce stable results. The other

adaptation is needed because, in different environment, the mobile device is interested in different contexts. For example, at night, the mobile device may pay attention to the context about whether there are artificial lights. But in the daytime, this context may be not necessary. The fusion-based context awareness algorithms, which compute other contexts according to the context artificial light, need to be able to adapt to this modification. The multi-sensor fusion system for mobile environment should be designed robust enough to adapt the continuous reconfigurations of both sensors and contexts.

In many former works, the sensor fusion can be classified into different levels according to the input and output data types [4, 5]. The fusion may take place in the data level, feature level and decision level.

In data level fusion, the raw data from sensors is used to extract features [6]. Varieties of the methods are developed in this level, and were applied in the image processing, visual & speech recognition, data compression and intelligent control [7, 8, 9]. The feature level fusion is to fuse the features extracted from multi-sensor data into new features or the final decisions. Because most features have well-defined structures, the fusion methods in this level can be based on statistical approaches and pattern analysis approaches [10, 11]. Decision fusion is a common problem in many research areas, such as decision theory and artificial intelligence. An example of the simple decision fusion is the voting system, in which every candidate has equal or not equal right to determine the final result [12]. Artificial intelligence techniques show new trends for the solution for decision fusion, for example the neural network [13]. There are two advantages of applying neural networks to fuse the decision. One is that the neural network is noise-tolerant and can process the input features with plenty of noise. The other advantage is that neural network allows the system to be reconfigured according to the specified application instance.

2. Layered architecture

The adaptation of the reconfiguration of the sensors and contexts in the mobile environment is the important factor in designing the architecture of the fusion software system. When the sensor is modified (being switched on/off or adjusted its sampling rate) in the system, there should be a feasible mechanism to let the related fusion processes know this change and make correct responds. On the other hand, when the user reconfigures a context in the system, the feedback of this adjustment should also activate the correct adjustment of the related processes and sensors. To develop a common and feasible reconfiguration fusion system, one method is that we define the whole fusion system with several independent layers. Each layer consists of certain structures and data processes, and keeps contact with next layers through defined interfaces. In this way, the reconfiguration in one layer can be controlled by the predefined function in this layer and the effect of the modification can be limited by the interface to the next layers. In other words, the result of the reconfiguration in one layer can be regarded as a kind of normalizing the input of the other layer, so that the adaptive fusion algorithms can be developed in different layers separately. In this paper, we describe a fusion architecture with four layers, see figure 1.

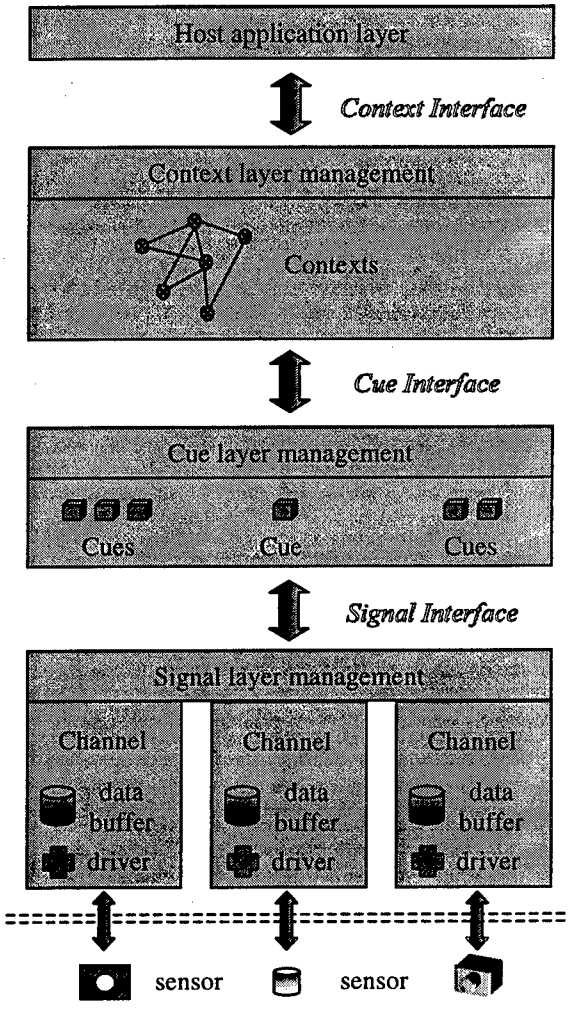


Figure 1. Four layers fusion architecture

2.1 Signal layer

The lowest layer is called signal layer, which connected with the sensors directly. The function of the signal layer is to control the data collection of the sensors and write the data into a uniform structure. A special kind of software channel is employed in this layer to adapt the reconfigure of the sensors. For each sensor, there is a channel with corresponding driver, data buffer and other attributes to manage it temporally. Three attributes of one channel are the logical name of the signal read from the sensor, which is used to identify the corresponding driver of the sensor; a time stamp system to manage the data stored in the buffer; a sampling frequency system, which is used to respond to the current available sampling statues. When the hardware of the system is modified, for example a sensor is added, a sensor is removed, a sensor is switched on/off and so on, the sampling frequency system of the related channels will detect the change automatically and adjust the value of the sampling frequency. This sampling frequency value can also be set by the system through software directly.

The output data of the signal layer is the raw signal data with a structured description. The description involves the information about the current data, such as the time stamp, the sampling frequency, the number of dimensions, and the size of the each dimension. Most of the signals employed in TEA project have one dimension, for example, light signals, audio signals, temperature, etc. There is also two or three-dimensional signal such as the acceleration signals.

2.2 Cue layer

The processes in the cue layer mainly focus on the time independent features extracting from each single channel data. The time independent features extractions transform the time-varied data space into time independent feature space. From our point of view, the information fusion can be regarded as a data compression process. The raw data from several sensors will be compress into the result space. The fusion across different sensors is to reduce the redundancy among the data of these sensors. The reduction of the redundancy among the data of on sensor is also a kind of information fusion. Except for the time independent features extractions, the data from multi-dimension sensors is transformed into independent feature space in the cue layer. The time-varied analysis in the cue layer is limited within only a short period of sample data. Long term analysis will be done in the higher layer.

We call these kinds of the self-independent features from single sensor channel as cue, in order to show their differences with the common concept of feature. The cue layer keeps a specified period of history of cues, which serves as a history description of the changing environment.

2.3 Context layer

The perceptible events in the environment are treated as the contexts of the activities of the host device in this layer. The current contexts can be derived from several cues, deduced from former or other current contexts, or combine the two approaches together. The system employs semantic nets to represent the former and current contexts. This semantic nets are designed with a limited verb set and probability description, for example, the current contexts can be represented like that "At 10:32, with 85% probability, (it) starts to walk, in the office". Each context keeps a value of its own respond frequency, which can be adjusted by the user according to his needs. More deep reconfigurations of the context, such as add a new context or training the context layer to recognize your new office room, need the cognition and deduce fusion approaches in context layer are self-adaptive or can be trained manually. Artificial neutral networks are good tools to support the deep reconfigurations of the context, because they can be trained through the examples automatically. The decision tree is another possible method to reconfigure the deduce algorithms. The context layer keeps the history of the contexts, which can be rewritten into the nodes in semantic nets to perform certain deduce algorithms.

2.4 Application layer

The application layer is developed within the operation system of the host and uses the result of the fusion system to improve the services of the host devices.

2.5 Interfaces

The communications between different layers rely on the fixed interfaces defined in the architecture. The interface between signal layer and cue layer is called signal interface. Through signal interface, the cue layer can read the data from each available channel and set the sampling frequencies of it. On the other hand, the signal layer can sent messages to activate the cue layer whenever the data is updated or the sensors are switched on/off. The cue interface is designed to keep contact between the cue layer and the context layer. By using this interface, the context layer can not

only access the current cues, but also has access to the stored history cues. The information about the updating of the respond frequency of the context can be sent to cue layer and further extended to the signal layer. Similar as in the signal interface, the cue interface also supports to send the cue-updating message from the cue layer to the context layer. The interface between the context layer and the application layer is the context interface. In order to apply the multi-sensor awareness device to different mobile devices, the context interface is designed as a one way interface, which offers the access only from the application layer to context layer. It offers a rich set of functions to the host applications, including reading current and history contexts, setting the respond frequencies of the contexts, setting the attributes of the contexts, recording the samples and training the algorithms in the context layer, adding a new context or deleting an old one, and so on.

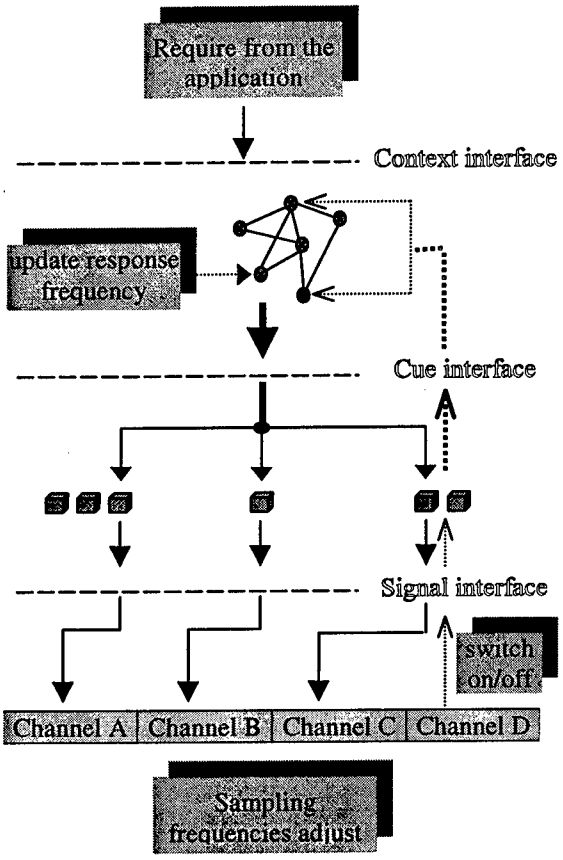


Figure 2. Reconfigure information feedback

2.6 Reconfigure information feedback

The information to reconfigure the system can be transmitted both ways: from the signal layer to the context layer and from the application layer to the signal layer.

The both feedback processes are depicted in figure 2. When the host application wants to modify the response frequency of a certain context, it sends a command to the context layer through context interface. In the context layer, first, the respond frequency of the specified context will be updated to the new value according to the command, if this new value is valid. And then, the new value will be transmitted to the related cues in the cue layer. Because one context may be the fusion result of several cues, and one cue may also be employed by different contexts, in the cue layer, the related cues decides whether they should adjust themselves to adapt the change of this context while do not affect other related contexts. If the cue chooses to change its respond frequency to the new value, this value will be transmitted to the corresponding channel in the signal layer. The channel, which receives this information, may adjust its sampling frequency after checking all the cues extracted from this channel.

When a sensor is switched off, the corresponding software channel should detect it and informs all the cues that based on this channel. This channel will be disabled under the signal layer management, but the related cues are still enabled because the history of these cues can be used for the future awareness. If a sensor is switched on, the signal layer will detect its signal, enable the channel and recover to send the updating message to the related cues. The context layer will check the time stamp of the cues before using them. A cue, which has not been updated for a long time according to its own respond frequency, will be regarded as unavailable resource. If this happens, the related algorithms in the context layer will be reconfigured with predefined methods.

3. Evaluation

In the experiment described in this section, we deployed the prototypical tea-device [14], a sensor-board that reads environmental parameters using a number of low cost sensors.

3.1 Hardware

The board consists of four major blocks: the sensors, the analog-to-digital converter, the microcontroller, and the serial line. The sensors measure the conditions in the environment and translate them into analog voltage signals on a fixed scale. These analog signals are then converted to digital signals and passed to the microcontroller. The microcontroller oversees the timing of the analog-to-digital converter and the sensors as well as manipulating the data from the analog-to-digital converter's bus to the serial line. Finally, the serial line connects to the higher layer, see Figure 3. In terms of the architecture described earlier, the hardware incorporates sensor and parts of the sensor dependent drivers (signal layer) implemented in a microcontroller. The communication between the sensor board and the mobile device is using a serial-line in a multiplex mode. In this prototype, the higher layers are emulated with a laptop, which connected between tea-device and the host device to control the experiment easily.

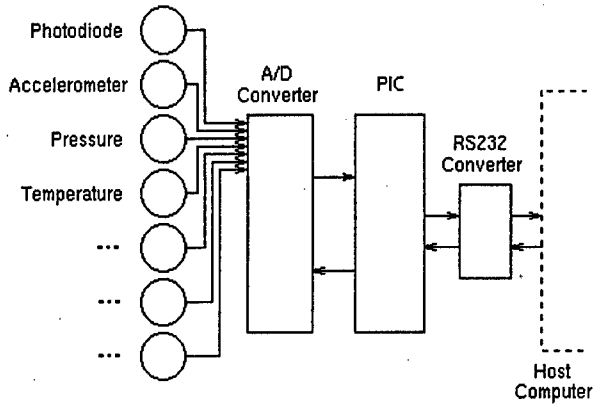


Figure 3. Schematic

3.2 Software and Interfaces

The context, cue, and signal interface are offered as C++ methods to the next higher layer. The context and cue layers are implemented entirely in C++, too. For the host application layer we used different host dependent implementations. The signal layer is partly implemented in C on the microcontroller and partly in C++.

3.3 Experiments and Results

In the experiment, we collected data of all sensors in different contexts cycle by cycle, as described in Table 1. Within each cycle, the sensors were activated and read according to their sampling frequency to feed the environmental parameters. The data for each context was collected over a time of about 100 seconds, or about 120 records. Selected parts of the data are depicted in the following figures.

Table 1. Contexts samples

Context	Description
Inside-1	office, artificial light, stationary
Inside-2	office, artificial light, walking
Outside-1	outdoors, daytime, cloudy, stationary
Outside-2	outdoors, daytime, cloudy, walking

Looking at the light data sample in Figure 4, it shows the values of brightness at cloudy outside and inside with artificial light. It is obvious to find the difference between inside and outside on the level of light as well as on the oscillation of the light. Comparing the acceleration data for a stationary device in figure 5. with the one for a moving device in figure 6, it can be seen that they differ significantly.

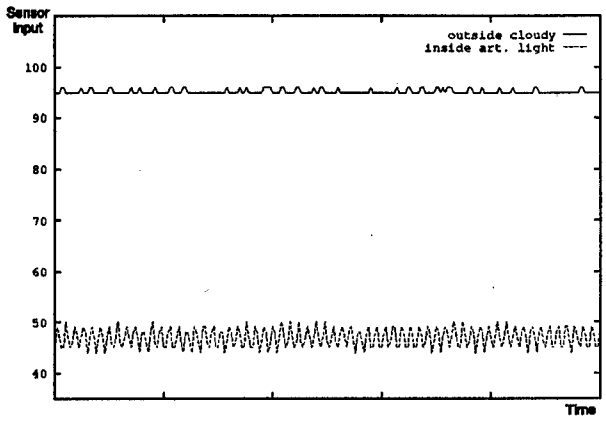


Figure 4. Light sensor data

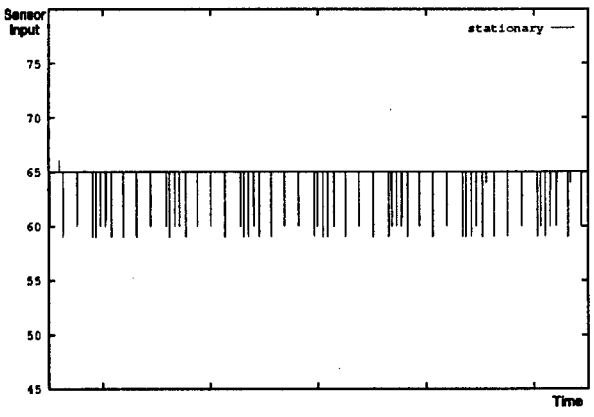


Figure 5. Acceleration sensor for stationary device.

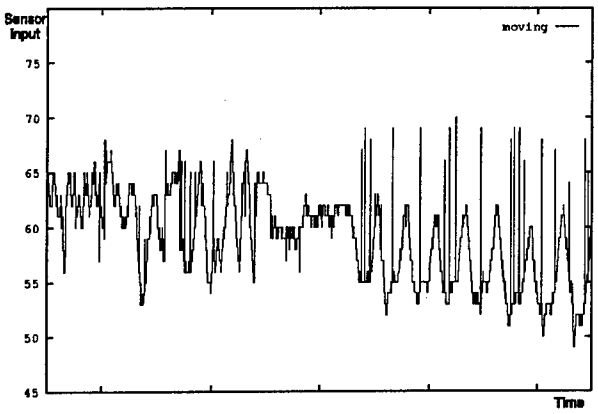


Figure 6. Acceleration sensor for moving device.

3.3.1 Cue extraction & context awareness

There also other sensors on the sensor board, such as the sensors of the temperature, the air pressure, the passive infrared and so on. Each cue is extracted from the data of one corresponding sensor with proper algorithm. In the figure 7, we can see a typical period data from passive infrared sensor when the user moves the device in hand (the X-axis represents the time & the Y-axis represent the value of the passive infra data). Using the sequence analysis algorithm, the cues leaving and closing can be recognized within one sampling cycle.

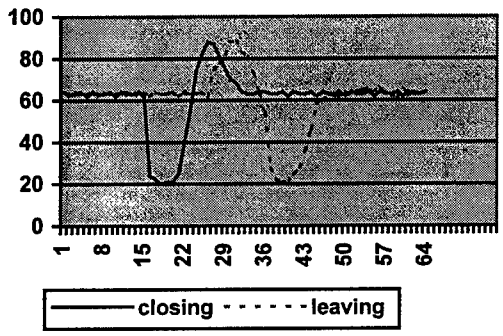


Figure 7. Passive infrared sensor for moving in hand

The data from some sensors, especially from light sensor, involves some random noises that usually occur with no more than two sequential values in one sampling cycle. Before analyzing the data from this kind of sensors, we suggest to use a mid value filter with 5-value-size window to do the preprocess.

Except for the cues extracted in time domain, the cue can also be the feature in frequency domain, for example the cue - base frequency. Base frequency represents the main frequency of oscillation of the light. The data from light sensor was transformed into frequency domain through FFT, and then used a linear window to find out the base frequency of in the date. This base frequency should be a stable value when there is artificial light near the light sensor.

Most of the awareness of the contexts is based on more than one cue and even other contexts. The cues and contexts are regarded as different dimensions of input vector of the fusion algorithm. Artificial neutral network and decision tree are investigated to fuse the input vectors into contexts. To describe the position of the mobile device, we employed three contexts: the device is in hand, the device is on the table, and the device is in a suitcase. The input vector has 15 dimensions, which corresponds with 15 cues from the sensor of gas (CO), temperature, pressure, light, passive infrared, and 2-dimensions acceleration. Automating the recognition, we used 297 samples (three classes, hand, table, suitcase; 99 vectors each) to train a neural network on them in a supervised mode. The other 297 samples were then used to test the recognition performance. With a standard back-propagation neural network we achieved a recognition rate of about 90 percent. Using a modular neural network, as described in [15], consisting of two input modules and on decision network we achieved a recognition rate of more than 97 percent.

3.2.2 Reconfiguration

The context "inside/outside" is used to describe the rough location of the host device is out door, inside of a building or a vehicle. The distinction of the inside and outside depends on the fusion result from the cues and contexts related with the light sensor and temperature sensor. The output data of the light sensor and temperature sensor are showed in the figure 3. Many cues are derived from the light sensor data in a standard period, such as the average brightness, standard deviation, base frequency, and so on. From the temperature sensor data, we get the cues: maximal and minimal temperature, average temperature. As showed in figure 8, two kinds of context are also useful to decide the context inside/outside.

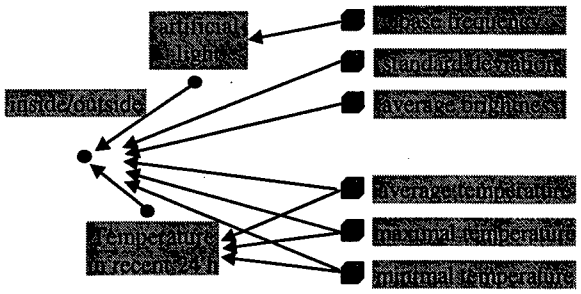


Figure 8. Deriving context inside/outside

The context "artificial light" indicates whether there are artificial lights in the current environment. The contexts "temperature in recent 24 h" describe the long-term statistic result of the temperatures in the past. We will simplify the decision process of "inside/outside" to show the reconfiguration of the awareness system.

In a normal situation, the decision tree of "inside/outside" is optimized by using the stored samples with all the attributes. In this decision tree, both the context artificial light and temperature related cues and contexts play important rolls (see figure 9).

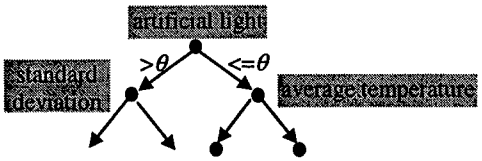


Figure 9. Decision tree for inside/outside

We discuss two reconfigure situations activated by disabling the context "artificial light" and switching off the temperature sensor. If the context "artificial light" is disabled by the host application, the decision tree has to be rebuilt according to the same stored samples but without the attribute "artificial light". The similar reconfigure process will also be done when the temperature sensor is switched off. The decision trees in these three situations can produce the recognition results, which are described in table 2.

Table 2. Recognition results

context	Total number of test samples	Recognition rate-normal	Artificial light disable	Without temperature
inside	512	93.0%	81.7%	91.2%
outside	512	98.0%	89.4%	87.6%

4. Conclusions & future work

The architecture presented in this paper is designed with a four-layer structure for multi-sensor fusion in mobile environments. The layered structure of the architecture allows the algorithms of the fusion system to be developed independently with sensors, data and the application demands. Through the interface defined between layers, the fusion algorithm face inputs with similar structure no matter whether they are real sensor data or the results of the other algorithms. The design of the layered architecture aims not only to develop the model to fuse the data from multi-sensor, but also to investigate the model to fuse the methods and techniques developed in the area of information fusion and other research area. Moreover, the layered structure makes it feasible to reconfigure the algorithms in each layer, which is important to enable awareness in mobile environments. The algorithms in the fusion system can be reconfigured properly to adapt the environment changes caused by the "movement" of the mobile devices, and produce more robust awareness results. Finally, the architecture keeps the interactions of host applications through different layers, which gives the opportunity for the host application to adjust the functions of the awareness device while also gives the chance for the fusion system to learn from the host.

Experimental results show that the awareness system we developed in this layer architecture performs robustly if all the possible situations of the mobile environment are known. If unknown situations occur in the environment, it is difficult for the system to produce the right and stable awareness results. The reason is that the awareness system can not find the new useful contexts in the environment by itself. Our future research will focus on application of data mining techniques in building the multi-sensor fusion system, which can adapt to unknown situations automatically. Furthermore, because the communication plays an increasingly important role in the application area of mobile devices, techniques for fusing the information from sensors with the information from communication channels will be investigated in our future work.

5. Acknowledgement

The research described in this paper is supported by the EC under the ESPRIT program, project TEA. We would like to thank people at TecO, Starlab Nv/Sv, Nokia mobile phone, and Omega Generation for many discussions surrounding this work.

References

[1] G. Reynard, S. Benford, C. Greenhalgh, & C. Heath, "Awareness driven video quality of service in collaborative virtual environments", in Proceeding of ACM CHI'98, pp. 464-472, 1998.

[2] S. Bly, S. Harrison, and S. Irvin, "Media spaces: Bringing people together in a video, audio, and computing environment", Communications of the ACM 36(1), pp. 28-46, 1993.

[3] Esprit Project 26900, Technology for enabling Awareness (TEA), www.omga.it/tea/, 1998.

[4] B. V. Dasarathy, "Information/Decision fusion – principles and paradigms", in Proceeding of the workshop on Foundations of Information /Decision Fusion, pp. 46-60, 1996.

[5] B. V. Dasarathy, "Sensor fusion potential exploitation – innovative architectures and illustrative applications", in Proceedings of the IEEE, pp. 24-38, January 1997.

[6] R. Luo & M. Kay, "Data fusion and sensor intergration: state-of-the-art 1990's", Data Fusion in Robotics and Machine Intelligence, pp. 7-136, 1992.

[7] K. Aizawa, Y. Egi, T. Hamamoto, M. Hatori, and M. Abe, "On sensor image compression for high pixel rate imaging", in Proceeding of IEEE international conference on Multisensor Fusion and Integration for Intelligent System, pp. 201-207, 1996.

[8] H. Kabre, "Performance and competence models for audio-visual data fusion", SPIE international symposium on intelligent systems and advanced manufacturing, vol. 2589, pp. 100-107, 1995.

[9] S. G. Goodridge, and M. G. Kay, "Multimedia sensor fusion for intelligent camera control", in Proceeding of IEEE international conference on Multisensor Fusion and Integration for Intelligent System, pp. 655-661, 1996.

[10] R. Bajcsy, G. Kamberova, R. Mandelbaum, and M. Mintz, "Robust fusion of position data", in Proceeding of the workshop on Foundations of Information/Decision Fusion, pp. 1-7, 1996.

[11] B. E. F. MacLeod, & A. Q. Summerfield, "Quantifying the contribution of vision to speech perception in noise", British J. of Audiology, Vol. 21, pp. 131-141, 1987.

[12] R. R. Brooks, and S. S. Iyengar, "Multi-Sensor Fusion - fundamentals and applications with software", ISDN 0-13-901653-8, Prentice Hall PTR, 1998.

[13] N. S. V. Rao, "Nadraya-Watson estimator for sensor fusion", Optical Engineering, vol. 36, pp. 642-647, 1997.

[14] A, Schmidt, J. Forbess, "What GPS Doesn't Tell You: Determining One's Context With Low-Level Sensors", the 6th IEEE International Conference on Electronics, Circuits and Systems, September 5 - 8, 1999.

[15] A. Schmidt, Z. Bandar, "A modular neural network architecture with additional generalization abilities for large input vectors", in Proceeding of international conference. on Artificial Neural Networks and Genetic Algorithms, pp. 35-39, 1997.

A Distributed VIPD Architecture with Central Coordinator

Li Yinsheng Zhang Heming Tong Bingshu

CIMS/ERC, Tsinghua University, Peking, PRC

Abstract: The paper illustrates what is related between VIPD (Virtual enterprise oriented Integrated Product Development) and information fusion, discusses some limitations of the distributed architecture, proposes a distributed VIPD infrastructure with central coordinator, and demonstrates its central coordinator and global database. By avoiding end-to-end communication and enhancing system security, the architecture will be a good attempt for manufacturing enterprises to unite and reorganize dynamically. On the other side, it will provide a lot of experiences and inspirations in applying multi-source information fusion technologies into complex systems.

Keywords: VIPD (Virtual enterprise oriented Integrated Product Development), information fusion, central coordinator, global database, PDM (Product Data Management)

1. Introduction

1.1 Background

The paper is supported by a research project called "Mode and technologies on VIPD", which is sponsored by Chinese National High-tech Plan/CIMS topic, and deserves to work on the theories, technologies, and modes for implementing VIPD (Virtual enterprise oriented Integrated Product Development). In the project, a digital product model, with multi-sources information including design history, assembly data, market information, and so on to be built-in, will be constructed [4]. And some advisable experiences and a reference mode for mid-scale and small enterprises to implement dynamic organization, and collaborate from remote locations, will be proposed [1]. The infrastructure, which is described in the paper, will be the basic architecture to integrate engineering environments of the project.

1.2 Arrangement

The paper proposes and illustrates an integrated infrastructure for VIPD, which makes advantages of both central management and distributed computation, and can integrate various polymorphous information derived from all the members of virtual enterprise and their subject product. With its central coordinator, it is more secure for member enterprises to exchange and share their data involved in product development. Moreover, the central coordinator will provide a facility for integrated product teams to collaborate their engineering transaction [10].

It is by seven sections that the paper illustrates the architecture for VIPD. In the next section, we will give an introduction to virtual enterprise and VIPD, which is helpful for readers to understand the infrastructure in the paper. In the third section, we will discuss information fusion in VIPD, which will do readers a favor to get clear for how VIPD will benefit information fusion theory. A simple development history of integrated modes will be given and the infrastructure for VIPD will be introduced in the forth section. After that, the theory and global database of the integrated architecture will be illustrated, and its features and advantages will be concluded. Finally, the paper will be ended with a

conclusion and the references.

2. Virtual Enterprise and VIPD

2.1 Virtual Enterprise

With the shorter and shorter product life span, enterprises are required to be enough flexible and agile to meet with smaller and smaller batch of orders. On the other side, however, it is very difficult for a single enterprise to get all the necessary technologies and resources for that. As a result, virtual enterprise (or virtual organization) emerges [2]. Virtual enterprise is one of the perfect styles for agile manufacturing. And what identifies its idea is as follows: in order to develop a new aborative product and win a market battle, some enterprises, which possess necessary resources and technologies for designing, manufacturing, and marketing the product, will make up a occasional union to cope with their rivals together. They select what they are good at and organize them to be a new enterprise, and make it as their behalf for profits [1].

All the components of a virtual enterprise are independent, self-determined, self-organized, and self-optimized, and they generally collaborate with each other and perform as coordinates. Besides, the members are often distributed in different locations, and can take part in more than a single virtual enterprise in the same time [2]. Another standout feature of a virtual enterprise is that a product opportunity determines its presence: when the opportunity occurs, it will be organized quickly; and when the opportunity fades away, it will be disjoint in the same speed.

It is due to their independent feature that many members of a virtual enterprise would be using various product development subsystems, which are usually polymorphous with each other. However, in order to have consistent and common product data so that they can collaborate in development, the subsystems are urgent to get integrated and harmonized [8].

2.2 VIPD

As we know, traditional CIMS (Computer Integrated Manufacturing System) and CE (Concurrent Engineering) have contributed a lot to information and process integration [5][9]. However, in product development, with more and more enterprises becoming dynamic organizations, many new problems, especially those about integration among multi-enterprises, have appeared and expected to be worked out [7]. As a result, VIPD is paid attention to and expected to benefit the problems [1].

VIPD is one of the key technologies for implementing virtual enterprise, which attempts to integrate both the utilities and product development processes into a uniform computer integrated engineering environment. And by drafting, designing, manufacturing, testing, and analyzing the product in a consistent environment, concurrent design and global integration among enterprises can be achieved based on information integration, function integration, and process integration. As a result, development will become faster and performances for concurrent and manufacturability in design can be fulfilled better [2].

Some objectives of VIPD are concluded as follow. Firstly, VIPD will expand its subjects from a part to a complete product, which will pay emphasis on constructing 3-D product model that merging assembly information into [4]. In the second place, VIPD will take up with developing the technologies to support remote collaboration among members of the product development teams, all of which are based on the distributed databases and network systems. Thirdly, VIPD will benefit integrating distributed PDMs (Product Data Management) on the condition of the wide-area network. Finally, as we know, when virtual enterprises get changed, the related supporting environments and technologies should be adjusted accordingly, and which in turn

requires those technologies in VIPD to be open and standardized.

3. Information fusion in VIPD

3.1 Information fusion in CIMS

As we know, CIMS is one of the typical complex systems, in which many kinds of data or information, such as those about administration, enterprise, organization, product, financial affairs, etc. will be dealt with. Considering their different features, formats, and sources into, not only different utilities, media, and devices will be used to edit, collect, or store them, but also different methods will be required to tackle with them. All of them, as we know, will resort to the theories and methods about information fusion. As a conclusion, the fusion (integration) of multi-source information will be one of the basic research topics, and will run through CIMS.

3.2 Information fusion in VIPD

VIPD is an advanced technology in CIMS and a landmark in integration. Fundamentally, VIPD results from the rapid progress in information and communication technologies. And when constructing its integrated environments, both the strategies of remote collaborative development, and the features of the product data in open, generality, and interchangeability between engineering fields or developing phases, including market decision, policies making, design, manufacturing, and maintenance, will be emphasized on. As a result, the information sources and supporting technologies in VIPD will be much more complex than those in integrating a single enterprise will. In the next paragraphs, six aspects will sum up several kinds of information multiformity in VIPD, together with their advisable solutions.

(1) The members are different in location. As we know, there are not any limits in position for the components of a virtual enterprise. Generally, by remote communication based on Internet, information exchange and sharing can be conveniently achieved between them.

(2) There exists an evident unconformity between the structures for member enterprises, especially in their resource organizations. One of the main causes for the unconformity is that these enterprises are specialized in different phases of making product, and they are often particular to each other. Consequently, it is necessary for VIPD to consider all the characters of different organizations in establishing security mechanisms and charts for examining, approving, and releasing the changes and proposals.

(3) There are often inconsistencies in rules for the members to instruct their engineers how to design products and by which standards and criterions to constrain their designs.

(4) There are many differences in the carriers of product data. As we know, product data can be stored by the data files (for example, documents and CAD/CAPP/NC files), databases (for example, metadata), and even hard copies. Moreover, there is severe unconformity in the database systems (for example, object-oriented database systems and relational database systems) and data editors (for example, MS Word contrast to VI, PRO/E contrast to UG, etc.). Therefore, in order to exchange and transform between them, it is necessary for engineers to develop various interfaces and front-ends.

(5) Product structures, BOM (Bill Of Material) reports, product management data, and flow charts for developing product, which will be dealt with in VIPD, are polymorphous. A popular resolution for the problem is to develop the corresponding transform forms.

(6) The supporting environments, including networks and database management systems, are polymorphous in a virtual enterprise. Owing to the different history of development environments and different enterprise backgrounds, software and hardware of the networks and database management systems are usually different and even non-compatible.

Fortunately, many providers of PDMs (Product Data Management system) have paid a lot attention to the problem and provided many commercial PDM systems, which are packed together with WWW servers, and can support across-platform navigation and operation [10].

All of the above make an indubitable conclusion that VIPD can be studied as a typical technology and practice for multiple-source information fusion, and the technologies and theories for information fusion are also basic supporting elements for implementing VIPD. In another words, research on information fusion will benefit VIPD a lot; and on the other side, study on VIPD will become one of the important branches and tendencies of information fusion. As will be illustrated in the next sections, the infrastructure for implementing VIPD can't only settle how to share and exchange polymorphous product data securely between the components of a virtual enterprise, but also support across-platform interoperation by WWW technologies [10]. Accordingly, we can declare that it isn't only a good resolution for information exchanging and sharing, but also a practical reference mode for information fusion of multiple-source polymorphous.

4. Limitations of the distributed structures and its solution

4.1 The history of integration structure

In the case of traditional central control, a failure from the master would make the entire system to break down, and which can often cause a terrible loss. As a result, distributed control systems appear and get used extensively. It is very evident that the distributed environments possess many advantages over the central control systems. For example, in the former case, each subsystem is independent of each other, and in the event that one of them goes wrong, there isn't any impact on other systems. As a result, distributed structures have got popular rapidly, and many of their modes, such as those based on agents, have been devised in the last decade.

4.2 Limitations of the distributed systems

As we have seen, distributed structures have been dominant in the last decade. However, with rapid growth in commercial hardware and software for information and communication, limitations of the distributed structure are becoming more and more evident. And in the event of collaborative development in a virtual enterprise, the conclusion is especially true. The next paragraphs will give some details for it.

Firstly, security is one of the critical difficulties in a distributed structure. As we know, all the components of a virtual enterprise are independent of each other, and it's only for commercial profits that they would get united. Generally speaking, because they often engage in some similar businesses, there are undoubted competitions between them. However, when product data are exchanged in VIPD, some protocols are inevitably required by a distributed system to get correct information. Therefore, it is a popular method to construct agents in all sites of the virtual enterprise. And in order to communicate and inter-operate with other comates, it is required that every agent be aware of information in all the other subsystems. However, on the other side, the agent is usually fully accessible to the local administrator. Consequently, component enterprises can't be assured that their product data are enough secure. As a conclusion, the VIPD environment can't be established smoothly without an active participation of the components

Additionally, to a distributed environment, there are many disadvantages in online updating between its subsystems. For example, when one of them gets changed or omitted, or a new one takes part in, it can't be assured that others be updated in time. Especially in the cases that servers with the agents built in are started and shut up frequently, it's difficult for them to keep

pace with. In such a condition, product data can never keep consistent between the member enterprises.

4.3 Prospects for the structure with central coordinator

In the case of central management to a virtual enterprise, because supporting systems for information exchange and sharing are based on Internet and independent collaborative subsystems, a centralized control architecture can seldom leads to a tragedy. An argument is that the integrated system is an enhanced environment, which is specialization and reorganization for random and anarchy data exchange and sharing, and one of its key functions is for collaboration between remote engineers. And in the event of central server failure, most of engineering activities can be continued in individual sites. After all, communication is not a continuous operation.

Just for the above consideration, most people are turning their attention around to central systems once again. In fact, under the condition of a virtual enterprise, because there is always a master member to head the union, which generally goes ahead and have an excellent leading power, it's probable and feasible to establish a single powerful server to control and coordinate all the communications between members.

Besides, with more advanced technologies and more reliable devices, disadvantages and risks from central control are lessened quickly. Accordingly, in our research project, by probing into the existing development environments, analyzing the motives and uniting modes for implementing virtual enterprise, and considering into the requirements and impacts of a virtual enterprise on the establishments, enterprise cultures and social settings, we conclude a integrated infrastructure for VIPD. The structure is based on commercial PDM systems [10], Internet, browser/server, and central coordinator, makes advantage of both central management and distributed computation, and is helpful to improve the competitive capacity of Chinese manufacturing industry.

5. The infrastructure for VIPD and its coordinator

5.1 PDMS in VIPD

As has been approved, involved with basic and complex development technologies and terrible workloads, it is acknowledged as a poor way to construct integrated architecture for product development by the foundational network and database systems. Fortunately, with Internet, object-oriented, and digitized technologies developed rapidly, PDM systems have been provided and used for integrating product development supporting subsystems.

PDM is one of the leading supporting technologies for concurrent engineering, which can be used to manage what are related to products (including information for components and parts, product structure configuration, documents and archives, resource organization, and security) and workflow for changing and releasing of item revisions. By taking product structures, development processes, and designers into a uniform platform, PDM can avoid those problems about versions, privileges, and data redundancy [6]. Its essential goal is to make a right person to receive right data and achieve a right task in a right way, right time, and a right location.

Current practices have showed that it's just a right way to apply PDM systems into VIPD. Moreover, to make use of PDMs' WWW servers and their client/server architectures and take it as a foundational supporting platform, can quicken and simplify the process for implementing VIPD. Moreover, the practice will show PDMs an improving way in the same time.

5.2 Introduction to the infrastructure

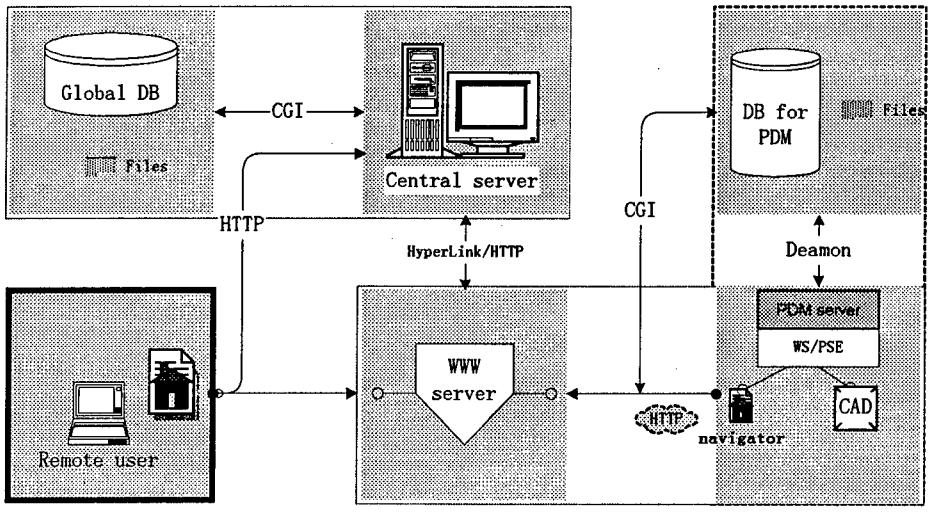


Figure 1 The architecture for VIPD

As is illustrated by figure 1, to construct an integrated environment for VIPD, every component of virtual enterprise is firstly required to possess an integrated subsystem for their developments. All of the subsystems are based on PDM systems and customized with WWW server and communication interfaces, so that any valid users can navigate and operate product data and applications by the allocated privileges in their browsers. In addition, a logical central server will be specified, with its global database and WWW server built in. In such an environment, product data, rules, and descriptions of subsystems can be stored and maintained in a global database, and bi-directional transformation and communication services from a global product structure model to various data views can be provided. Moreover, any valid user, wherever he comes from and locates at, can access the central server by his privileges from his sites.

In such a mode, the central server can provide a powerful coordinator for integrated product development in virtual enterprise, with global databases and file systems in it. And a local WWW servers in a member enterprise can't only deal with its local transactions normally, but also update in time with central server. In this case, there are explicit links between the local product data and items in global databases, and the other data in a subsystem won't be impacted at all.

With the architecture, self-determination of a member enterprise can't be disturbed anyway, and its distributed transaction can also function as well as anywhere. In addition, one of its most important advantages is that the communication functions of a commercial WWW platform can be inherited by using a local WWW server. As a consequence, workloads in developing communication interfaces will be reduced a lot, and the generality, modularization, and standardization of a subsystem will benefit.

5.3 Global database and coordinator

It is a general global product database that describes the subject product of a virtual enterprise in all ways and is used as an essential facility for exchanging and sharing between the subsystems. And status, configuration, and site information for the member subsystems are all stored in the global database by standardized forms. By the database, some necessary services can be provided during the interactive activities between the members. These services include name service, query service, schedule service, transform service, and add and cancel service.

Generally speaking, by name service, a subsystem can call for the others with their names or IDs, instead of knowing their correct addresses. And by query service, a subsystem can ask the central server for what the other members can provide. In such a case, the server will search for the tables for related members and their

features and services, and once an item to match with the query conditions is found out, a response message will be sent back to the request. As for the schedule service, in the case that no member can provide any facilities for the request, it can generate an item in a space called "blackboard" to record the requirements. Taking into account that the other members can access the blackboard for the current requirements, when there is a member can provide a requirement, a response will be made and informs the corresponding request.

Another service, which will enhance the VIPD architecture, is transformation performance of the global database. It is by transitional forms that the service can assist different subsystems to achieve an exchange for their polymorphous product data, such as BOM reports, etc.

Finally, by add and cancel service, when a virtual enterprise wants to get a new member or cancel an existent one, what it will have to do is to change the lists in the global database.

5.4 Features

In the architecture, the WWW servers and their customized services are very similar in all members, so a central server with its global database is the only key facility for VIPD. With ODBC and across-platform programming languages, all components can be standardized easily. As a conclusion, the architecture is good at compatibility and practical in implementing.

Under presence of the central server, subsystems can contact the central server directly for a communication, instead of communicating with their destinations in point-to-point mode. Besides, by the infrastructure, when an enterprise want to join, what it is required is to customize a little for its PDM system and publish information to the central server by the procedures, so that the latter can update its status, rules, security and privilege tables in time. Because its description lists can be updated online, items can be added or changed at any time. As a result, it is more convenient to add or cancel a new member or applications and change meta-information for a product.

Accounting that the members can communicate with each other by only contacting with the central server, end-to-end interaction can be avoided and the system will become more secure and practical. To sum up, we can declare that such a system is open and compatible, and can take full advantages of both the distributed computation and the central management.

6. Conclusion

The paper identifies virtual enterprise, VIPD, and PDM, illustrates how VIPD and information fusion theories impact each other in their system architecture. Additionally, the limitations of both the central and the distributed architecture are discussed. Afterwards, an open and compatible infrastructure for implementing VIPD is proposed and its coordinator is demonstrated. Based on commercial PDM systems and browser (client)/server, the architecture can take full advantages of the distributed computation and the central management. Moreover, by avoiding end-to-end communication between the members of a virtual enterprise, the integrating mode based on the architecture can reduce what to do for configuring subsystems, enhance system security, lessen data transfer, benefit information sharing and exchange in a virtual enterprise. As a conclusion, it will be a good way for Chinese manufacturing enterprises to unite and implement virtual enterprise. And on the other side, the architecture will provide many useful experiences and inspirations for applying multi-source information fusion technologies into complex systems.

7. References

[1] Zhang Heming. Advanced manufacturing oriented integrated product development and its technologies. postdoctoral research

summarization in Tsinghua University, 1997.12

[2] Wang Qinmo. Agile manufacturing and virtual enterprise. Chinese Machinery Engineering. Vol.8. No.4:51-56, 1997.8

[3] BJØRN P.MUNCH. REIDAR CONRADI. etc. Integrated Product and Process Management in EPOS. Integrated Computer-Aided Engineering. 3(1):5-19(1996). Vol.3. No.1, 1996

[4] Ram Anantha. Glenn A Krama and Richard H Crawford. "Assembly Modelling by Geometric Constraint Satisfaction". Computer-Aided Design. Vol.28. No.9:707-722, 1996

[5] Research team on feasibility of integrated frame for concurrent engineering. Report on feasibility of key technologies for integrated frame in concurrent engineering, 1996.12

[6] Hannu Peltonen. Olli Pitkänen. Reijo Sulonen. Process-based view of product data management. Computers in Industry 31(1996):195-203

[7] Robert de Graaf. Luk Kornelius. Inter-organizational concurrent engineering: A case study in PCB manufacturing. Computers in Industry 30(1996):37-47

[8] Hyunbo Cho and Mooyoung Jung. Enabling Technologies of Agile Manufacturing and its Related Activities in Korea. Computers Ind. Engng. Vol.30. No.3:323-334, 1996

[9] Merle Thomas Jr. Concurrent Engineering: Supporting Subsystems. Computers Ind. Engng. Vol.31. No.3/4:571-575, 1996

[10] Li Yinsheng. etc. Integrating Polymorphous PDMs. CADDM. Vol.8. No.2:50-56, 1998

Agent–based Information Processing System Architecture

Luo Zhongyan
Department of Automation
Tsinghua University
Beijing, P.R.China

Liu WenHuang
Department of Automation
Tsinghua University
Beijing, P.R.China

Liao Lin
Department of Automation
Tsinghua University
Beijing, P.R.China

Abstract

The fast development of information technology and rapid expansion of information demand have challenged contemporary information systems. This paper presents architecture of an information processing system with intelligence, coordination and adaptability. A prototype of typical news system is then implemented as an example. The principal techniques used in the system are analyzed and investigated in-depth and the architecture proposed is evaluated.

Key Words: agent, information processing system, system analysis and design, workflow

I. Introduction

We are in an information era with the development of technologies such as computer, network and database accelerating greatly the maturation of information techniques. Information has been regarded as one important resource and commodity with inestimable value. Information processing system is just the key element to produce and provide new information. Whereas, due to the existence of various unstructured information, information processing not only need routine manipulations of computer but also need the interaction between human & machine and the cooperation among people, especially on such aspects as contents, quality and responsibility. Many achievements have been acquired on partial techniques of information processing, but the relationship of each part and the whole architecture of information system have not been deeply investigated. Meanwhile, new information systems will face a lot of vital problems and opportunities, such as distribution and diversification of information, coordination and intelligentization of workflow management, diversity and individuality of users' need, and so forth. Therefore the adjustment and reconstruction of system structure and framework is inevitable.

The development of distributed artificial intelligence has provided us the ideas and methods of agents. In a multi-agent system, the autonomy, social ability, responsibility and pro-activeness of agents make them coordinate to accomplish systemic integrated functions, which is just what a new information processing system needs.

This paper presents architecture and design method of an agent-based information processing system, which consists of five subjects: information entry, information processing, information publication, information resource management and service, system management and decision. A typical news system is designed and analyzed to show agents' specific functions and communication protocols. Then the operation process and cooperative relation among agents are described and

Correspondence to: Luo Zhongyan, Liu Wenhuang, Liao Lin, email: liuwh@cims.tsinghua.edu.cn, Address: Automation Department, Tsinghua University, Beijing 100084, P. R. China.

analyzed. Further investigation and discussion are also made on the difficulties and key problems in system design such as workflow management and control, inheritance and polymorphism of agents, and the cooperation and fusion among agents.

I. Agent-based information system architecture

1. Basic concepts and structure of agent

The internal structure of an agent is first given in figure 1:

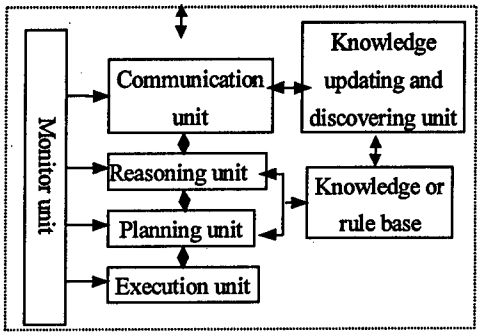


Fig.1 General agent structure

Communication unit: This unit receives, sends all kinds of information; accomplishes information intercommunication among agents; provides interface and communicates with outside environment.

Reasoning unit: Based on contents in knowledge or rules base, it makes reasoning on information, examines if it is valid and realizable, and makes corresponding message respondence.

Planning unit: It schedules undertaken tasks according to capability of each agent and informs execution units.

Execution unit: This unit executes and accomplishes some kind of function based on the plan designed by the planning unit.

Monitor unit: The unit monitors internal states and task executions.

Knowledge updating and discovering unit: It discovers new knowledge and rules from outer messages and former work summary; receives instructions from superior agent to expand knowledge or rule base.

Knowledge or rule base: store contents relevant to agent functions, message grammar, semantic knowledge and rules.

A fairly complete internal structure of agent is already explained. In practical system design and implementation, it is necessary to simplify or strengthen some units.

2. Cooperation and communication mode of agent:

Many methods of cooperation among agents have been investigated, and here we will discuss the method of registration table (figure 2):

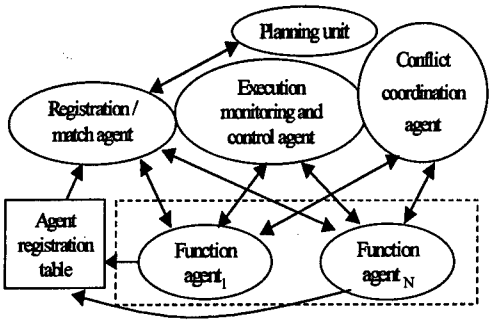


Fig.2 Cooperation among agents based on registration table

Planning unit: Through the cooperation with registration / match agent, it acquires tasks and divides them into some executive small ones.

Registration / match agent: This agent manages and maintains function agents and agent registration table; receives tasks divided by planning agent, matches tasks according to registration table and distributes them to each function agent.

Execution monitoring and control agent: It monitors and controls each function agent executing tasks.

Conflict coordination agent: The agent coordinates and solves problems when diverse function agents' goals conflict.

Communication among agents in the system will adopt the communication mode of dividing grouping blackboard based on agent subjects (figure 3)

3. Architecture of agent-based information processing system

Since a large-scale information processing system contains a great number of agents, if we

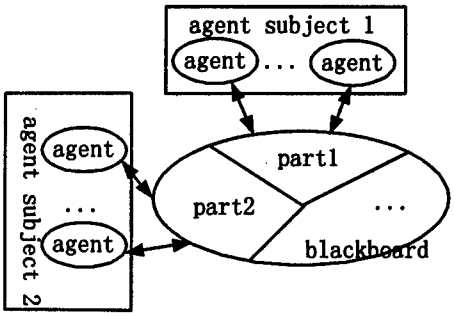


Fig.3 Communication mode based on blackboard

create these agents without classifying, it will induce chaos of system management, increase the difficulty of system maintenance and aggravate the load of system operation and communicating. Accordingly, we adopt the idea of grouping (subject dividing) while designing the system framework. According to the logical relations and functions of agents in the system, they are divided into some subjects. Each agent subject maintains its relative independence on logic and function. And they accomplish the information processing cooperatively through the interchange of data information and transmission of control information (figure 4).

General functions of every part of agent-based

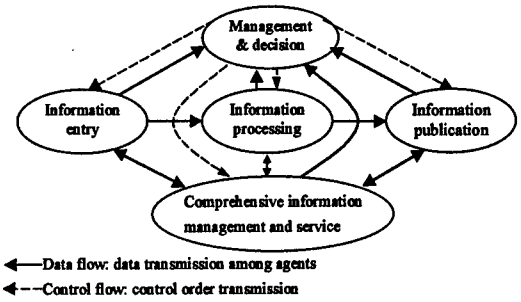


Fig.4 Structure of agent-based information processing system

information processing system are as follows:

1) Agent subject of information entry: These agents collect original information for the system through various channels and ways, arrange and classify information and provide materials for further processing.
2) Agent subject of information processing: On the basis of certain processing mechanism, such agents process original information and generate information products. Agent subject of information processing is the kernel of the whole system. And the quality of the information products is determined by its functions. However, the design of information processing subject has close relation with the processing mechanism and operation mode of the system, which will be exemplified in detail later.
3) Agent subject of information publication: These agents publish and distribute information products to outside, manage relevant transactions, provide information services for environment and collect feedback information from users.
4) Agent subject of comprehensive information management and service: Such agents manage comprehensively system information resources and intermediate information while processing, provide convenient and efficient services of storage and inquiry for the system.
5) Agent subject of management and decision: The agents administer the whole information processing system in a high level, analyze data synthetically and control the operation tactics of the system.

II. Agent-based news system

I) Background and logical structure of Computer Integrated News System

We have implemented Computer Integrated News System (CINS) for Science & Technology Daily office, a good-sized newspaper office. CINS will enhance the management level, competence and adaptability all-around. It integrates such systems as collecting and editing, manuscript delivery, typesetting, printing and publishing, and distributes news quickly. Moreover, it can obtain feedback information from users and demand information from outside in time, control and adjust system strategies of information collection, processing, distribution and newspaper publication. We will introduce the specific design and structure of agent-based information processing system using this typical news system as an example.

I) Agent-based system structure (5 parts)

1. Agent subject of information entry

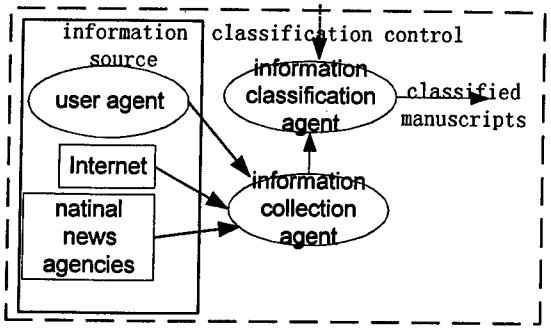


Fig.5 Subject of information entry

- User agent: So-called system users include information source provider, information processor, information user and information manager and so on. Therefore, user agents not only possess general functions of information interaction, but different agent subjects have different specific functions. First we will show elementary functions:

a) User interface: provide fast and easy operational interface, receive users' inquiry demand and display inquiry results.

b) Examine validity of users' input and translate vague, incomplete demand to standard communication description of agents.

c) Communicate with the agent of comprehensive information management and service.

d) Provide functions and communication ways relevant to agent subjects.

Users: domestic and overseas correspondents, free contributors.

- Information collection agent: It collects raw materials for the whole information processing system; receives and processes manuscripts from everywhere, including reports and news from domestic and overseas correspondents, free contributors, national news agencies and government-relating institutions; searches relevant information from Internet.

Functions:

a) Receive and process regularly or irregularly, routine or mobile transferring manuscripts.

b) Search information on Internet.

c) Submit raw information materials to information classification agent.

- Information classification agent: It classifies preliminarily raw information materials on hand and prepares for further processing.

Functions:

a) Receive raw information materials from information collection agent.

b) Classify raw information according to relevant knowledge and rules.

c) Store information about classification results.

d) Notify information processing agent of new materials.

e) Adjust classification based on control information.

2. Agent subject of information processing

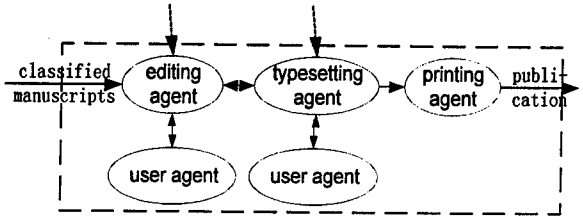


Fig. 6 Subject of Information Processing

- Editing agent: The agent filters out raw materials; edits, pre-signs and reviews selected manuscripts by some rules; completes the processing of information contents.

Functions:

a) Receive classified manuscripts.

b) Select manuscripts.

c) Edit, preview, review and sign manuscripts following designed procedure.

d) Save edited reports and notify typesetting agent.

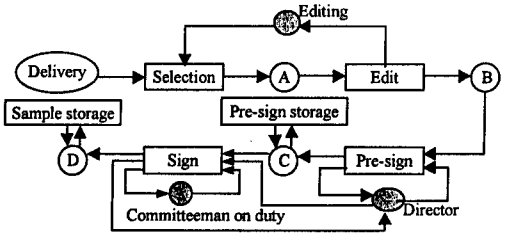


Fig.7 Editing agent workflow

- Typesetting agent: This agent is responsible for designing layout of edited manuscripts and completing typesetting.

Functions:

a) Receive signed reports.

b) Typesetting.

c) Modify layout in cooperation with editing agent.

d) Save layout information and inform printing agent.

- Printing agent:

Functions:

a) Receive layout information file.

b) Adjust layout, make films and PS2 format file by laser scanning.

c) Printing.

- User agent: It supports relating operators to do information processing.

3. Agent subject of information publication

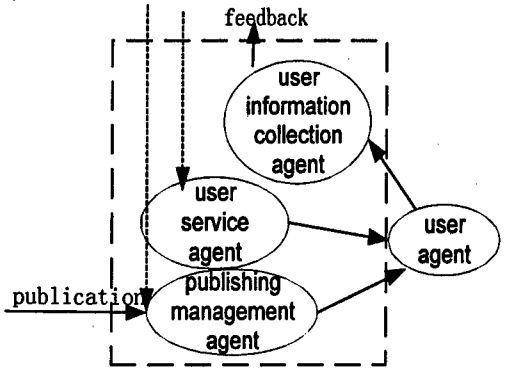


Fig.8 Subject of information publication

- Publishing management agent: This agent distributes or publishes final information products (including press publication and electronic publication), manages information and publishing process.

- User service agent: It provides publication information and information inquiry service for users.

- User information collection agent: The agent collects order information and feedback from users.

4. Agent subject of comprehensive information management and service:

- Service management agent: This agent receives service demands; for those demands it can handle, it divides tasks and does planning, collates results and returns them to demand agent.

Functions:

a) Receive demands from other agents, and examine their validity.

b) Accept reasonable demand.

c) Do reasoning and divide tasks to the other two agents.

d) Optimize return results and send them to demanders according to standard agent communication mode.

e) Manage dynamically data access agent and format conversion agent.

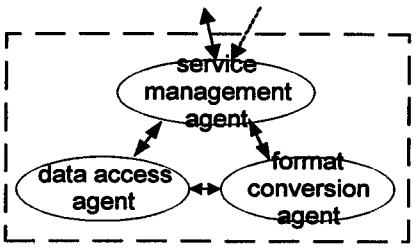


Fig. 9 Subject of comprehensive information management and service

- Data access agent: This agent provides services of inquiring and storing data information from different databases.

Functions:

a) Receive demands from service management agent; inquire and store data from databases.

b) Transfer relevant data to format conversion agent and ask for unique format.

c) Operate on data results after conversion and send results to service management agent.

d) Monitor changes of data sources dynamically.

- Format conversion agent: it takes charge of converting different types of data, including structural data conversion and multimedia data conversion, to standard formats.

Functions:

a) Receive data conversion demand from other agents.

b) Examine validity and give reply.

c) Convert data format.

d) Return converted data information to demanders.

5. Agent subject of management and decision

- User agent: It supports daily management for managers and decision-makers; supports decision discussion.

- Data analysis and information fusion (DAIF)

agent: This agent arranges and analyzes internal and external information, does information fusion, summarizes and discovers useful rules in or out of systems and provides evidences for management and decision.

- Cooperative DSS (CDSS) agent: It organizes management agents to have meetings and discussions, coordinates work at a high level and establishes coordination strategies.

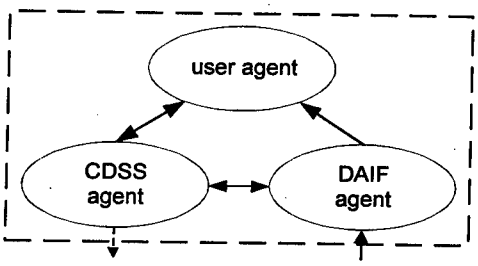


Fig.10 Subject of management and decision

IV. Further discussion

• Management of workflow and virtual edit department

The definition of workflow:

The workflow in information processing system refers to the whole work process from collecting and arranging raw information materials to distributing information products and providing relating services. Since information processing is the kernel part, here we mainly discuss the management and control of workflow in that module.

Usually the workflow of a news information system consists of three stages as collection & editing, typesetting and printing. And each stage can be divided into specific processing procedure (as described before). Since main resources (except hardware) consist of information resources and human resources, the management and control of workflow should pay attention to the following three points:

1. How to customize information-processing flow to guarantee completing tasks efficiently and qualifiedly.
2. How to allocate human resources in the information processing flow to achieve efficient, timesaving and low-cost operation of the system.
3. How to define the responsibilities of personnel in the system, i.e. role definition.

Virtual edit department:

Usually a newspaper office consists of some edit departments, which are responsible for different types of news reporting respectively. But to some sudden, important news incidents, it is necessary to select personnel from relating departments and form a temporary organization. Such organizations are called virtual edit departments because of their high time demand and short lifecycle. And to guarantee completing such tasks as reporting sudden incidents quickly and accurately, there should be some difference on the organizational and operational mode.

Through redefining the role of virtual edit departments and customizing workflow, it is possible to simplify procedures of manuscripts' processing, reduce processing time and improve the timeliness. A description of virtual edit department workflow is given below:

In virtual edit departments, correspondents' manuscripts will be delivered directly to editors for selecting and editing without classifying. After they have been modified, the director is responsible for finalizing and signing and then the manuscripts are stored in sample depository for typesetting and publishing. After we simplify procedures of manuscripts' processing, system efficiency will be improved and distribution time will be shortened. But it is necessary to redefine personnel's roles and it also means the augmentation of personnel's rights and responsibilities. For example, since customary procedures of pre-signing and finalizing manuscripts

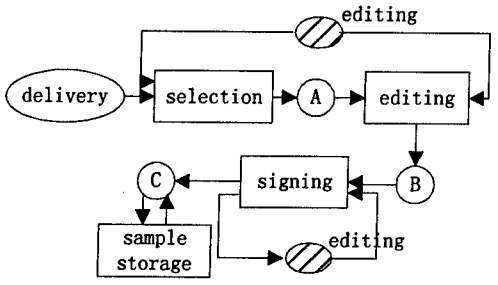


Fig.11 Redefinition of workflow

are replaced by once signing, the director should be more capable and take more responsibilities.

- **The cooperation and function fusion between editing agent and typesetting agent while doing typesetting**

Since the agent-based information processing system in this paper is a MAS (Multi-Agent System), the cooperation among agents is very frequent, which is also the basic requirement to accomplish the whole function of system. However, since the goal and evaluation standard of each agent may differ, it is possible to cause conflicts among agents and increase difficulty of cooperation. Quite a few papers have investigated the problem of cooperation among agents on theory. Here we only explain it in a practical system, then present and discuss the method of agent fusion.

In examples before, three steps are necessary for manuscripts processing as collecting and editing , typesetting and printing. Each time manuscripts are signed by the editor, they are delivered to typesetters. Since it is usual to modify manuscripts repeatedly, typesetters will redo their work time after time. However, the evaluation methods of editing and typesetting are different. For editor, less modification times means higher quality of editing. But for typesetters, to achieve better visual effect usually needs more modification times and more work. Editors expect less modification but typesetters want more. Therefore goal conflict appears between editing and typesetting on manuscript modification.

We present agent fusion to solve goal conflict discussed above. Usually after the first time manuscripts are read and edited, the contents and size of articles will not change much. That is to say, the modification of manuscripts is only local adjustments on layout. Therefore, we combine the work of re-editing and re-typesetting and let it done by only one person, not two as usual. In system architecture, the work done by two agents are accomplished by single agent, which will produce agent fusion. Since the contents of reports are significant and local adjustments of layout are comparatively easy, the work of re-editing and re-typesetting can be assigned to collecting and editing agent. And typesetting agent will confirm the modification. The problem of cooperation under goal conflict can thus be solved. It should be noticed that agents' functions and rules are adjusted too. Another advantage of agent fusion is reduction of workload and number of typesetters.

- **Base class and polymorphism of agent**

In a MAS, a group of agents may have identical functions, but to other similar functions their focuses differ. In order to construct, maintain MAS conveniently and decrease cost, we can borrow the concepts in Object Orientation (OO) technology and introduce the mechanism of inheritance and polymorphism for agent class.

Inheritance:

Among the agents in a MAS, part of the functions or knowledge in some agents may be identical. It is reasonable to create a base class with the common part and other agent can inherit those functions or knowledge from the base class. An inheritance tree forms after some levels of inheritance.

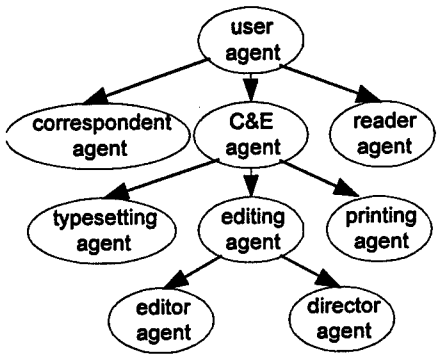


Fig.12 Inheritance tree

In the example above, all the user agents, such as correspondent agent, collecting and editing (C&E) agent, reader agent, share many common functions, such as basic interface support and information inquiry and so on. These functions and relevant knowledge can be sum up and form a base user class. The user class in each subject can be created through inheriting base user class. The agent inheritance tree of the example above is shown below.

Polymorphism:

Inheritance solves the problem of using same functions and knowledge among agents. But some kind of agents often have similar but not identical functions, which can not be solved by inheritance but need the assistance of polymorphism. Based on the mechanism of polymorphism, those offspring agents created by inheritance can modify and enhance functions, knowledge or rules of father agent according to practical need, which make individual offspring agent have different functions (or reasoning, planning etc.) under the same interface.

In the above example of news information system, as to the function of information inquiry, correspondent agent will get the information of all the relevant manuscripts, editor agent will receive those manuscripts being modified and be able to store results after each modification, and reader agent can acquire information of printing or electronic publication which is convenient to be browsed. These behaviors exhibit the polymorphism of information inquiry. To meet practical need, polymorphism is also introduced to the mechanism of reasoning, knowledge and so forth.

Inheritance and polymorphism are important features of OO technology. The introduction of these two mechanisms will help to simplify the concepts and structures of systems, make them easy to understand, and facilitate systems' building and maintaining. Inheritance in OO technology is limited within properties and functions, and polymorphism is mainly used in functions. However, inheritance and polymorphism in agent-based system has been expanded to include knowledge base and reasoning rules, which is a significant problem in the application of agent technology.

V. Summary

This paper presents the architecture of agent-based information processing system. It not only analyzes the functions of all the subjects and structural relationships among them in the view of the whole system, but also provides the agent technique to facilitate the realization of intelligence and adaptability. The entities of agents bring the object-oriented characteristics, which make the system more flexible and easier to build and rebuild. In the analysis and implementation of the news system, we describe basic analyzing approaches and communication protocols of agent-based systems. The investigation of workflow management provides the system with more flexibility, adaptability and better ability to deal with sudden affairs. The technique of agent fusion plus workflow management realizes the cooperation under goal conflicts and improves the efficiency and quality of the system. And the discussion of inheritance and polymorphism will facilitate the management and decrease the workload of developing and maintaining a system.

Acknowledgements:

We want to thank Yuan Yuan, Hong Jun and Zhang Yunyi for their sincere help with the paper.

References:

[1] Shigeru Fujiata. Agent-based Design Model of Adaptive Distributed System. Applied Intelligence. 9:57—70, 1998.

[2] Donna S. Haverkamp & Susan Gauch. Intelligent Information Agents: Review and Challenges for Distributed Information Sources. Journal of the American Society for Information Science. 49(4): 304--311, 1998.

[3] Liu Hong. An Information Process Modeling Approach Based on Agent. Computer Research & Development. Vol. 35, No.8 Aug, 1998.

[4] Xing Yanhui. United Use of Heterogeneous Database Based on Multi-agent System. Journal of Software. Supplement. Jun, 1998.

[5] Zhang Haohan. Cooperative Sketching System Based on Multi-agent Model. Journal of Software. Supplement. Jun, 1998.

[6] Liu Hong. A Constructing Approach for Mechanical Design Oriented Agent System. Journal of Software. Supplement. Jun, 1998.

Session RA3
Information Fusion for Decision Support
Chair: Galina Rogova
Calspan/CUBRC, NY, USA

Engineering a Model of the Subjective Judgement of Experts in a Data Fusion Application

Jane O'Brien[1] and Mark Bedworth[1]
Signal Processing and Imagery Department, DERA Malvern, UK

Abstract – *Airborne anti-submarine warfare includes fusion carried out by the Tactical Navigator when making decisions about the deployment of sonobuoys. This can be automated using a Bayesian Belief Network. However, the effectiveness of such a system depends on the accuracy of modelling the Tactical Navigator's decision-making process. This in turn relies on a practiced Tactical Navigator supplying the correct decisions to make and judgmental information on how these decisions would be affected by other factors. Such knowledge is only gained through experience and is difficult to quantify. Knowledge engineering methodologies and tools are available to aid with such knowledge acquisition and quantification.*

This paper will describe the airborne anti-submarine warfare problem and highlight the need for knowledge engineering techniques to develop a successful solution. It will provide a general background to knowledge engineering and describe methodologies, including CommonKADS, for carrying it out. The paper will then detail the application of CommonKADS to the development and implemenation of an automated decision-making aid for the Tactical Navigator.

Keywords: Anti-submarine warfare, Bayesian Belief Network, Expert, Information Fusion, Knowledge Engineering

1. Introduction

Airborne Anti-Submarine Warfare (ASW) is a complex military task where many decision-making problems cannot be explicitly solved either theoretically or by taking measurements [1]. One way to study the decision-making process in this case is by developing a computer simulation of the situation. An area where this has been evaluated is in assessing how the Tactical Navigator (TacNav) determines what action to take next when flying a mission. This has been tackled by developing an automated decision aid to the TacNav. The different information available to the TacNav indicates that this is a data fusion problem. In addition to providing an insight into operational problems, this aid can also be used to evaluate the possibility of fully automating the airborne anti-submarine warfare task.

The development of such a system depends on identifying the correct information to fuse. Since part of this information is encapsulated in the TacNav's experience the process is not as simple as it might be and depends on using knowledge engineering (KE) techniques. There are many KE techniques to use; this application has provided an opportunity for the tool, CommonKADS, to be evaluated.

Following this introduction, this paper will describe the ASW application, indicating where the fusion process takes place and highlighting the key modelling issue. It will then describe KE and provide an overview of CommonKADS. A description of the ASW case study using CommonKADS and its implementation will be provided, followed by conclusions on the use of KE techniques in general and CommonKADS in particular.

2. An Anti-Submarine Warfare Task

The aim of the ASW mission under examination is for a Maritime Patrol Aircraft (MPA) to detect and locate a submarine through the deployment of sonobuoys. As

[1] This work was carried out while the authors were employed at the Defence Evaluation and Research Agency, UK.
For communication purposes email {jane.obrien,mark.bedworth}@datafusion.clara.net

such this is a search and track mission which is deemed to be complete when a stable track is established.

Intelligence information provides the MPA crew with the target type and an indication of the submarine's location and, hence, the area to be searched. Sonobuoys are dropped in one of a set of patterns dependent upon factors such as the speed and direction of the target, presence and strength of previous detections, the type of the crew and its workload. The particular sonobuoy pattern used is the decision of the TacNav. A procedure manual of tactics regarding what action to take in any situation is used and a newly qualified TacNav will closely follow these rules. However, as he becomes more experienced he will start to apply his experience and adapt the rules to better fit the situation. Hence, there is no well-specified algorithm to determine what pattern of sonobuoys should be deployed in any given set of circumstances.

2.1 Fusion in the ASW Application

Anti-submarine warfare is a task that includes two levels of manual fusion. The first level takes place when the sonar operators declare a target detection based on information provided by a set of sonobuoys previously deployed. The second level of manual fusion is carried out by the TacNav who uses the qualitative detection information provided by the sonar operators with track data and the perceived crew work-load to make decisions about the next deployment of sonobuoys. The results of this latter fusion process may be inconsistent and dependent on the experience of the TacNav employed at the time.

The objective of this work was to combine the encapsulated experience of the TacNav with the rules provided by the tactics manual to provide a consistent advice tool.

2.2 The Key ASW Modelling Issue

The effectiveness of the system described above depends on the accuracy of modelling the TacNav's decision-making processes. Although the tactics regarding sonobuoy deployment are specified (for most circumstances) in a tactics manual, it has been found that the exact decision made will vary both between and within TacNavs. Thus the problem is one of modelling, not only the simple heuristics, but also the imprecise knowledge encapsulated in the mind of the experienced TacNav.

Various issues were identified, including the fidelity of the individual rules, the representation scheme used and the software implementation approach. It quickly became clear that a methodical approach to engineering the model as a whole was more important than the optimisation of these individual components.

The knowledge held in the TacNav's mind is a valuable asset and can be utilised in a disparate range of areas such as sonobuoy-use reduction, personnel training and mission optimisation. The TacNav provided judgmental information on how these decisions would be affected by external factors such as whether the crew was aggressive, whether or not the target had already been detected, the workload of the crew, etc. Such knowledge is not specified in the tactics manual and is only gained through experience.

Other knowledge is held in the tactics manual and records of previous missions. This makes it too disparate to be directly useful for our purposes. All of this knowledge needed to be collated and represented in a way that could be exploited to our advantage. It was felt that this stage would be usefully separated from the actual implementation.

The foregoing issues can be dealt with by KE, which is discussed in the next section.

3. An Overview of Knowledge Engineering

Knowledge within any organisation is commonly scattered between a number of personnel, documents and / or computer systems that may not even be located at the same site. Knowledge acquisition is the process of extracting knowledge from an expert. KE, of which knowledge acquisition is a component, focuses on the acquisition, modelling and management of this distributed fundamental domain knowledge, as well as any personal expertise.

KE covers a range of techniques including mathematical modelling, neural networks, genetic algorithms, knowledge-based and expert systems, data mining, natural language processing, intelligent agents, virtual reality, data visualisation and case-based reasoning. Expert systems are considered particularly beneficial.

As with anything else, there are advantages and disadvantages with KE. Disadvantages include a mistrust of the concept and hence little acceptance of the techniques. This is probably due to the fact that there is no proven track record in the field and that it appears to take a long time to develop anything usable. Another disadvantage is that there are very few knowledge acquisition and knowledge engineering

tools available. This means that development is usually conducted in a hybrid manner. On the other hand, advantages include an increased understanding of the processes under consideration at the end of the KE experience by both the expert and the knowledge engineer and a more robustly optimised process. Feigenbaum [2] summarises the three main advantages of KE as cost reduction, automated information processing and the gaining of new knowledge.

3.1 Approaches to Knowledge Engineering

The individual with the responsibility for collecting and structuring this knowledge and for developing the model with which to fuse it all is known as the knowledge engineer. A knowledge engineer has to elicit and manage large amounts of information-rich but ill-structured expertise data and needs a structured approach to help in this process.

A pre-requisite to any form of structured approach is a multi-disciplinary team. This has the advantage that a wider view of the knowledge available is obtained than if a single person is involved in the task.

The major tools for knowledge acquisition include interviewing, data analysis, text analysis, behaviour analysis and machine induction, the first two being the most popular. It is rare for one technique alone to be used in any knowledge acquisition task [3]. Some of these are briefly described below.

- Interviewing (learning by being told) provides information directly from the people with the knowledge and involves the knowledge engineer in studying verbal exchange, questionnaire responses, etc;
- Data Analysis is knowledge acquisition through analysing historical data records;
- Text Analysis is knowledge acquisition through the use of books, manuals, the internet, etc. It is a litle used method but has the advantage that access to a busy expert is not necessary;
- Behaviour Analysis (also known as learning by observation) involves the knowledge engineer observing the expert in action and the expert justifying his actions;
- Machine Induction theoretically speeds up the process by collecting information in the form of case studies. A computer extracts the appropriate information to produce the required knowledge.

Winston [4] defines the basic questions to be posed regarding knowledge as:

1. What kind of knowledge is involved?
2. How should the knowledge be represented?
3. How much knowledge is required?
4. What exactly is the knowledge needed?

3.2 Methodologies for Knowledge Engineering

It is often difficult to go directly from the elicited knowledge to an implemented system. One reason for this is the confounding of different types of knowledge, *i.e.* task knowledge and domain knowledge, making it unclear how the system ought to be developed.

There are prescribed methodologies for KE, the right one to use at any one time depends on the situation. A generic KE life cycle appropriate for predictable systems, with rigid specifications that allow fixed price development and a disciplined manner of progression, includes:

- feasibility study including assessing the scope of the system, determining which parts of the system should be knowledge engineered and which parts should be conventionally programmed, which techniques to use, software integration issues, determining data and information availability, appraising cultural issues and identifying appropriate experts;
- requirement specification including defining and validating the knowledge, data representation and maintenance requirements, agreeing the users expectation of the system and how they wish to interact with it, determining mandatory and desirable requirements and producing performance specifications;
- system design;
- module design;
- module coding;
- module integration;
- acceptance testing requiring the availability of test data sets. This also covers the problem of how to validate knowledge and how to test safety critical systems. (Incremental testing could help in overcoming some of these problems.) Acceptance testing requires awareness of the original scope of the problems and identification of the quality of the tests being carried out;
- commissioning is similar to testing with the added problems of resistance to new technologies by the intended users. Commissioning requires feedback from users to assess the implementation and expectation issues.

Each stage should be formally documented and signed off before proceeding to the next stage. This gives rise to extra administrative costs and additional time if it is decided at a later date that earlier stages need altering.

For flexibly specified systems with vague or uncertain requirements and outcomes, the above life cycle needs to be moderated accordingly, but feedback should be tightly monitored. Examples of tools to use with these less formal methods include Rapid Application Development (RAD) [5] and Dynamic Systems Development Method (DSDM) [6], an implementation of RAD. These are general iterative prototyping approaches to software development The idea of achieving a certain level of functionality within a fixed time period is not new, but these tools to facilitate it are.

3.2.1 CommonKADS

The two methodologies described above place the implementation of the acquired knowledge at the centre of the design process. This often leads to bespoke systems with little reuse of existing knowledge processing modules. During the last decade there has been a move away from this implementation-centric view to a knowledge-centric view in which the knowledge model and its implementation are maintained as separate entities during the design phase. CommonKADS is such a methodology, which is currently finding favour in European KE applications. It is a results-oriented methodology for developing a Knowledge-Based System (KBS) from application selection to design and testing. It is derived from KADS [7,8] that was developed during European Union funded ESPRIT projects (Projects 1098 and 5248) that ran between 1983 and 1994. The work was extended to develop KADS to become a European standard in the form of CommonKADS [9]. KADS is now widely used within European Union countries as a practical KBS development methodology.

The use of CommonKADS to develop a KBS is fundamentally a process of multi-perspective modelling. To this end CommonKADS provides a framework of representations and process suggestions for producing system descriptions at different levels of abstraction through the use of diagrams, text and / or graphical notations. These diagrammatic representations are considered to be the most useful parts of the approach.

The methodology can be split into three main components as shown in Figure 1 - the feasibility study, knowledge modelling and design and implementation.

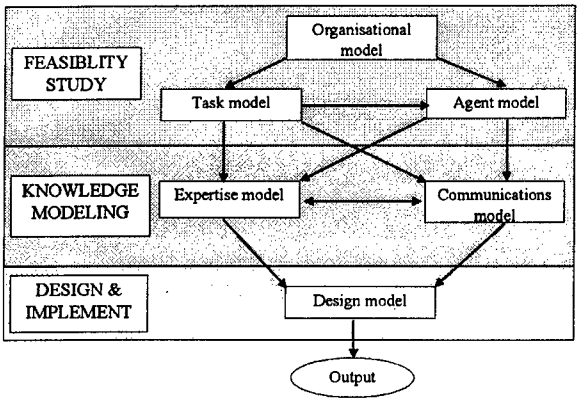


Figure 1.The CommonKADS Template

The feasibility study comprises the production of:

- an organisational model which models the organisational environment in which the system will operate;
- a task model which describes, at an abstract level, the tasks which are necessary to realise some function within the organisation;
- an agent model which models the capabilities of the people and / or the computer systems that perform the tasks identified above.

The knowledge modelling comprises the production of:

- a communications model that models the communications among the agents involved in a task. The purpose of this model is to identify some of the risks associated with the user interface;
- an expertise model which models the problem solving capability of the agents involved in the task. The knowledge required in this model can be separated into three types:
  - *domain knowledge* which is knowledge about the physical and conceptual systems being tackled;
  - *inference knowledge* which describes the inferences that can be made using the domain knowledge;
  - *task knowledge* which specifies the goals and activities making up the task and the order in which the inferences will be used.

The design and implementation phase includes:

- the design model which describes the structures and mechanisms of the systems which are involved in the task.

The CommonKADS methodology facilitates a library of re-usable models or part-models for frequently used types of task.

4. Developing the Anti-Submarine Warfare Model

A multidisciplinary team was available to work on the tool. This included mathematicians, staff familiar with different aspects of the ASW application and an experienced TacNav.

Despite its widespread use in the KBS community, there was no evidence that CommonKADS had been applied to a data fusion system. It was decided that CommonKADS would provide a valuable development tool in many data fusion applications and that the entire ASW TacNav aid development could be addressed using the CommonKADS methodology. The results are shown in Figures 2-7, although it should be noted that for classification reasons, the models shown might not always be complete.

4.1 The ASW Feasibility Study

The Organisational Model established a basic organisational context within which the TacNav aid would operate. This is shown in Figure 2.

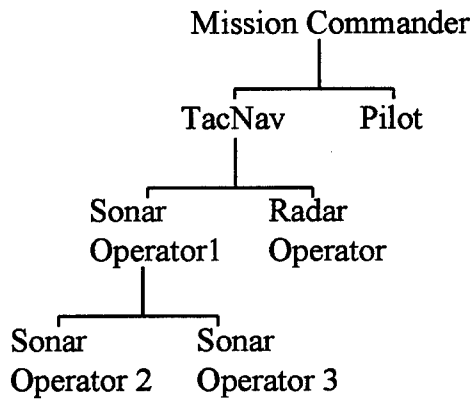


Figure 2. The Organisational Model for ASW

The Agent Model identified the main human and computer elements present in the organisation and detailed their capabilities as shown in Figure 3. From this, it was decided to ignore the existence of the radar and radar operator in the initial computer model.

The Task Model identified and related the different tasks (excluding radar) performed during an ASW mission. This is shown in Figure 4.

AGENT	CAPABILITY
Sonobuoys	able to passively detect sub-surface targets and to provide intensity and Doppler information
Radars	able to detect surface targets (omitted from initial model)
Sonar operators	able to assess the sonobuoy information in context to call target contacts and to judge their confidence
Radar operators	able to assess radar information in context to call target contacts (omitted from initial model)
Tracker	able to maintain an estimate of target location using sonar derived bearing estimates
TacNav	able to assimilate the information from the above and other sources

Figure 3. The Agent Model for ASW

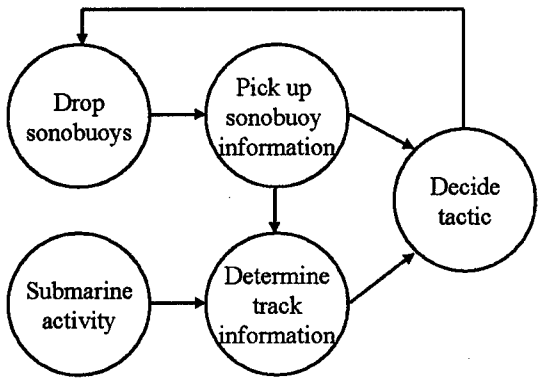


Figure 4. The Task Model for ASW

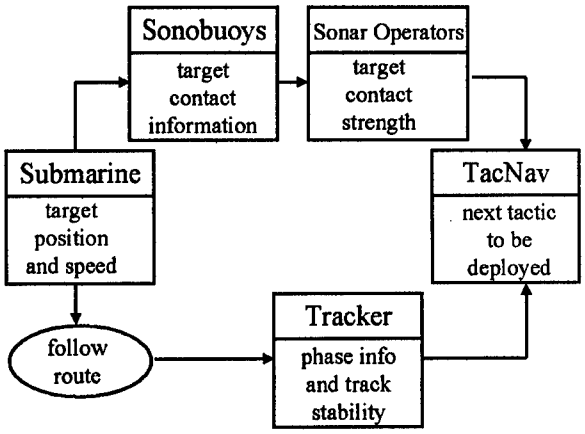


Figure 5. The Communications Model for ASW

4.2 The ASW Knowledge Modelling

The Communications Model identifies what information is passed, where it comes from and where it is passed. This is shown in Figure 5.

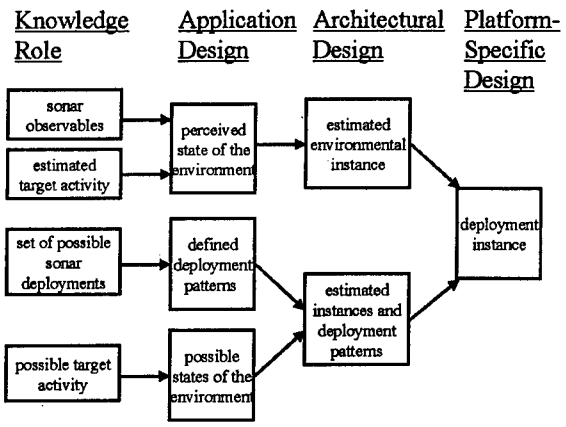


Figure 6a. The Domain Expertise Model for ASW

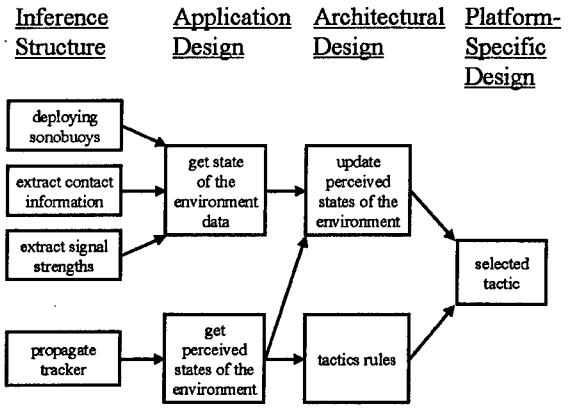


Figure 6b. The Inference Expertise Model for ASW

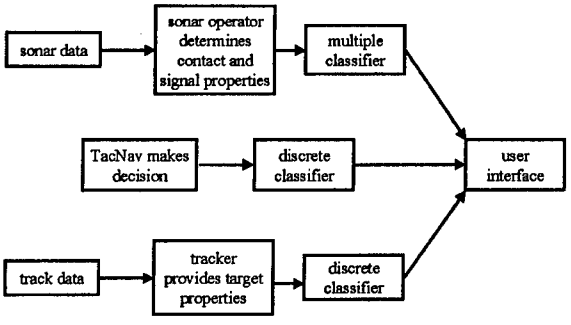


Figure 6c. The Task Expertise Model for ASW

The Expertise Model is separated into three (Figures 6a-c) and identifies the knowledge relating to the ASW problem domain, the conclusions that can be reached using the domain knowledge and the activities to be carried out and the order in which this is done.

4.3 The ASW Design and Implementation

The knowledge models were used to design a block structure for the decision aid that was connected using information derived from the communications model. The assignment of functions to particular software modules was done with reference to the organisation, agent and task models.

4.3.1. The ASW Design

A computer model of the ASW scenario was developed. This model was greatly simplified by making some assumptions including:

- there is only one target being limited in course and speed and whose action is not affected by that of the MPA;
- there is only one aircraft searching with a fixed maximum number of sonobuoys being deployed at any one time;
- the sonar operators as a group and the Tactical Navigator are of average ability;
- all crew members are aware of the area of interest and target type.

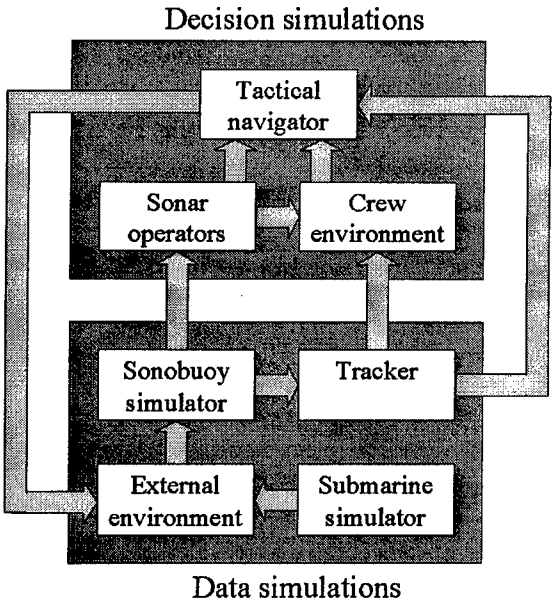


Figure 7. The Design Model for ASW

The model comprised three components illustrated in Figure 7:

- a decision simulator to simulate the TacNav's decision-making process;
- a data simulator to provide time-varying parameters, such as submarine position, of the external environment ;

a set of interfaces to provide linkages between different parts of the model.

4.3.2. The ASW Implementation

It was decided to implement the design shown in Figure 7 in two parts. The data simulator and the interfaces were written in C++ using standard software engineering practices. The decision simulator was implemented using a Bayesian Belief Network (BBN).

The purpose of the decision simulator in the ASW computer model was to predict the next action to be taken by the TacNav. The TacNav makes his decision by fusing a variety of data and information, some of which is uncertain, and then evaluating all of his options to produce a set of actions each of which is associated with a likelihood. Work at DERA has previously shown that complex military applications can be modeled using BBNs [10]. Since these allow incorporation of uncertainty into the model and produce an uncertain output, the use of a BBN for this application was considered appropriate.

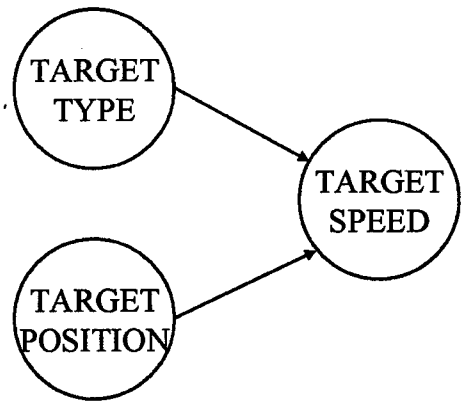


Figure 8. An Example BBN Structure

A Bayesian Belief Network comprises nodes and directional links that depict the relationship and dependencies between uncertain data. Nodes may have parent nodes and child nodes. A parent node is one whose value affects a child node. An example of a simple BBN is shown in Figure 8 where the value of *target_speed* depends on the values of *target_type* and *target_position* and hence *target_type* and *target_position* are the parent nodes of *target_speed.* Similarly, *target_speed* is a child node of both of *target_type* and *target_position.*

Each node may assume one of a number of states. For example, the *target_speed* node Figure 8 could take the values *fast*, *medium* or *slow.*

BBNs allow information about uncertainties associated with any node to be propagated through the network and the uncertainties of parent and / or child nodes to be updated based on this new information. So if *target_speed* is *slow*, but a new measurement has just been made which shows the new target position to be a long way from the previous target position (assuming periodic measurements), then *target_speed* can be updated to *medium.*

Each node has associated with it a set of conditional probabilities, known as the Conditional Probability Matrix (CPM). This indicates the probability of each state of the node given all combinations of the parent node states. The default values of the CPMs of nodes without parents are the prior probabilities of the states. When something happens to change these prior probabilities, this change is propagated through the BBN updating all subsequent CPMs using probability theory. One problem with the CPMs is size. If, in the above example, there are only two target types (*A* and *B*), three target positions (*same*, *close* and *far*) and three target speeds (*slow*, *medium* and *fast*) it can be seen that for even such a small network, the CPM for the node *target_speed* is large. In cases where a node has more than two parent nodes and / or any of the nodes have many states, the size-problem can become unmanageable.

A BBN can also work backwards. In the example above we could ask "Given that the *target_speed* is *fast*, what is the probability that the target is of type *A*?" This can be found using Bayes theorem.

The commercial package HUGIN [11] was used to implement the ASW BBN. This was chosen because previous work had indicated that it was suitable for the purpose as well as being readily available, able to run on a PC and operable from within a C++ program using an application programming interface (API).

The BBN developed for this application was only used for forward propagation, although there is no reason why it could not be used for backwards propagation as well to perhaps assess the performance of other components of the model.

The decision simulator itself was split into sub-components:

- the crew environment which models the estimated crew work-load;
- the sonar operators which are modeled as a group rather than individually. This sub-component fuses the detection and signal strength information from the sonobuoys to provide qualitative detection and strength estimates for target contact;
- the tactical navigator which fuses contact information from the sonar operators and the estimate of the crew work-load to produce a tactical decision.

The completed implementation has subsequently been tested by domain experts and is currently being considered for further development. Full details of the whole computer model can be found in [1].

5. Conclusions

The authors had previously taken an algorithmic approach to data fusion system development, and regarded the inclusion of judgmental information as outside their domain. In developing this data fusion system it became clear that the problem of including judgmental information had to be addressed. After some unstructured preliminary attempts, it became clear that a methodical approach needed to be followed. We would recommend the use of a sound knowledge engineering methodology in such cases. We found CommonKADS to be a useful, albeit somewhat unwieldy, approach.

Our difficulties in using CommonKADS included:

- the representation of the knowledge was different at the different layers;
- the diagrams could not make recursive processes explicit;
- even a small system produced a large quantity of documentation.

Advantages we have observed in using CommonKADS included:

- the solution was captured irrespective of the final implementation;
- the specification of system functionality was (properly) documented;
- the different conceptual types of knowledge were appropriately distinguished making the final model easier to understand;
- it seemed that large systems would be more easily maintained.

References

[1] Unpublished DERA / MoD data

[2] E. A. Feigenbaum. "Knowledge Engineering: The Applied Side" in "Intelligent Systems: An Unprecedented Opportunity", edited by J. E. Hayes and D. Michie, Ellis Horwood Limited, ISBN 0-85312-646-1, 1983

[3] D. M. Cleal and N. O. Heaton. "Knowledge-Based Systems: implications for human-computer interfaces", Ellis Horwood Limited, ISBN 0-7458-0152-8, 1988

[4] Patrick Henry Winston. "Artificial Intelligence", Addison-Wesley Publishing Company, ISBN 0-201-08454-6, 1997

[5] http://www.custom-built-software.co.uk/

[6] Jennifer Stapleton. "DSDM Dynamic Systems Development Method : The Method in Practice", Addison Wesley Publishing Company; ISBN: 0201178893, July 1997

[7] S. Hayward, J. A. Breuker and B. J. Wielinga, "The KADS methodology: Analysis and design for knowledge based systems", ESPRIT P1098 Deliverable Y1, STC Technology Ltd., Alborg, 1987

[8] D. S. W. Tansley and C. C. Hayball. "Knowledge Based Systems Analysis and Design: A KADS Developer's Handbook", Prentice Hall, Englewood Cliffs, NJ, 1993

[9] A. Th. Schreiber, B. J. Wielinga, J. M. Akkermans, W. Van de Velde, and R. de Hoog. "CommonKADS: A comprehensive methodology for KBS development", IEEE Expert, 9(6), December 1994

[10] S.-J. Farmer. "Belief Networks", DERA internal report no DRA/CIS(SE1)/651/HIP/GIP/RP3/0.B, February 1997

[11] http://www.hugin.dk

Assembling a Distributed Fused Information-based Human-Computer Cognitive Decision Making Tool

Erik Blasch
Air Force Research Lab
2241 Avionics Cir, WPAFB, OH 45433 USA
Email: erik.blasch@sensors.wpafb.af.mil

Abstract *– A human presented with a variety of displays is expected to fuse data to obtain information. An effective presentation of information would assist the human in fusing data. This paper describes a multisensor-multisource information decision making tool that was designed to augment human cognitive fusion.*

Keywords: Cognitive-Level Fusion, Belief Filtering, tracking and ATR.

1. Introduction

Many psychologists, engineers, and computer scientists design interfaces for man-machine systems. One of the inherent assumptions in these designs is that the human fuses information from a variety of displays. To understand human sensory processing, many theories have been proposed such as Gibson's work in *ecological* optics [1]. Gibson proposed that the environment affords the user with information and that ecological information contains structure. An affordance is information made available to the human; however, man's attention is needed to take advantage of potential information. Neisser [2] described perception in the form of *schemas*, where a schema is a mental codification of experience that includes a particular organized way of cognitively perceiving and responding to a complex situation or set of stimuli. A schema includes an anticipatory sensory signal, plan of action, and manager of information flow. Recently, researchers have adapted Neisser's schema to include situated action *plans*. A third paradigm is that of information processing [3] that seeks to map man and machines together. The *information processing* theory models man as a symbol manipulator with filtering and memory processes.

In dynamic environments, man's reliance on his sensory information fails for a couple of reasons: 1) sensory information is too rich to gather reliable data, 2) attention is focused on multiple tasks, and 3) complete information is not observable. For example, a pilot looking for ground moving targets is inundated with a vast amount of information, while flying the aircraft and looking for targets, and is only one observer of the complex battlefield, shown in Figure 1. In the first case, the human needs to augment his sensory capability by utilizing other sensory information such as radar. In the second case, the pilot's attention is divided between target identification and successful control of the aircraft. In the third case, the pilot is a member of a competitive dynamic situation. The pilot is a distributed battlefield processor; however, through communication links, the fusion of information over space can be resolved in a computer interface to afford the person with information from other aircraft or satellites. Additionally, a fusion interface design localizes his field of view, can augment his sensing capability, and provide information for flying the aircraft and identifying targets.

Focus on data and information fusion has relevance for cognitive interfaces. Data fusion integrates sensor signals, whereas information fusion processes signals for meaningful constructs. Researchers have effectively been working in data fusion (Waltz and Llinas [4], Varshney [5]), information fusion (Mahler [6]), and decision fusion (Dasarathy [7]). At the cognitive-fusion level [8], the human utilizes information to develop a parsimonious fused perception of the world. Gathering information from

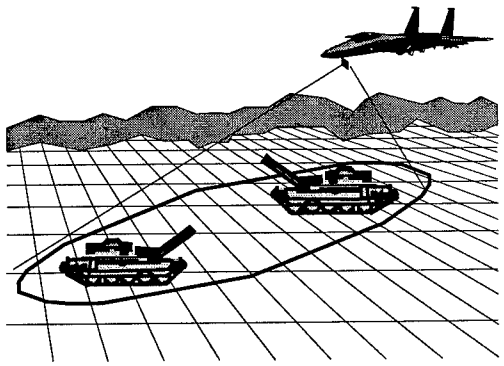


Figure 1. Target Sensor Management.

an interface, the human must make an evaluation of the information and form not only a fused perception, but also a fused action as shown in Figure 2. Cognitive fusion includes goals, decisions, and a *fused action.* Managing sensors for target identity is an example.

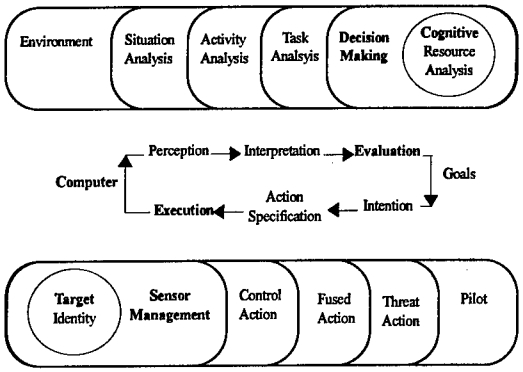


Figure 2. Fused Ecological Information perceptual evaluation and action execution.

Cognitive psychologists, such as Rasmussen [9, 10] and Flach [11], have been addressing issues for designing interfaces to augment complex decision making. Bennet and Nagy [12,13,14], have design concepts to enhance user performance and minimize human errors. Their approach is to employ ecological interfaces that afford functional abstraction. In addition, others have focused on design interfaces that effectively afford the user relevant information. Relevant information includes movement and color representations of targets. We seek to address human motion processing to augment these displays for spatial and temporal fusion [15]. Finally, the role of information accumulation is also one of uncertainty reduction. Researchers such as Bisantz and Llinas [16,17] are investigating uncertainty minimization through trust in automation.

Cognition for moving ground targets from synthetic aperture radar (SAR) and high range resolution (HRR) sensors has been a topic of recent discussion. Kuperman[18,19,20] is assessing crew aiding systems for subjective assessment of SAR imagery, which includes cognitive fusion [21]. Blasch [4] has proposed a cognitive fusion algorithm for SAR and HRR processing and an adaptive action algorithm [22]. Blasch's algorithms are based on the multiple levels of fusion including data, information, and cognitive level fusion. The integration of computer and human fusion is a new field and a topic of research interest.

Humans form hypotheses about the world and then seek information to confirm these hypotheses. One important issue is the processing of moving information. Watamaniuk [23] has shown that people process a local and global speed signal and has used to the information to guide the presentation of moving information [24]. Additionally, Wamataniuk's work in random dot displays is like the clutter in a SAR image [25]. We seek to utilize movement information for man-machine radar target identification.

For this paper, we seek to assemble an interface that fuses SAR and HRR information, integrates multisource spatial and temporal information, and affords the user with an ecological perception of the battlefield for distributed cognitive decision making of ground moving targets. Section 2 formulates the ground target identification problem and Section 3 details issues in cognitive fusion ATR. Section 4 describes the interface and Section 5 discusses issues relevant for further discussion and research.

2. Ground Target Identification

When performing target identification, a pilot focuses on salient information, such as threats to survival and control of the aircraft. Threats are difficult to measure because they are situation dependent and require reactive navigation [22]. While navigating a scenario, a pilot seeks to increase target-identity confidence by fusing and anticipating sensor measurements. Given a sensor suite, the pilot must adaptively view the correct sensor to discern the target of interest. In the multisensor/multitarget scenario, the pilot desires information that affords the best set of information to identify targets.

Recursive decision making under uncertainty is prominent in sensor fusion strategies. Sensor fusion includes automatic signal filtering, measurement association, target threat estimation, and cognitive sense prediction. Figure 3 shows a cognitive fusion model, based on the JDL levels of fusion, in which kinematic data is processed for situational and threat information. After fusion of data for information, a sensor manager, such as a human, must take a plan of action to choose the next set of sensor measurements. A target recognition and tracking plan includes a domain representation, a dynamic environment understanding with risks and uncertainties, and acknowledgement of situation complexity arising from many possible sensor actions and outcomes. Such recognition problems have been studied for

engineering and cognitive tracking research [22].

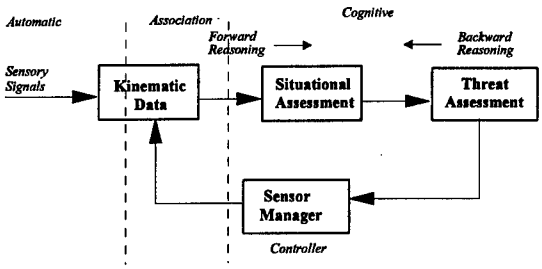


Figure 3. Sensor Fusion.

A method for automated sensor fusion and sensor action plan selection would assist pilots in time-critical target tracking, identification, and threat assessment [4]. For instance, tracking a moving target includes *searching* measurements, *predicting* target types, *extracting* information and *matching* sensed and expected information. Performing such a task requires measurement *action selection* to minimize the number of measurements and optimize target identity. Roboticists, who are researching man-machine systems, have developed algorithms for planning [26, 27], perceptions [28], and assessing goals [29].

An interface design can be an effective tool if the user trusts displayed information [17]; however, if the interaction is not mutual, either the human trusts the interface or neglects the interface completely. If the uncertainty is high and interface confidence is low, the human chooses not to use interface such as in the case where a human turns off the display and visually looks for a target on the ground. If the pilot must maintain a high altitude, visual scanning is not possible. The pilot must put full faith in the interface information. We seek to augment the human-machine fusion by operating in the domain of the human, such as presentation of sets of information with confidence values related to the uncertainty in the measurement system. An effective and efficient interface can aid target identification, but presenting fused information is not well understood.

3. Cognitive ATR Decision Making

Gibson referred to the cockpit environment as affording information to the user. While the environment is man-made, we can take advantage of the interface design so as to afford the user with fused information for decision making. Decision-making processes require the management of vast amounts of information. The human mind unfortunately is limited in its capabilities to manage, recall, and sort information. However, computers are adept in data collection, manipulation, and fusion tasks. One advantage of humans is fusing information for decision making by bounding sets of information. Computers can support the human decision making process by presenting sets of information to enhance ATR speed and quality while the human can create and manage sets of information.

The *cognitive information fusion* concept is implemented in a computer interface which utilizes target sets, confidence values, and color-coding. The interface filters radar data, presents salient information, and captures incomplete knowledge. By using a hierarchical structure for information and data fusion, the human can bound the selection of fused information. Thus, high-level information and low-level data-fusion bound the information database. Further insights can be gained from the database through "*belief filters*"[4], which represent the current situational fused belief. A unique interface feature is the ability to display any information-fusion level to allow for multiresolution decision-making.

3.1 Data Fusion

Time-critical scenarios, where multiple sensors can look at the environment, force the pilot to adaptively select sensors for target track updates as depicted in Figure 1. However, there is a tradeoff of sensing time and confidence. The difficulty is that only a few sensors can measure a target before an updated track is needed. Hence, to save time, certain sensor measurements may be ineffective for target recognition, or lack information-producing actions and track updates. The interface must provide reliable, real-time feedback to support decision-making.

3.2 Information Sets

Fitts and Posner presented a way for humans to learn new tasks [30]. They presented three stages of development as cognitive, association, and automatic. In the case in which a human is presented with a new and complex problem, they first use declarative knowledge in acquiring new facts to understand the *cognitive* problem. In the *association* stage, evidence is accumulated to prune or eliminate extraneous facts. Additionally in this stage of conflict resolution, facts are matched in order to develop relationships between the targets. Finally, in the third and final

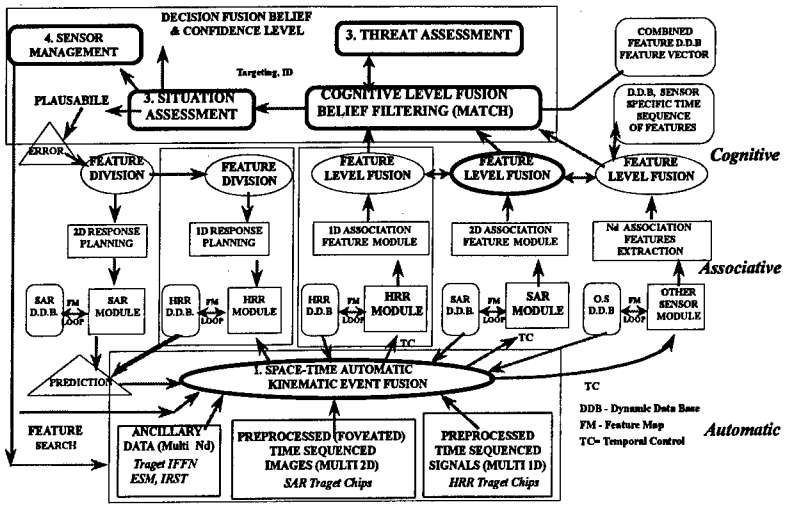


Figure 5. The Cognitive HRR/SAR Control Hierarchy.

stage, association rules are used to *automatically* perform the task. Like Fitts and Posner, we chose to employ these stages, as shown in Figure 5. We modify the initial idea for the ATR problem. We view incoming data in the automatic stage as a set of data since raw data gathered by the sensors is converted to facts, features, or information based on learned rules and phenomenology. The second difference is that the data association is resolved into information sets. Finally a cognitive stage uses fused confidences based on the information sets to identify unknown target types.

The action confidence level determines the amount of clutter measurements. The tracking system processes the clutter for target recognition and chooses to move forward, avoid threats, or seek mission targets which is displayed in the interface. The scenario is similar to one in which a pilot monitors multiple target perspectives and selects the set of sensor actions that confirms threat beliefs.

3.3 Situation and Threat Information Fusion

Situational information fusion requires a learned set of adaptive actions producing a goal-directed behavior. The problem is complicated due to target-threat importance, measurement uncertainty, and order of actions. The mission specific goal is to get to a desired target while avoiding threatening targets. Since the threatening targets are random, off-line learning will not help; however, some time is available for coordinating a set of next-state sensor measurements to discern threats, which is a human-machine cooperation task.

The adaptive action algorithm fuses sensor and dynamic information such as target maneuverability. The system reasons over possible sensing actions for threat assessment. Actions are prioritized based on target of lethality or desirability. Using the action plan, the pilot reasons over track updates to identify a target. For adaptive sensing actions, the interface presents target confidences to the user.

An action is *information producing* if it has a causal relationship. The target threat update increases confidence when a causal relationship occurs. For instance, a causal relationship exists for sequential processing of the identity and its threat, but not the reverse. Updating the threat belief with only the threat measurement results in a minimally reinforced belief. To conduct the analysis, the person must carry out sensing plans that are adaptable to the sensed information. Although the pilot does not process probability measurements, he does compare relative probabilities as confidences compared to other target identities from a set of targets. A pilot cares only about the decision, not how it was derived. To calculate *belief confidences,* association of space-time event action probabilities is fused. The *belief association probability* summation is used to develop confidences in sensed information. Once the belief is updated, a confidence level is presented based on the fusion of spatial associations and temporal target state estimates.

4. Interface Design

While the interface is only one of many possibilites, it serves as a model from which the fusion community can discuss issues in presenting fused information for decision making.

4.1 Data Fusion

From the onset, it was decided that the signal-level information would be difficult for the human to process, but the person would want access to the data. For instance, HRR information is a 1-D signal that captures the movement of the target. The human

has a 1-D sensor for audition, so audible information is available for target identity similar to radar Doppler processing. Additionally, by presenting the 1D signal, (shown in Figure 6, top right) fusion of visible information can verify if the correct signal is obatined, the relative size of the target, and whether the signal is above background noise.

For a stationary target, the radar information is displayed as a SAR image (shown in Figure 6, top left). The SAR image is cluttered, however, the user can choose a region of the image to process. Typically, a moving target indicator (MTI) provides access to all the targets in the field of view; however, the human must determine which target is of interest. In the case of multiple targets, tracking information can provide visual cues as to the position of the targets (shown in Figure 6, top middle). Thus, the human acts as a sensor manager to select targets, from a pushbutton interface, and regions of interst to focus the radar sensor data collection (shown in Figure 6, top).

4.2 Fused Information Sets

Information fusion is a result of the data and signal analysis. The SAR and HRR data types are fused by the computer or by the human. Since the human tries to compare the data with learned perceptions of targets, he is performing a search, predict, extract, and match for targets. For instance, in the battlefield, certain types of targets are assumed to be moving together like tanks. The human must parsimoniously limit the matching of targets from a set of hypothesized targets. Likewise, the interface processes sets of information and presents confidence values (shown in Figure 6, lower right). The control of target set sizes is done by choosing a minimum set of target types to analyze. Initially the belief in all targets is possible, but through accumulated sensed evidence, the correct target identity increases. This is done interactively between the human and the interface through set management. Additionally, targets that are not plausible are pruned from the plausible set. The difference between the believable targets and the plausibility of targets can be used as a confidence measure (shown in Figure 6, lower right). Thus, the human and the interface both process confidences for suspected target identity and location that can be assessed through receiver operator curves, (shown in Figure 6, lower left).

4.3 Cognitive and Decision Fusion

Since the pilot is only one of many in the battlefield, additional information is processed to determine the

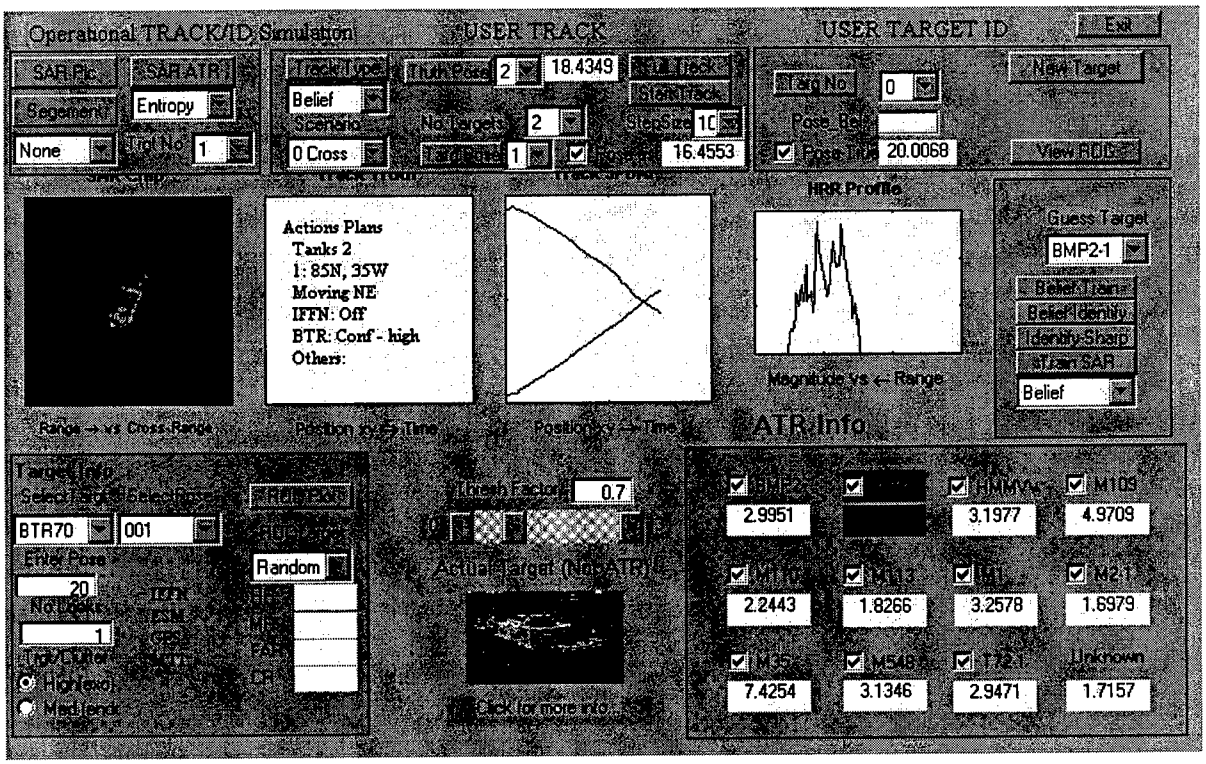


Figure 6. Initial Interface Design for Integration Fusion of Information.

targets (shown in Figure 6, lower left). The case of a multiplatform scenario affords the user information from other aircraft, with their respective sensors. This spatial information is provided to the user and included in the calculation of the confidence measures. Additionally, the temporal information fusion is available from the target tracker (shown in Figure 6, top center).

At the cognitive fusion level, additional information is needed such as Identification of Friend, Foe, or Neutral (IFFN) target affiliation (shown in Figure 6, middle center). Decision fusion is one in which the interface helps select the targets of interest. When suggested targets are assessed, the human confirms which targets to prune or add to the target set.

4.4 Fused Action

The purpose of the paper is to discuss issues in human-computer interface fusion; however, for the sensor management case, the human makes decisions serially. Likewise, the computer makes sequential decisions, albeit at a faster data rate than a human so as to appear to be processing in parallel. Cognitive fusion can be called parallel processing, however, we do not discuss the issue, since the interface is limited to sequential decisions. Since the human can only take one action, it should be a fused action based on the information and decision chosen.

4.5 Initial Human-Computer Interface Issues

The analysis of the interface is the result of one human assessing the information and is subject to the designer's preferences. Color, motion, and size are all cues that augment the perception of the targets. Tracking and motion cues help to direct attention to the targets of interest. Additionally, colors, well separated in the color space, help to clarify target confidences. Studies have shown that the human is adapted to processing 7 ± 2, pieces of information [4]. At all times, the interface seeks to take advantage of the limited numbers of information. For example, color separations was limited to 7 colors for processing.

Kuperman [17], used a SAR rating system and found that operators preferred image enhancements to the SAR imagery which consisted of reducing the image sizes by statistical means and a fuzzy set enhancement of the image. In the interface design, we use SAR image enhancement by segmenting the MTI plot with multiple targets, to that of a single target with image smoothing and size enhancements. It was found that the human was better at identifying the target when size was increased and performed slightly better with the smoothed image, rather than the raw data alone.

5. Discussion and Conclusions

The interface design is the initiation of work in augmenting image analysts and pilots for assessing ground moving targets. While many issues could serve to enhance the work, none should be ruled out. The research goal is to design effective and efficient interfaces that present a fusion of information from the computer for the human. The research goal is to integrate the two systems through the interface design.

Many issues will need to be tested to determine the validity of the design. Hence, assembling the interface, as opposed to the successful analysis of the design is the key to the work. Research in engineering data, information fusion, and decision fusion were used to develop the *signal-processing* and research in psychology and perception motivated the *display design*. Cognitively, engineering and psychology motivate assembling an interface to afford the user with effective and efficient ways for target identification for cases in which a purely visual analysis is not available, such as a high altitude aircraft with radar sensors.

The author invites any comments and suggestion from which to spawn a new field of research in human-computer evaluation and execution fusion interface designs.

References

[1] J.J. Gibson, *The Senses Considered as Perceptual Systems*, Waveland Press, Inc., Prospect Heights, IL, 1966.

[2] R. Lachman, J. L. Lachman, and E. C. Butterfield, Cognitive Psychology and Information Processing, Lawrence Erlbaum Associates, Hillsdale, NJ, 1979.

[3] U. Neisser, *Cognition and Reality*, W. H. Freeman and Co., New York, NY, 1976.

[4] E. Blasch and L. Hong. "Sensor Fusion Cognition using belief filtering for tracking and identification," *SPIE Int. Sym. On Aerospace/Defense Simulation and Control, ATR, Vol. 3719,* Orlando, FL, 13-17 April, pp. 250 –259, 1999.

[5] E. Waltz and J. Llinas, *Multisensor Data Fusion*, Artech House, Inc. 1990.

[6] P.K. Varshney, "Multisensor Data Fusion," *Electronics and Communication Engineering Journal*, vol. 9. pp. 245-253, Dec. 1997.

[7] R. Mahler, 'Random Sets in Information Fusion', in *Random Sets: Theory and Applications*, Eds. J. Goutsias, R.P.S. Mahler, H.T. Nguyen, IMA Volumes in Mathematics and its Applications, Vol. 97, Springer-Verlag Inc., New York, pp. 129-164, 1997.

[8] B. V. Dasarathy, *Decision Fusion*, IEEE Computer Society Press, Los Alamitos, CA, 1994.

[9] J. Rasmussen, A.M. Petersen, and L. P. Goodstein, Cognitive Systems Engineering, New York: John Wiley and Sons, inc., 1994.

[10] J. Rasmussen, "Skills, rules and knowledge; signals, signs and symbols; and other distinctions in human performance models," IEEE Trans. Syst., Man, Cybern., vol. SMC-13, pp. 257-266, Jan. 1983.

[11] J.M. Flach, P. Hancock, J. Caird, and K. Vicentre (Eds.) *Ecology of Human-Machine Systems*, Hillsdale, NJ, Earlbaum, 1994.

[12] K.B. Bennet, A. L. Nagy, and J. M. Flach, " Visual Displays,' in *Handbook of Human Factors and Ergonomics*, G. Salved (ed.), John Wiley and Sons, pp. 659-696, 1997.

[13] J. M. Flach, K. B. Bennett, "A theoretical framework of representational design, in R Parasuraman and M . Moula (eds.) *Automation and Human performance: theory and applications*, Mahwan, NJ: Lawerence Earlbaum Associates, pp. 65-87, 1996.

[14] K.B. Bennet and J.M. Flach, "Graphical Displays, implications for divided attention, focused attention, and problem solving, *Human Factors*, 34(5), pp. 513-533, 1992.

[15] K. B. Bennet and A. L. Nagy, "Spatial and temporal considerations in animated mimic design," *Displays*, 17(1), pp. 1-14, 1996.

[16] Llinas, J., Bisantz, A., Drury, C. G., Seong, Y., and Jian, J. Studies and analyses of aided adversarial decision-making. Phase 2: Research on Human Trust in Automation. April, 1998. Center for Multisource Information Fusion, State University of New York at Buffalo.

[17] Bisantz, A. M. Modeling Environmental Uncertainty to Understand and Support Dynamic Decision Making. Ph. D. thesis, School of Industrial and Systems Engineering, Georgia Institute of Technology, December 1997.

[18] G.G. Kuperman, "Human Systems Interface (HIS) Issues in Assisted Target Recognition (ASTR)," NAECON97, pp. 37-48, 1997.

[19] G. G. Kuperman and D. Shaya, "Subjective Assessment of SAR Imagery Enhancement Algorithms," Tech Report AL/CF-TR-1997-0133, 1997.

[20] J. E. See, I. Davis, and G.G. Kuperman, "Aided and unaided Operator performance with SAR imagery," NAECON98, pp. 420-427, 1998.

[21] G. G. Kuperman, "Cognitive Systems Engineering for Battlespace Dominance," NSSDF98, pp. 91-103, 1998.

[22] E. Blasch, "Adaptive Action Learning," *Fusion 98*, Las Vegas, NV. July 6-9, pp. 183-190, 1998.

[23] Bravo. M, and S. Watamaniuk. "Evidence for Two Speed Signals: A Coarse Local Signal for Segregation and a Precise Global Signal for Discrimination," *Vision Research*, Vol. 35, pp. 1691-1697, 1995.

[24] S.N. J. Watamaniuk, R. Sekular, and D.W. Williams, Direction perception in complex dynamic displays: the integration of direction information, Vision Research, 29, pp. 47-59, 1989.

[25] S. N. J. Watamaniuk, and S.P. McKee, Detecting a trajectory embedded in random-direction motion noise, Vision Research, 35, pp. 65-77, 1995.

[26] D. W. Payton, "Internalized Plans: A Representation for Action Resources." *Designing Autonomous Agents – Theory and Practice from Biology to Engineering and Back*, Ed. P. Maes, MIT/Elsevier Boston, MA, 1994, pg. 89-104.

[27] P.E. Agre and D. Chapman, "What are plans good for?," *Designing Autonomous Agents – Theory and Practice from Biology to Engineering and Back, 3rd.* Ed., P. Maes, MIT/Elsevier Boston, MA, 1994, pg. 17-34.

[28] R. C. Arkin, "Integrating behavioral, perceptual, and world knowledge in reactive navigation." *Designing Autonomous Agents – Theory and Practice from Biology to Engineering and Back.* Ed. P. Maes, MIT/Elsevier Boston, MA, 1994, pg. 105-122.

[29] P. Maes, "Situated agents can have goals." *Designing Autonomous Agents – Theory and Practice from Biology to Engineering and Back*, Ed. P. Maes, MIT/Elsevier Boston, MA, 1994, pg. 49-71.

[30] P.M. Fitts and M. I. Posner, *Human Performance.* Belmont, CA; Brooke's Cole. 1967.

Hybrid Approach to Multiattribute Decision Making under Uncertainty

Galina L. Rogova
Center for Multisource Information Fusion
Calspan-University at Buffalo Research Center
Buffalo, NY, U.S.A
rogova@cubrc.org

Paul Losiewicz
Analytical System Engineering Corporation
Rome NY, U.S.A
losiewiczp@rl.af.mil

Abstract - *The research described in this paper addresses issues of designing a computationally effective decision support system which can assist a decision maker in making an optimal choice between several discrete alternatives. A new hybrid approach to multi-attribute decision making under uncertainty incorporating Neural Networks and the Dempster-Shafer theory of evidence is introduced. A neural network is employed for representation and quantification of a decision maker's pairwise preferences of one alternative over the other. The Dempster-Shafer theory of evidence is used for combining the evidences representing these preferences for modeling the choice of the most preferable alternative. The designed method also includes consideration of subjective judgments about attributes representing aggregated concepts along with quantitative attributes. The case study considered in the research has demonstrated the feasibility of the application of this approach to Fusion 2/3 Level problems, namely to threat assessment.*

Key words: decision support, subjective judgment, multiattribute decision making, neural networks, the Dempster-Shafer theory of evidence

1. Introduction

In today's world, decision makers face continually increased amounts of data coming from multiple sensors, communication systems, and large databases. They also have to respond more quickly. At the same time, human decision making capabilities remain limited: short-term memory, the base for perception and processing, is limited to four chunks of information [1]. These factors require the development of computerized decision aids that model the decision making process and help to overcome human limitations.

The research described in this paper addresses issues of designing a computationally effective decision support system based on the multi-attribute decision theory. The multi-attribute decision theory is used to model subjective judgment of an expert who has to make optimal choices between several discrete alternatives. The judgment modeling is based on the notion of the underlying multi-attribute expected utility (cost) or value of the expected future outcome associated with each alternative that reflects how well the alternative is rated against a chosen goal. The optimal choice corresponds to the maximal expected utility or to the minimal cost associated with the alternatives. In many cases the decision situation is very complex, making it almost impossible to evaluate existing alternatives. However, it is often possible to represent each alternative with a set of features (attributes) and evaluate each alternative based on the value of the attributes associated with it.

Generally, the multi-attribute decision making process comprises two phases: the interpretation phase and the reasoning phase.

The interpretation phase includes:

- construction of decision alternatives
- choice of attributes (qualitative and quantitative)
- prediction of expected values of each attribute for each alternative

The reasoning phase includes preference based

evaluation of alternatives and selection of the alternative corresponding to the optimal choice.

There have been several methods developed for modeling the alternative selection process of decision makers in many applications such as manufacturing, market research, transportation, etc. (see, e.g. [2,3]). Most of these methods are based on explicit modeling of the underlying utility (cost) as a function of the attributes characterizing the alternatives. However, the methods that use explicit utility functions have to make an assumption about the form of this function [4-6]. These assumptions constitute constraints that may lead to decreased adequacy of the model. Other methods used in modeling multi-attribute decision making do not require explicit construction of a utility function [7,8], and use heuristic search to find the most attractive alternative. However, this type of method demands considerable input from decision makers during the knowledge acquisition stage of the development of these methods, and in most cases, the burden put on experts is substantial [7]. Another drawback of these methods is that they may not work efficiently in the case when the number of alternatives changes.

This paper presents a computationally efficient, connectionist decision-support system which simplifies the knowledge acquisition process without putting any constraints on the form of the utility function. This method also incorporates uncertain and incomplete quantitative as well as qualitative representations of attributes. The system is also capable of adapting to any potential change of decision makers' preferences and/or changes in the decision situation.

The introduced hybrid method utilizes a connectionist approach in order to represent qualitative expert preference of one alternative over the other in numeric form. Then, the Dempster-Shafer Theory of Evidence [9] is used to combine these preferences and make a decision about the most preferable alternative.

The Dempster-Shafer Theory of Evidence is a tool for representing and combining measures of evidence. This theory is a generalization of Bayesian reasoning and it is more flexible than the Bayesian one when our knowledge is incomplete, and we have to deal with uncertainty, ignorance, and conflicting information.

The Neural Networks possess many computational and representational capabilities which make them especially suitable for representing qualitative expert preferences [10-12]:

- ability to learn from available data and to construct, verify, and validate themselves
- ability to cope with the brittleness problem;
- ability to easily adapt themselves to changes in decision environment and decision makers preferences

The detail description of the introduced hybrid approach is presented in the next sections. In Section 2, we give detailed description of our multi-attribute decision making system, Section 3 describes the process of quantification of the qualitative attributes, Section 4 describes the NN architecture for expert knowledge representation, Section 5 presents the evidential decision making process, Section 6 shows the applicability of the designed method to threat prediction and describes experiments and results.

2. Hybrid system for multi-attribute decision making

We consider here the problem of modeling subjective judgment of a single decision maker who may have imperfect knowledge about the decision situation. As it was mentioned above, the multi-attribute decision making process consists of interpretation and reasoning steps. We assume here that the interpretation step is completed and we have already chosen a set of decision alternatives and a set of attributes and defined the expected values of the quantitative attributes and evaluation grades reflecting subjective judgment about the qualitative attributes. Our effort will be concentrated on developing an approach to a computationally effective reasoning process of automated selection of the most attractive alternative. In our approach, we neither make any assumption about the form of the utility function or explicitly model it. Instead, similar to [10], at

the knowledge acquisition stage we model pairwise preferences of the expert with the neural networks. The expert is asked to compare pairs of alternatives and to order them according to his preference. Attributes of these alternatives along with the expert preferences are used to train the neural network. Utilization of pairwise comparisons only instead of comparison of all the alternatives simultaneously reduces the burden put on the human expert during the knowledge acquisition stage since it is easier for him to chose between only two alternatives. It also reduces the number of input nodes for the Neural Network and, therefore, the number of patterns and amount of time required for training of the system. This is especially important in real-life applications when construction of a large training set may be very expensive or impractical. The results of pairwise comparisons are used to compute a belief in the level of preference for each alternative considered for decision making. Utilization of the Dempster-Shafer theory of evidence instead of the heuristic search usually following the result of pairwise comparison of alternatives, allows us to deal with conflicting information that inevitably follows the results of pairwise comparison of alternatives with the NN. The conflict appears due to uncertainty related to imperfect expert knowledge about the value of numeric attributes, subjective judgment characterizing non-numeric attributes, occasional inconsistent judgments of the human expert, and the neural network errors. Instead of choosing the alternative corresponding to the maximum utility, we make our decision based on maximum belief in the level of preference computed with the Dempster combination rule as a function of quantified pairwise preferences.

Since we consider both numeric and non-numeric attributes represented by the expert subjective judgments often evaluated through a number of related factors, a quantificational preprocessing for these qualitative attributes may be required.

The designed decision support system consists of the following components:

- a process for quantification of the qualitative attributes
- an NN-based pairwise comparison model
- an evidential decision making process

All the components of the system will be described in detail in next sections. The information flow of the system is presented in Figure 1.

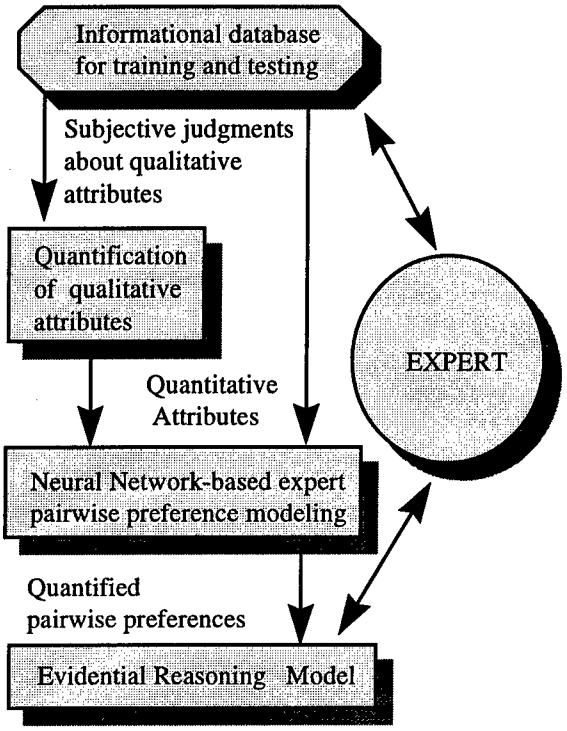


Figure 1. Hybrid Decision Support System

3. Quantification of the qualitative attributes.

As was mentioned in the previous sections, in our system we consider a situation where a set of attributes Y defining the alternatives comprises two subsets, $Y = Y_1 \cup Y_2$, where the attributes $y_k \in Y_1$, $k = 1,\ldots, K_1$ are numeric and attributes $y_n \in Y_2$, $n = K_1 + 1,\ldots, K$ are non-numeric. One way of incorporating numeric and non-numeric attributes is to quantify the non-numeric attributes. In the simplest situation, the states of the attributes for each particular alternative often can be represented by evaluation grades assigned by the decision maker. These evaluation grades reflect the decision maker's subjective judgment about the

quality of the state of the attributes, and define the preference degree. It is possible to employ these evaluation grades for numerical representation of the qualitative attributes. In decision making under uncertainty, the expert can assign more then one evaluation grade to the attribute. For example, the attribute can be good with a certain degree of confidence and, at the same time, excellent with a certain degree of confidence. In more realistic cases, the qualitative attributes present aggregated concepts and can be only evaluated through a number of related factors. The expert can assign single or multiple evaluation grades with some level of confidence only to the factors and it may be necessary to combine these levels of confidence in order to numerically assess the qualitative attributes.

The quantification process adopted here utilizes the Dempster Shafer theory of evidence for combination of multiple evaluation grades for multiple factors characterizing qualitative attributes and is similar to the method presented in [13].

Let a set of possible evaluation grades $G = \{g_1,\ldots, g_m\}$ define a frame of discernment $\Theta = \{\theta_1,\ldots\theta_m\}$ with hypothesis θ_i: the quality of attribute n corresponds to evaluation grade g_i. Let $\varphi \in 2^{\Theta}$, where 2^{Θ} is a set of all subsets of Θ. Evaluation grades G are sorted in non-decreasing order and g_1 and g_m are the worst and the best grades, respectively. We assume also that the number of evaluation grades is the same for all qualitative attributes and that, due to uncertainty, the quality of attributes can correspond to more than one evaluation grade. The objective is to define confidence levels for evaluation grades for the qualitative attribute y_n through subjective judgment of the decision maker about factors $F_n = \{f_n^j\}, j = 1,\ldots, J$ influencing the evaluation of y_n in each alternative. We consider levels of confidence assigned to evaluation grades as weights of evidence in support of hypotheses $\varphi_i \subset \Theta$. For alternative A_l, let $m_{\theta_i}(f_n^j(A_l))$ be a basic probability assignment in support of hypothesis θ_i based on the quality of f_n^j and $c_{\theta_i}(f_n^j(A_l))$, a confidence level that the decision maker assigns to hypothesis θ_i. Then we can write:

$m_{\theta_i}(f_n^j(A_l)) = \alpha_j\, c_{\theta_i}(f_n^j(A_l))$,

where α_j is a coefficient defined by the relative importance of the factor f_n^j in evaluation of attribute y_n of alternative A_l. Confidence level is defined during the expert knowledge acquisition phase such that $\sum_i c_{\theta}(A_l) \le 1$.

Combining basic probability assignment defined for all the factors with the Dempster rule of combination [9], we obtain the evaluation grades for qualitative attribute A_l:

$m_{\varphi}(y_n(A_l)) = \underset{j,\theta_m}{\oplus} m_{\theta_m}(f_n^j(A_l))$ for any $\varphi \in \Theta$.

For simplification of calculations needed for implementation of the Dempster Rule of combination we adopt the "rationality assumptions" [13] that assume that a decision maker supports no more than two consequent grades, for example, bad and very bad or fair and good. Since the basic probability assignments participating in the combinations are not zero only on singletons, the result of the combinations $m_{\varphi}(y_n(A_l)) \neq 0$ only if $\varphi = \theta_i$ or $\varphi = \Theta$.

The existing solution methods for the multi-attribute decision making require a single value assigned for each attribute. In [13], for example, the confidence levels supporting the evaluation grades are converted into a single preference degree as follows:

$p(y_n(A_l)) = \sum_i (m_{\theta_i}(y_n(A_l))p(\theta_i)$

$+ m_{\Theta}(y_n(A_l))p(\Theta))$,

where $p(\theta_i)$ is the scale of θ_i and assumed to be an increasing function defined on [-1,1] with $p(\theta_1) = -1$ and $p(\theta_m) = 1$. However, the form of function $p(\theta_i)$ is arbitrary which may contribute to the overall uncertainty existing in the problem. Utilization of the NN for modeling pairwise preference of an expert introduced in

our system does not require a single preference degree assigned to a qualitative attribute and allows us to avoid additional uncertainty related to the unknown function $p(\theta_i)$.

4. Neural network for representation of pairwise decision maker preferences

In our decision support system, a NN serves as a tool for transforming the qualitative preferences of a decision maker for a pair of alternatives into numerical values. At the training stage, we present examples of alternative pairs from the training set as inputs along with corresponding expert preferences as outputs. During the training process the system adjusts itself to respond correctly to this training set. After completion of the self-adjusting process, the system will respond correctly to an unknown input. A supervised fully connected NN is used for our system.

Each pair of alternatives used for the NN training is represented by a $2N$-tuple $T(A_i, A_j) = (T_1^i, ..., T_N^i, T_1^j, ..., T_N^j)$, where T_l^i, T_l^j, $l = 1, ..., N$ are attribute vectors of alternatives A_i and A_j, respectively. N is the number of elements in each tuple:

$N = K_1 + m\sum_{i=1}^{K_2} J_i$, where K_1 is the number of quantitative attributes, K_2 is the number of qualitative attributes, J_i is the number of factors characterizing quantitative attribute i, and m is the number of evaluation grades.

A set of target outputs for the NN comprises 2 2-dimensional binary vectors: (1,0) if alternative A_i is more preferable for the expert then alternative A_j and (0,1) otherwise. As the result of training, the NN weights will adjust themselves in such way that the NN outputs will be as close as possible to the respective targets. Therefore in our decision support system, the NN represents a transformation function $R(A_i, A_j)$ of qualitative expert preferences on a pair of alternatives into a two-dimensional vector (O_1, O_2) with the following decision rule:

$A_i \succ A_j$ if $O_i > O_j$ and $A_i \succ\prec A_j$, if $O_i = O_j$.

Here $A_i \succ A_j$ denotes that alternative A_i is more preferable than alternative A_j and $A_i \succ\prec A_j$ denotes no preference.

During the decision making phase, attributes of each pair of alternatives from a set under investigation will be presented to the trained NN. The output vectors $\{(O_1, O_2)_i\}$ will be employed to represent a measure of confidence in the choice of more preferable alternatives. These measures of confidence will be then combined in the framework of the Dempster-Shafer theory of evidence.

5. Evidential decision making process

This section describes a decision making process that evaluates preference relationships within pairs of alternatives represented by the NN outputs and chooses the most preferable alternative. Generally, if we have the preference relationships for each pair of alternatives and these relationships are non-conflicting, we are able to find the most preferable alternative using one of the existing methods, such as mathematical programming or heuristic search. In our case, when the information presented to the expert is noisy, incomplete, and contains qualitative attributes, the set of obtained preference relationships might be conflicting. In practice, we also should not expect the expert to supply consistent preference relationships on pairs of alternatives during the knowledge acquisition stage, especially when the number of attributes and/or the number of required evaluation grades is high. The NN that produces these outputs is only a model of these preferences and it can also incur this conflict. In order to combine this conflicting information and be able to make the decision about the best alternative, we employ the Dempster-Shafer

theory of evidence which is very efficient when we need to combine conflicting information coming from several sources.

Let us consider a set of M alternatives $A = \{A_i\}$ and an NN trained to model pairwise expert preferences. At the decision making stage, for all different pairs of alternatives $\{(A_i, A_j)\}$ we obtain a set of NN output vectors $O = \{(O_i, O_j)\}$ with $|O| = M(M-1)/2$. Let a set $O^k \subset O$ be a set of outputs containing O_k, where $|O^k| = M$. Let $\Xi = \{\xi_1, \ldots, \xi_M\}$ be the frame of discernment, where ξ_i represents a hypothesis that the most preferable alternative is A_i. We can also consider an NN output (O_i, O_j) for a pair of alternatives A_i and A_j as independent evidences for hypothesis ξ_i and ξ_j, where O_i and O_j are the values that reflect the measure of belief in hypothesis ξ_n. For each $(O_k, O_j) \in O^k$, $j = 1, \ldots M$, $j \neq k$, we can consider a simple support function $m^j_{A_k} = O_k$, $m^j_{A_k}(\Xi) = 1 - O_k$ with focal elements ξ_k. Then a separable support function m_{A_k} representing a combined belief in ξ_k based on all pairs from O^k is a combination of $m^j_{A_k}$:

$$m_{A_k}(\xi_i) = 1 - \prod_j (1 - m^j_{A_k}(\xi_k)), \text{ if } i = k,$$

$$m_{A_k}(\xi_i) = 0, \text{ if } i \neq k,$$

$$m_{A_k}(\Xi) = \prod_j (1 - m^j_{A_k}(\xi_k)).$$

Combining all the m_{A_k} $(k = 1, \ldots, M)$, according to the Dempster rule of combination, we can obtain now

$$m(\xi_i) = \frac{m_{A_i} \prod_{j \neq i} (1 - m_{A_j})}{\sum_i m_{A_i} \prod_{j \neq i} (1 - m_{A_j}) + \prod_j (1 - m_{A_j})}.$$

Since ξ_i is an atomic hypothesis, $Bel(\xi_i) = m(\xi_i)$ and the most preferable alternative A_m corresponds to the highest combined belief: $m(\xi_m) = \max_{1 \le i \le M} m(\xi_i)$.

6. The hybrid system for threat prediction: experiments and results

The introduced hybrid approach to multi-attribute decision making is problem independent and can be used for designing a decision-aid tool in various applications such as study of consumers' attitudes and preferences, analysis of investment alternatives, situation assessment and prediction, etc. In order to demonstrate viability of the introduced method and its applicability to the Level 2/3 Fusion problems [14], we conducted a case study where we applied this approach to modeling threat prediction. Specifically, we model a procedure of selection of the most likely threat direction (decision alternative) by considering the relative level of danger from force aggregates in different sectors of an unclassified North Korean Tactical Scenario developed for the study. For designing a case study, North Korea was divided into three zones with each zone divided into six equal geographical sectors representing a possible threat direction. For training and evaluation of the designed hybrid system we built 17 scenarios for each zone.

The level of danger was based on the analyst's implicit awareness of Combat Compound Value (CCV) represented an "underlying value of threat" for any of the defined sectors The CCV for each sector included subjective judgments about terrain and quantitative information regarding each type of intelligence data used in the assessment. Terrain within the CCV represented an aggregated concept and was evaluated through relevant factors (mobility and detectability) that were qualified as POOR, AVERAGE or GOOD with some confidence levels. Qualitative judgments of mobility was based on difficulty in Cross Country Movement (CCM) over the terrain while qualitative judgments of detectability were made based upon the concealment potential of terrain. Quantitative information

was represented by lethality that measured force projection capability based on a total ordering of the in-garrison power projection capability of each type of unit (Armor, Infantry, Artillery, Anti-Air), and the number of units found in a sector.

The informational database of 51 scenarios was used for evaluation of the designed hybrid system. The objective of this evaluation was twofold. First, we intended to demonstrate the ability of the NN to model expert pairwise preferences with both quantitative and qualitative attributes. Second, we intended to show that the combination of the NN outputs with the Dempster rule of combination allows us to model the decision maker choice of the most preferable alternative (the most likely direction of treat in the threat prediction problem).

A fully connected three layer NN with 27 hidden nodes trained with the back propagation algorithm [15] was employed for modeling an expert's pairwise preferences. The training and testing were performed with the "one-scenario-taken-out" method. For training we used 750 pairs of directions as input patterns along with corresponding expert preferences as outputs, while a test set at each cycle contained 15 directions. Each direction is represented by the number of units to reflect lethality and the subjective judgment of the analyst about terrain characteristics. The training and testing results for modeling the decision maker preferences between two alternatives are shown in Table 1.

Table 1. Prediction accuracy in modeling pairwise preferences with the NN

	Training result	Test result
Accuracy of pairwise preferences prediction	99.59%	95.82%

The NN output vectors corresponding to the results of pairwise comparison of directions from one scenario were combined following the evidential routine introduced in Section 5. The result of combination was tested against the analyst choice of the most likely direction of threat. The accuracy of prediction of the most likely direction of threat by the decision support system is shown in Table 2.

Table 6. Prediction accuracy in modeling the choice of the most likely direction of threat by the decision support system

	First choice	First & second choices
Prediction accuracy in modeling the most likely direction of threat	84.1%	100%

The results of our experiments with simulated data demonstrated a high degree of agreement between the system and the decision maker. Consideration of two choices (the best and the second best) allows us to further improve the system accuracy while introduction of the second choice does not degrade the utility of the system since it is still easier for experts to make a choice between two alternatives. More experiments with larger databases and more realistic scenarios are required in order to make final conclusions about performance of the system.

6. Acknowledgments

This research was supported by the US Air Force Research Laboratory (AFRL) Information Directorate under contract No. US AFS, F30602-97-C-0181 whose support is gratefully acknowledged. The authors also express their gratitude to Michael Hinman for valuable discussions and Michael Mandry for creating the extensive database of scenarios used in training and testing the system.

8. References

[1] *Command Decision making Modeling: Technology assessment*, Compiled by the US Army Artificial Intelligence Center, July, 1996

[2] C. L Hung and K. Yong, *Multiple Attribute Decision methods and Applications, A State-of-the-Art Survey,* Berlin: Springer-Verlag, 1981.

[3] P.E. Green, V. Srinivasan, Conjoint Analysis in Consumer research: issues and outlook, J. Consumer Res., Vol. 5, 103-123, 1978

[4] C.C White III, A.P. Sage, S. Dozono, A Model of Multi-attribute Decisionmaking and Trade-off Weights Determination Under Uncertainty, *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. 14, 2, 223-229, March 1984.

[5] S. Zions, A Multiple Criteria Method for Choosing Among Discrete Alternatives, *Eur. J. Op. Res.,* Vol 7, 143-147, 1981.

[6] M. Paolucci, R. Pesenti, A New Cost Function to Solve Multi-Attribute Decision Making Problems with Non Separable Attributes, *IEEE Int. Conf. on Systems, Man, and Cybernetics*, 1961-1965, 1991.

[7] M. M. Koksalan, M. H. Karwan, S. Zionts, An Improved Method for Solving Multiple Criteria Problem Involving Discrete Alternatives, *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. 14, 1, 24-34, January 1984.

[8] J-B Yang , P. Sen, A General Multi-Level Evaluation Process for Hybrid MADM with Uncertainty, *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. 24, 10, 1458-1470, October 1994.

[9] G. Shafer, A *Mathematical Theory of Evidence*, Princeton, MIT Press, 1976.

[10] J. Wang, M. Bender, Connectionist Decision Support Systems for Multiple Criteria Decision Making, *IEEE Int. Conf. on Systems, man, and Cybernetics*, 1955-1960, 1991.

[11] S. J. Andriole, Advanced Information Technology for Next Generation Decision Support, *Information and decision technologies*, Vol. 15, 243-257, 1989.

[12] S. I. Gallant, Connectionist Expert Systems, *Comm. ACM*, Vol. 31, 2, 152-169, 1988

[13] J-B Yang , N.G. Singh, An Evidential Reasoning Approach for Multiple-Attribute Decision Making with Uncertainty, *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. 24, 1, 1-78, 1994.

[14] E. Waltz , J. Llinas , *Multisensor Data Fusion*, Artech House Publishers, Norwood, MA, September 1990.

[15] D.E., Rumelhart, J. L. McClelland, (Eds.) *Prallel Distributed Processing.* Cambridge,:MIT Press, 1986.

AN ASSESSMENT OF ALTERNATIVE SAR DISPLAY FORMATS: ORIENTATION AND SITUATIONAL AWARENESS

Gilbert G. Kuperman
Human Effectiveness Directorate
Air Force Research Laboratory
Wright-Patterson AFB OH

Michael S. Brickner
PAMAM Human Factors
Engineering Ltd
Ramat Hasharon, Israel

Itzhak Nadler
Israel Air Force Human Factors
Engineering Branch
Tel Aviv, Israel

Abstract - *This study explores the operational utility of fusing synthetic aperture radar (SAR) imagery and digital terrain map (DTM) data. Specifically, the two-dimensional (2D) display of SAR imagery was compared against a two and a half dimension (21/2D) display of SAR overlaid on corresponding DTM data. Eight imagery analysts (IAs), assigned to the Israeli Ground Corps Command Imagery Analysis Unit and to the Israeli Air Force, and two weapon system officers served as subject matter experts. The measures employed in this comparison included both an assessment of operator situational awareness (SA) and of performance in an information extraction task. Based on the SAR imagery which was used in the experiment, performance measures (accuracy and speed in feature location) and SA measures did not yield significant performance differences between the 2D and the 2 ½D displays. The average time required to complete each task was significantly longer for the 2 ½D displays. Based on experience, the SME's opinion was that the 2 ½D imagery display may be potentially helpful in the performance of various imagery analysis tasks and in enhancing SA.*

Key Words: Synthetic Aperture Radar (SAR), Digital Terrain Map/Elevation Data, Imagery Exploitation, Situational Awareness, Information Fusion.

1. Introduction

1.1 Background

Synthetic aperture radar (SAR) sensors offer two compelling advantages over conventional (electro-optical) sensing technologies: stand-off range and adverse weather capabilities. SAR images can be formed with effectively no loss in resolution out to the limits of the system's stabilization and motion compensation capabilities. SAR sensors can "see" through clouds and through light rain. Further, depending on their coverage mode and data processing limitations, SAR sensors can be capable of high area coverage rates. These attributes make SAR imaging a valuable resource for tactical and theater airborne reconnaissance, surveillance and target acquisition applications.

The air forces of both the United States and of the State of Israel have great interest in exploiting these capabilities. The United States Air Force has operational SAR capabilities in the B-1B, F-15E, J-STARS, and U-2 systems and plans to include SAR as a primary imaging mode in the Global Hawk uninhabited air vehicle. The Israel Air Force has operational SAR capabilities in their Phantom 2000 and F-15I multi-role aircraft and has other SAR capabilities in development. (A prior study [1] explored the benefits of SAR display enhancement algorithms in an image interpretability task.)

SAR, however, is a non-literal imaging sensor. That is, the imagery produced by a SAR does not resemble a photograph taken of the same scene. The intensity values in the SAR image are proportional to the radar cross sections of the corresponding points in the ground scene (and not to their visible wavelength reflectance). The impulse response function of the SAR (the fundamental determinant of system resolution) includes side lobes. Thus, the return from a point on the ground may include energy contributed by adjacent scatterers. The "shadows" in a SAR image are caused by the active illumination of the scene by the emitting radar (and not by the sun angle). The perspective of a SAR image is that of an observer looking down on to the scene from directly above, while it is being illuminated by the radar from one side (the location of the SAR).

Because of the non-literal nature of the SAR image, operational questions exist regarding how well an imagery analyst (IA) can orient it against a map reference. A fundamental imagery exploitation task is to confirm (or plot) the actual ground coverage of a collected image against a map reference. (A recent survey of IA tasks and workstation functional requirements is presented in [2].) Several other standard imagery exploitation tasks (e. g., landform analysis, traversability studies) require that the operator interpret the image so as to assess the basic geologic and terrain characteristics, including judgments of the heights of terrain features and the grades of slopes. Further, orientation may require the IA to locate salient terrain features and to match them

against their map references. Understanding of the terrain contributes significantly to the establishment and maintenance of situational awareness (SA), affording the context within which other imagery interpretations may be made. The human operator is unique in having the ability to apply contextual information to the interpretation of complex visual stimuli (such as reconnaissance imagery).

Endsley [3] has been a primary researcher in studying situational awareness. This study attempts to extend her model (Figure 1) to the intelligence exploitation domain. Within the definitions implicit in her model, the SA metrics employed correspond to Level 1 SA.

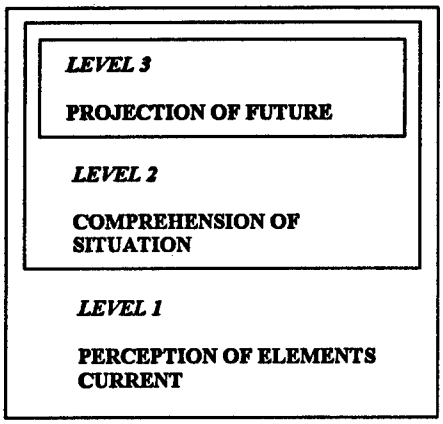


Figure 1. Model of Situational Awareness (from Endsley, 1994)

SAR is not the only technology which may support these operational requirements. Digital terrain map (DTM) data, consisting of elevation "posts," equally spaced in latitude and longitude, provide another source of information regarding the heights and slopes of the terrain. DTM data can be viewed in two dimensions (2D), as elevation contours, or as a continuous depiction in which elevation is coded by luminance values or colors. 2D image formats may be rotated so that North (or any arbitrary direction) is toward the top of the display. Alternatively, DTM data may be displayed in 2 ½D in which a 3D "model" of the terrain, with a shading scheme applied as if it were illuminated by the sun, is projected on to the 2D display surface. 2 ½D DTM displays may be rotated in both azimuth and elevation to change the effective viewpoint of the observer.

Fusion also offers potential capabilities to support enhanced orientation, situational awareness, and information extraction capabilities. Disparate data sources, such as SAR imagery and DTM elevations, may be combined (overlaid) so as to support a 2 ½D display of the SAR images. Figure 2 [4] depicts the model of the levels of fusion adopted by the US. The fusion of SAR and DTM, as in this study, correspond to Level 1 in this model.

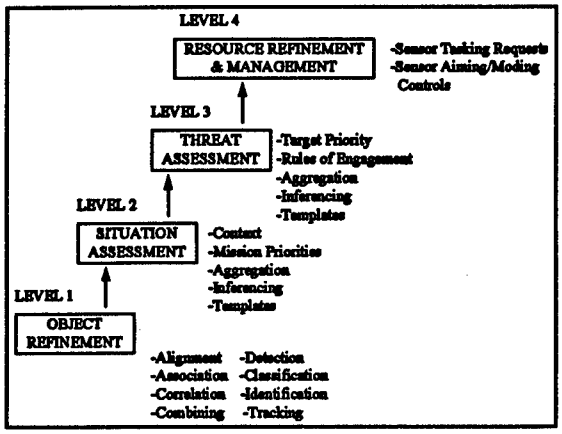


Figure 2. Model of Data Fusion

1.2 Objective and Approach

The objective of this study was to perform an operational assessment of the relative utility of 2D and 2 ½D displays of SAR imagery. In the 2D case, the SAR images were viewed conventionally. In the 2 ½D case, the SAR image was overlaid on the corresponding DTM model. Subject matter experts (SMEs), primarily military IAs assigned to the Israel Air Force (IAF), the Intelligence Command, or the Ground Corps Command, performed orientation and information extraction tasks using both display formats. The study was conducted at the facilities of Synergy Integration Ltd., Tel Aviv, with the support of PAMAM Human Factors Engineering Ltd., during the period 19 August through 17 September 1998.

2. Method

2.1 Imagery

The SAR imagery used in this experiment was acquired by a developmental sensor flown on the Israel Aircraft Industry's Boeing 737 multi-mode radar testbed aircraft. The imagery had a nominal resolution of 1.2 m. The imagery, in detected form, had a nominal dynamic range of 8 bits (or 256 gray levels). All imagery was acquired at high grazing angles (approximately 45 degrees).

Three swaths were provided by the Israeli Ministry of Defense. The first included coverage of the Armored

Command Museum at Latrun. The second included the area of Rosh Ha'ayin and the third included coverage of Ben Gurion International Airport. Thirty-eight stimulus images were extracted from the Latrun and Rosh Ha'ayin swaths. Six images, used only for familiarization with the task and practice with the apparatus, were extracted from the Ben Gurion coverage.

The Rosh Ha'ayin and the Latrun swaths differed in scale. In the Rosh Ha'ayin swath each centimeter of the image represented approximately 60 meters on the ground. In the Latrun swath each centimeter of the image represented approximately 92 meters on the ground. As a result the width of Rosh Ha'ayin swath was approximately 1 km by 1 km and of the Latrun swath approximately 1.5 km by 1.5 km. The resolution of each of the images was 700 by 700 pixels.

2.2 Selection of "Targets"

The experimental design was constrained, to some extent, by the coverage of the available imagery. Since the objective of the experiment was to investigate the effect of SAR imagery overlaid on a three-dimensional terrain elevation database and viewed in a 2 ½D display on both orientation and situational awareness, no buildings were included. A senior and highly experienced IAF IA first determined the coverage of the SAR imagery against a 1:50,000 scale survey map. Features (such as river bends, confluences/divergences of streams, the intersections of dirt roads, etc.) were selected from the map information for use as designation "targets" and their Universal Transverse Mercator (UTM) coordinates were read and recorded. These same features were then located within the SAR images and the corresponding pixel location was read and recorded. This process was repeated until all 38 stimulus targets and the six practice targets had been selected. The target location coordinates were maintained as the "school solution" for scoring the accuracy of the designation portion of the task. The imagery was then divided into 22 matched pairs (one half of each pair to be presented in 2 ½D and the other half in 2D.) The pairings were made on the basis of containing similar targets within similar backgrounds.

2.3 Overlay of SAR Imagery onto DTM Data

Commercial, off-the-shelf software (MultiGen II Pro, from MultiGen Inc., San Jose, California) was used to convert the SAR pixel coordinates into UTM coordinates, the reference system used for the DTM data. Multiple control points were selected from each of the SAR images and their geographic reference locations were carefully determined from the map. A transformation program, using these control points, was used to convert every pixel location into its corresponding UTM coordinates. One SAR image from each matched pairing was then overlaid onto the corresponding DTM elevation data (using the same software package). The product of this procedure was a 2 ½D representation of the area (as compared to the 2D representation of the original SAR imagery).

No additional exaggeration to the elevation data was introduced. Thus, the displayed image of the overlaid SAR and DTM depicted ground distances (x and y) and heights (z) in the ratios of 1:1:1.

2.4 Apparatus

The images were displayed and designation coordinates and response times were recorded using a Silicon Graphics Incorporated (SGI) ONYX graphics workstation equipped with an Infinite Reality Engine multi-processor. The workstation was also equipped with a SGI model CM2187ME 533 mm (21 inch) diagonal color monitor. The display resolution (full screen) was 1280 by 1024 pixels. The brightness and contrast controls of the display were set by the Experimenter. The apparatus was located in a laboratory setting and was used to support both stimulus preparation and data collection. All stimulus imagery was displayed using commercial, off-the-shelf software (the VEGA general visualization environment from Paradigm Simulations Inc., Dallas, Texas). The displayed image (700 by 700 pixels) was approximately 200 by 200 mm (8 by 8 inches) on the monitor.

2.5 Subject Matter Experts

Five enlisted IAs from the Israel Defense Force Ground Corps Command's Imagery Analysis Unit, three IAs from the IAF, and two Weapon System Officers (WSOs) of the IAF, served as subject matter experts (SMEs). All were male. They ranged in age from 19 to 25 years. Their experience in tactical imagery exploitation ranged between six months and 6 ½ years. Four of the IAs and both WSOs had some SAR imagery experience; all of them had experience in the exploitation of electro-optical (photography and television) sensor collections and all had previous experience in performing softcopy imagery exploitation. None of the SMEs had had previous

experience in exploiting high resolution SAR imagery (as was used in the present study). All SMEs had 6/6 (20/20) vision, uncorrected or corrected, and all had received formal military training in imagery analysis during a three month duration Service school.

2.5 The SME Task

Figure 3 depicts the sequence of events which composed the experimental task. Upon arrival at the laboratory facility, the SMEs were informed as to the purpose of the study and instructed regarding the conduct of the experiment. The instructions to the SMEs explicitly placed primary emphasis on the accurate performance of the designation component of the task but also emphasized the requirement to complete the task as rapidly as possible. The instructions also included the caution that the imagery was more recent than the map and might contain (extensive) differences with respect to the addition of man-made structures such as buildings and roads.

Information regarding the SME's background, training and imagery exploitation experience was elicited through a brief questionnaire which included questions regarding their training and experience in the exploitation of SAR imagery and their experience in interpreting softcopy imagery. The SME was then seated at the graphics workstation.

At the beginning of the task, each SME was shown an extract from a 1:50,000 scale, color, topographic Survey Map of Israel. The map, oriented north-up and covered approximately 2 km by 2 km in area, had been annotated to depict the coverage of a SAR image at a different orientation and included a red dot marking the location of a target. This map allowed the SME to understand the relative differences in coverage between succeeding map extracts and their corresponding SAR images. They were then instructed in the use of the apparatus for the imagery orientation and target designation portions of the task. The practice images were used to allow the SMEs to gain proficiency in the use of the equipment, the orientation and target designation components of the experimental task, and the nature of the SA questions. Any remaining questions that the SMEs might have regarding the task were answered by the Experimenter at this time. When the SMEs reported that they were confident in the execution of the task, the data collection trials were initiated.

At the beginning of each of the 38 data collection trials, a 1:1 scale extract from a 1:50,000 scale, color topographic Survey Map of Israel was provided to the SME. The map extract was always oriented North-up and covered approximately 2 km by 2 km in area. The header on the map copy described the type of target to be located (e. g., dome, intersection of a dirt road and a stream, etc.) while the exact location of the specific target of interest was depicted on the map itself by a small red dot. The map extracts were mounted as successive pages in a flip chart-type booklet. The Experimenter initiated each trial (by depressing a specific function key on the keyboard). The SME was permitted 15 seconds for map study. During this interval, the image display region was blank (showing a solid, medium luminance, light blue field). The Experimenter informed the SME whether the current trial was a 2D or 2 ½D display format. The SAR image, containing the target, then appeared on the workstation display. The images were always presented so that the radar shadows pointed toward the bottom of the display (i. e., as if the radar were illuminating the ground from along the top edge of the display). No restriction was placed on the viewing distance between the SME and the workstation monitor.

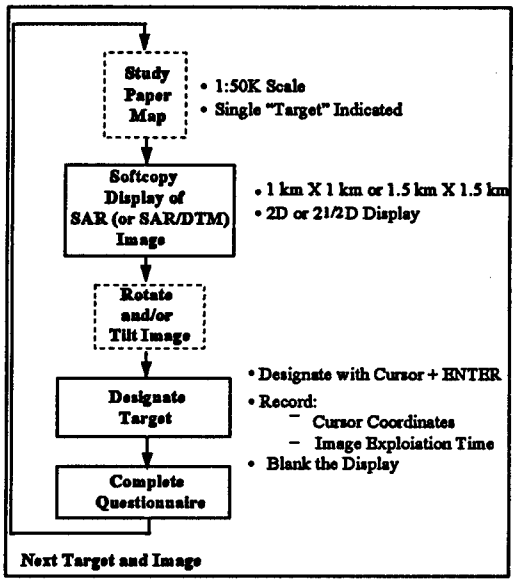


Figure 3. Flow Diagram of the SME's Task

The SMEs were permitted up to three minutes (180 seconds) during which they were required to orient themselves to the SAR image in the context provided by the map information (which was available throughout the trial), to locate the pre-briefed target, and to designate the target. At the completion of the tasks the display automatically went blank and performance time was recorded. If the SME did not respond within 180 seconds the display went blank and the trial was recorded as having "timed out".

During this three minute period, the SMEs could use the left and right arrow keys on the workstation keyboard to rotate the image through a full 360 degrees of azimuth. The up and down arrow keys "tipped" the image through 90 degrees of "elevation." Rotation in both azimuth and elevation were continuous and could be applied in any combination.

For each SME, half the stimulus images were presented in overlay on the DTM elevation data. In these cases, rotation of the displayed image produced a 2 ½D view. In the other half of the trials, a 2D view was presented, the arrow keys could still be used for tip and rotation but no elevation data were overlaid on the SAR images. The mouse was used to drive an "arrow" cursor on the display to point on the image. When the SME had located the target, the ENTER key on the keyboard was used to record the target location into the data file for that trial. (The keyboard ENTER key was preferred to the mouse buttons in order to prevent involuntary motion of the mouse cursor during designation).

Upon designation, the display was blanked and the location of the designated point was automatically recorded, along with the time between stimulus onset and the act of target designation. The SME then flipped the page in the map booklet (thus precluding any further reference to the map) and found two questions regarding the image presented during the just-completed trial. These SA questions dealt with absolute or relative terrain height judgments or with the relative location of other objects in the SAR image. The answers to the questions were recorded manually by the Experimenter. (This allowed for immediate answers to any SME requests for clarification of the SA questions.)

2.6 SA Questions

Two SA-related questions were developed by the Experimenters for each target image. The questions dealt with absolute or relative terrain height judgments (e.g., which bank of a stream was higher?, which slope of a dome was steepest?) with the direction of objects (e.g., what was the direction of the stream?) or with the relative location of objects in the SAR image (e.g., in which direction from stream bend were two large buildings?). The SA questions were presented in multiple choice form, three possible answers to each question were presented and the SME had to select the correct one. No time limit was imposed in answering these questions.

Once the SA questions had been answered, the trial was completed. The SME then indicated readiness to proceed with the next trial. This sequence was repeated until all 38 images had been presented to the SME. The SMEs were given a short break after each group of eight to 12 trials (while the Experimenter loaded a different SAR swath).

2.7 Rating Scale Questions

After all 38 stimulus images had been presented, the SME was asked to complete a series of rating scale questions regarding overall impressions of the task and of the two different display formats. Each scale consisted of seven points with semantic anchors at each endpoint (as shown in Figure 4).

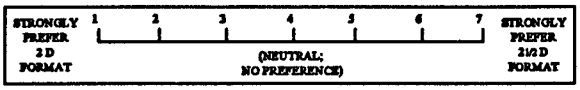


Figure 4. Rating Scale with Semantic Anchors

A rating of one always meant that the 2 ½D display greatly degraded the SME's ability to perform the referenced function while a rating of seven always meant that the 2 ½D display greatly enhanced that ability. The first group of questions dealt with comparisons between the 2D and 2 ½D display formats with respect to: performing general orientation, assessing the structure of the terrain, assessing differences in terrain heights, and assessing terrain slopes. The next scale required the SME to rate utility of the 2 ½D display format in supporting general imagery interpretation tasks. Another set of questions related to SA. The SME was asked about the differences between the 2D and 2 ½D display formats in supporting giving answers to the SA questions. The SMEs were also asked to comment on whether they relied primarily on the map extract or on the SAR imagery in answering these questions. They were also requested to comment on the relevance of the SA questions to their current military duties. Provision was also made for the SMEs to record any overall impressions or comments regarding the entire experiment

Upon completion of the rating scales, data collection was ended and the SME was thanked for participation in the experiment. Each SME participated for approximately two hours, including instruction, practice, data collection, and completion of the questionnaire.

2.8 Experimental Design

A mixed, within-subject experimental design was employed. Half of the SMEs were presented with one half of the matched SAR image pairs overlaid on to the DTM data; the other half of the SMEs were presented with the alternate half of the image pair presented in non-overlaid format. Half of the SMEs were presented with the experimental imagery in the reverse order from that presented to the other SMEs. This counterbalance was to protect against learning effects. Thus, there were four unique combinations of imagery presentation: order of presentation and DTM or non-DTM underlay (the independent variable of interest).

3. Results

3.1 Designation Accuracy

Accuracy of the terrain feature designations was measures in cm on the displayed image. The mean accuracy score for the 2D display was 1.33 and for the 2 ½D display 1.39. This difference is not significant.

3.2 Response Time

Designation time tended to be longer for overlaid SAR-DTM images. The average response time for the 2D images was 51.9 seconds and for the 2 ½D 60.6 seconds. This difference is statistically significant (p=0.001). Designation times for the Latrun swath (51.79 seconds) were significantly shorter than for the Rosh Ha'ayin swath (60.68 seconds) (p=0.001). (The shorter response times for the Latrun swath may be due to the higher availability of salient human-made features in the images of the Latrun area.)

3.3 SA

Each trial was followed by two SA questions. A score of 1 was assigned to each correct answer and 0 to wrong answers. SA scores were computed for trials with correct and partially correct target designations only. The final SA scores were computed as the sum of points for each trial. The mean SA scores for the 2D images was 1.08 and for the 2 ½D images 1.04. This difference is not significant.

3.4 Rating Scales Responses

The first four questions on the rating scale dealt with the strength of the SMEs preference for either the 2 ½D or the 2D SAR display format in the context of supporting the IA's ability to orient to the terrain scene. The first scale addressed general orientation, the second addressed the assessment of landforms/terrain structure, the third understanding of terrain height differences, while the fourth explored understanding of differences in terrain slopes. As depicted in Figure 3, the SMEs, as a group expressed a marked preference for the 2 ½D display format. (In the Figure 5, a mean rating of 4.00 reflects no preference between the two formats.)

The fifth rating scale required the SMEs to express their preference in the context of the utility of the display format to support imagery interpretation in general. A preference for the 2 ½D format was found.

The sixth rating scale explored the two display formats in the context of SA. Again, a preference for the 2 ½D was elicited.

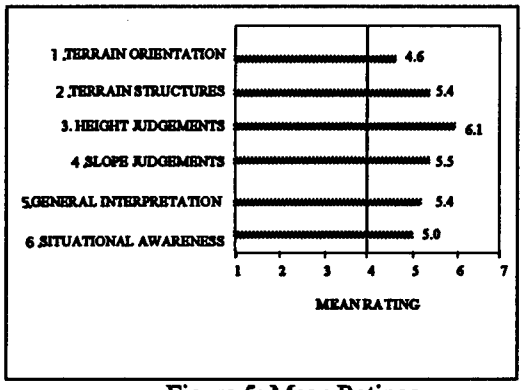


Figure 5: Mean Ratings.

All ratings were significantly higher than the neutral score (4.0). Table 1 presents the statistical summary for 10 SMEs.

Table 1: Mean ratings and T scores for the six rating scale questions:

Question	Mean Rating	T score	Probability
1	4.6	2.64	0.05
2	5.4	4.24	0.01
3	6.1	13.74	0.001
4	5.5	3.53	0.01
5	5.4	5.22	0.001
6	5.0	2.55	0.05

3.5 Observations

Before discussing the implications of the results from the formal measures used in the study, some observations on the part of the Experimenters, made during the data collection runs, may give the reader insight into the study.

None of the SMEs had any apparent difficulty in employing the display/controls mechanization (arrow keys, mouse, enter key) used in this study.

Although none of the SMEs had any experience in the exploitation of high resolution SAR, they were all able to complete the target designation task without any reported difficulty.

All IAs had received training in landform and traversability analysis as part of their IA school curriculum.

Some SMEs indicated that the effective usage of 2 ½D images may require experience and perhaps even formal training.

Wide-ranging, individual differences were observed with regard to the strategies employed by the SMEs in viewing the SAR display. Some SMEs physically rotated the paper map to match the orientation of the SAR (regardless of whether DTM data were available). This kept the radar shadows pointing toward the bottom of the display – a technique that IAs are taught to employ to avoid a "false" reversal in apparent elevation / depression of the scene. Others appeared to first rotate the SAR display (again regardless of the format) and then to quickly tilt the displayed image, apparently to gain an appreciation for terrain relief.

4. Conclusions and Recommendations

4.1 Conclusions

High resolution SAR imagery, collected at high grazing angles, does not appear to present any of the difficulties conventionally associated with low and medium resolution non-literal imagery at least in the context of the present salient landform designation and terrain-based SA tasks. This also suggests that only minimal impact to the training support system may be encountered as these systems become operational.

Designation scores with the overlaid SAR-DTM imagery (2 ½D) produced slightly higher accuracy scores than SAR alone (2D). However, these differences were small and did not reach statistical significance. The general pattern of results did not change when only selected targets, which contained mountainous areas and no salient human-made features, were analyzed. The elimination of the most difficult and the easiest trials from the statistical analysis increased the differences between the 2D and the 2 ½D scores, but this difference too failed to reach statistical significance. Several factors may have affected the potential effects of an overlaid SAR-DTM imagery on the accuracy of target recognition:

The sets of SAR swaths used in the study were rather limited in size and included only small areas which were both mountainous and free of salient human made objects. Hence, the number of sections in which the SAR-DTM overlay could provide significant advantages was rather small and the variety was very limited.

Because of the limited width of each swath and the small variety of useful terrain areas, the size of the area displayed during each trial was significantly smaller than the size of area which IA use in their regular routine. This may have made the use of terrain features more difficult than usual to exploit.

The use of the overlaid SAR-DTM seems to require some training. This was indicated by the results which show a larger improvement in SAR-DTM performance than in SAR alone, and was pointed out by some of the SMEs (in their comments) as well.

Response times were approximately 17 percent longer for the 2 ½D trials than during the 2D trials. This is not surprising given that the 2 ½D images contain more information. Additionally, during the 2 ½D trials SMEs made more extensive use of the tilt option which provided them with different views of the terrain, whereas, tilting the 2D images was possible but did not provide any additional information.

Situation awareness as measured by the questions at the end of each trial did not benefit from the overlay of SAR-DTM. Two reasons may have affected the results. First, the answers to the SA questions could be extracted from the maps as well as from the SAR images. At the end of the experiment SMEs were asked about the extent to which their SA answers were based on the SAR as compared with the map. During debriefing, most SMEs reported that the maps were an equal or a dominant source of SA information. Obviously, the use of the map obscures SAR imagery effects. Secondly, although all IAs considered the SA

questions as relevant to their jobs, they also indicated that the level of details required tended to be higher than is usually required on the real "object recognition" job, (e.g., comparing the slopes of two adjacent domes). Several SMEs indicated that this level of detail would be more relevant for determining traversability. Hence, some of the SA questions were perceived as an additional secondary task rather than as part and parcel of the main target acquisition task.

Individual performance differences were quite large and seem to be related to the level of experience. Interestingly, the more experienced SAR interpreters seemed to have benefited less from the SAR-DTM overlay then the inexperienced SMEs. However, these findings were not significant and require further investigation.

In their subjective ratings at the end of the experiment, SMEs expressed their faith in the potential of the 2 ½D imagery, as an aid for image analysis, improving SA, enhancing general orientation, understanding the structure of terrain and perceiving height and slope differences.

4.2 Recommendations

Future studies should include exploration of the 2 ½D SAR and other sensors (e. g., electro-optical), in a fused display format, to support IA confidence in performing SA and information extraction tasks. (This recommendation is based on observation of the SMEs strategies in carrying out the tasks.)

The use of a DTM overlay should be studied in conjunction with various types of sensor imagery under conditions where sensor imagery may disappear or fade out (e.g., passing through a cloud, degraded conditions for thermal imagery). It is hypothesized that under these conditions, the DTM may serve as an anchor, prevent loss of orientation and thus enhance orientation and object recognition performance.

SME training and individual differences may have played an important role in the present study. These issues require further investigation.

5. Acknowledgements

This study was conducted with the joint participation of the IAF Human Factors Engineering Branch (HFEB) and the Crew System Interface Division of the United States Air Force Research Laboratory's Human Effectiveness Directorate (AFRL/HEC). In any exploration of this type, the true knowledge resides with the SMEs of the operational units. The authors are greatly indebted to the young men and women of the Israeli Ground Corps Command and Air Force who shared their expertise so willingly with the experimenters in support of this research.

6. References

[1] Kuperman, G. G., & Shaya, D., *Subjective Assessment of SAR Imagery Enhancement Algorithms*, AL/CF-TR-1997-0133, Armstrong Laboratory, Wright-Patterson Air Force Base, Ohio, September 1997.

[2] The Technical Staff, Adroit Systems Inc., *An Evaluation of Required Tools for the Imagery Analyst*, AFRL-HE-WP-TR-1998-0005, Air Force Research Laboratory, Wright-Patterson AFB OH, October 1997.

[3] Endsley, M. R. (1994), *Situation Awareness in Dynamic Human Decision Making: Theory*, in Gilson, R., Garland, D., and Koonce, J. (eds.), *Situation Awareness in Complex Systems*, Daytona Beach, FL: Embry-Riddle Aeronautical University Press, 27-58.

[4] Joint Directors of Laboratories Sub-Panel on Data Fusion (1991), *Functional Description of the Data Fusion Process*, Department of Defense, Washington, D. C.

Human Performance and Data Fusion Based Decision Aids

Ann M. Bisantz , Richard Finger, Younho Seong, and James Llinas
Center for Multi-source Information Fusion
Department of Industrial Engineering
State University of New York at Buffalo
Amherst, NY 14260
bisantz@eng.buffalo.edu

Abstract – Decision-aids based on data fusion technologies may be applied to support decision-making in a variety of environments, ranging from military command and control situations to intelligent transportation applications. In any situation, the ultimate performance of human decision-maker/decision-aid system depends not only on the quality of the aid, but on the human decision-maker's utilization of the information provided by the aid. This utilization can be affected by many factors, including the degree of trust the decision-maker has in the aid, and the form in which information is presented to the decision-maker. This paper describes a framework for investigating trust in data-fusion based decision aids, and results from a pilot experiment in which distorted and blended graphical forms were used to represent uncertain information.

Key Words: decision-aids, trust, information displays.

1. Introduction

1.1 Data Fusion Based Decision-Aids

Decision-aids based on data fusion technologies may be applied to support decision-making in complex, dynamic environments such as military command and control, non-destructive testing and maintenance, and intelligent transportation. These aids provide operators with situational estimates which can aid in the decision-making process. For instance, in a military environment, data fusion based decision-aids may provide commanders with estimates of an entity's identity or threat potential. Regardless of environment, such aids provide decision-makers with information that has an associated level of confidence or uncertainty, through the application of automated algorithms and processes. The ultimate performance of such systems, consisting of both the human decision-maker and the automated decision-aid, depends on the human decision-makers' utilization of the information provided by the aid. Such utilization can be impacted by many factors, including the level of risk, time pressure, nature of the information display, and level of trust the decision-maker has in the automated aid.

This paper describes a research approach addressing the latter two factors in the context of a military environment. In a military context, data fusion has been identified as a means to perform assessments of identities, situations, and threat potential based on information derived from multiple electronic and intelligence sources. In these situations, the inherent risks, time pressure and large volume of data have led to the need for computerized aids performing automated data fusion (Walts and Llinas, 1990).

The process of data fusion in a military context includes multiple levels, each of which provides information at a different level of abstraction. For instance, different levels would address the detection and identification of potential targets, the association of targets into organized groups with certain behaviors, and the estimation of the threat potential of those groups. Thus, the results of data fusion processing can provide input to the situation assessment activities of battlefield commanders (Llinas, Drury, Bialas, and Chen, in press). Ultimately, information resulting from the data fusion process is presented to the human decision-maker through a computer interface.

1.2. Decision Aiding in an Adversarial Environment

Aided-adversarial decision-making (AADM) refers to military command and control decision making in environments in which computerized aids are available, and in which there is a potential for adversarial forces to tamper with and disrupt such aids. Hostile forces may attempt to compromise tactical decision-making through offensive activities conducted to attack or interfere with an adversary's

information systems. Information warfare can impact an adversary's operations through information disruption, denial, and distortion (Llinas, Drury, Bialas, and Chen, in press). For instance, disrupting or denying access to sources of information may make it difficult for decision-makers to assess situations and take appropriate actions. Distorted information, through manipulation and addition of incorrect information, may fool adversaries into taking actions desirable from a friendly perspective.

1.3 Human Trust in Automated Aids

Given the potential for information operations to disrupt and corrupt information provided by data-fusion based aids, it is necessary to understand the extent to which decision-makers rely on or use these aids, and factors affecting that reliance. A possible source of information regarding these issues is research that has been performed in the area of human trust in automated systems (e.g., Lee and Moray, 1992; Muir and Moray, 1996; Parasuraman, Molloy, and Singh, 1993; Sheridan, 1988). Researchers have suggested that trust can affect how much people accept and rely on increasingly automated systems (Sheridan, 1988).

Generally, research from both social science and engineering perspectives agree that trust is a multi-dimensional, dynamic concept capturing many different notions. For example, Rempel et al. (1985) concluded that trust would progress in three stages over time from predictability, to dependability to faith. Muir and Moray (1996) extended these three factors, and developed an additive trust model that contained six components: predictability, dependability, faith, competence, responsibility, and reliability. Sheridan (1988) also suggested possible factors in trust, including reliability, robustness, familiarity, understandability, explication of intention, usefulness, and dependence.

Empirical results have shown that people's strategies with respect to the utilization of an automated system may be affected by their trust in that system. For example, Muir and Moray (1996) and Lee and Moray (1994) studied issues of human trust in simulated, semi-automated pasteurization plants. These studies showed, among other results, that operators' decisions to utilize either automated or manual control depended on their trust in the automation and their self confidence in their own abilities to control the system. Additionally, results showed that trust depended on current and prior levels of system performance, the presence of faults, and prior levels of trust. For example, trust declined, but then began to recover, after faults were introduced (Lee and Moray, 1992). Lerch and Prietula (1989) found a similar pattern in participants' confidence in a system for giving financial management advice: confidence declined after poor advice was given, then recovered, but not to the initial level of confidence.

In the context of AADM, there exists the potential for several circumstances in which trust in data-fusion based decision aids could be affected. For instance, information warfare techniques could be used by an adversary to distort the information provided by decision aiding systems, disrupting (appropriately) commanders' trust in, and utilization of, such systems. Alternatively, an adversary might act deceptively, fooling a commander into trusting and acting based on information in a way favorable to the adversary. Finally, an adversary might disrupt a commander's trust in an aid that is providing good ("trustworthy") information. For these reasons, it is necessary to investigate human trust in AADM situations, in order to better understand how data-fusion based decision aids will impact the decision-making process under different circumstances.

2.0 Investigations of Decision Aiding in Adversarial Environments

2.1 Theoretical Framework

To structure the investigation of aspects of human trust in data fusion-based decision aids, a multi-dimensional framework was developed (Llinas, Bisantz, Drury, Seong, and Jian, 1998). The framework integrates and systematically varies a set of dimensions which may affect trust in decision aids. The following dimensions are included in the framework:

1. Locus of Attack. One potential factor is the location at which the potential for corruption exists. Two potential dimensions can contribute to this factor: the component dimension, and the surface-depth dimension.

a) Component Dimension. Information could be corrupted at a variety of components, or levels, in the AADM environment. Information could be corrupted at the level of the tactical situation (by interfering with sensors), within the information processing and data fusion algorithms that comprise the decision aids, or at the level of the human-computer interface.

b) Surface-Depth Dimension. A second related dimension along which investigations of performance in AADM systems can vary is a surface-depth dimension. The surface level corresponds to the

information available about the environment (as formalized in Brunswik's Lens Model; Cooksey, 1996; Hammond, Stewart, Brehmer, and Steinman, 1975), whereas the depth level corresponds to the actual state of the environment. In an AADM environment, surface level features would be the observable outputs from sensors, or data fusion processes. Depth level features would be the actual operations of the sensors or algorithms themselves.

2. Malfunction Level. Information aids for AADM can fail or be corrupted in qualitatively different ways, either failing completely, or being partially degraded, resulting in two malfunction levels:

- Element failure. System components can fail completely resulting in a loss of data.
- Element degradation. The quality of information provided by the system component can be degraded, resulting in partial information loss, or increased ambiguity and uncertainty.

3. Causes of Failure or Corruption. Information can be corrupted through different causes or intentions, ranging from naturally occurring system failures (e.g., hardware malfunctions), to deliberate attacks on the information systems, to deliberate attacks which are disguised by the adversary.

4. Time Patterns of Failure. A final dimension reflects the dynamic or time-dependent characteristics of the degradation. Failures, sabotage, and subterfuge can occur not only as failures or degradations at a particular point in time, but also in a continuing fashion. Additionally, failures can occur with patterns that are either predictable or unpredictable.

2.2 Framework-based Experiments

This framework is being used to develop experiments in the area of human trust in data fusion-based decision aids. At present, experiments are planned to investigate changes in trust in, and reliance on, a data-fusion aid when the situation is framed as either one in which the aid may be unreliable due to hardware failures, or one in which the aid may be subject to deliberate tampering by an adversary. Participants will perform a simulated military command and control task in which they will identify unknown aircraft moving on a radar screen.

During the task, participants will be able to access both non-aid information (e.g., altitude, radar emission, and speed information) about unknown aircraft, as well as an identity estimate from a simulate data-fusion aid. The identity estimate will be in the form of a probabilistic range (e.g., that an aircraft is friendly or hostile). Participants will request access to either type of information, and will be limited in the number of requests, forcing a tradeoff between information sources.

Participants will perform the experiment over six scenarios, during which time the speed and altitude of the aircraft will vary within pre-defined, overlapping ranges. Prior to the experiment, participants will be given conditional probability information about the chance that an aircraft is hostile, given that it is flying at a particular speed altitude, or has a particular radar signature.

After several three normal scenarios, a fault (either a constant shift in the probabilistic range, or a gradually increasing range) will be introduced into range provided by the data-fusion aid. Participant's reliance on either form of information (either the decision aid, or the other available information) will be measured before and after the insertion of the error to assess the potential loss of trust in the aid subsequent to the error.

3.0 Investigations of Data Presentation

As noted above, one factor which may influence the utility of data fusion based decision aids, and the influence of these aids on the decision making process, is the form in which the uncertain information determined by these aids is presented to decision-makers. Uncertain or probabilistic Information can be shown in a variety of formats ranging from simply text to graphical representations to text/graphical hybrids. Past research has focused on representing position, direction and identity uncertainty in a format that reveals the true probabilistic nature behind the data (Andre and Cutler, 1998; Banbury, Selcon, Endsley, Gorton, and Tatlock, 1998; Kirschenbaum and Arruda, 1994).

Position uncertainty deals with how to represent the possible places an object may inhabit. Environments in which this type of uncertainty plays an important role include commercial aviation and military sonar/radar. Andre and Cutler(1998) investigated this form of uncertainty with the use of a task in which a pilot would have to play "Chicken" with a circular object, they called a meteor. The pilot's goal was to come as close as possible to the meteor without collision. To represent the position uncertainty a circular ring surrounded the meteor. The ring varied in size dependent upon uncertainty level. Collision frequency was found to be far less when the ring was displayed: without the ring, participants appeared to dismiss the fact that uncertainty was present in the system. Kirschenbaum and Arruda (1994) conducted a similar experiment which investigated the effect of

different displays of position uncertainty on a decision-making task as to when and where to fire at a target. Participants were shown either a graphical representation of position uncertainty in the form of an ellipse around the target or a verbal indicator that ranged from poor to fair to good. The elliptical aid was found to be superior to the verbal in cases of moderate to high difficulty scenarios. Overall it appears that the use of a visual position uncertainty aid helped the performance of the user.

Aids which present heading uncertainty attempt to display all the possible future directions an object may move. Andre and Cutler (1998) tested three different types of heading uncertainty aids in a simulated anti-aircraft task: a textual description and two graphical representations that utilized either arcs or rings. The three aids improved user performance when compared with a no aid condition. The arc-based aid, which represented the uncertainty in direction by utilizing an arc that covered the entire angle of possible movement heading, provided a slight advantage over the other two aids.

Finally, identity aids strive to give the user an idea of how accurate the identification of an object is. Currently most aids display this information in the form of probabilities. Banbury et al. (1998) investigated how the context in which information is displayed affects a decision-making task. Participants were asked to make a shoot/no-shoot decision based on a probabilistic estimate of an aircraft's identity, presented as a numeric percentage. Results showed an impact of estimate uncertainty - participants were found to have a reluctance to shoot when uncertainty was greater than 9%. Additionally, presenting a secondary target identification (e.g., not just the chance that is a hostile fighter, but also the chance that it is a friendly aircraft) also impacted decisions to shoot. Participants were more hesitant when a secondary, friendly, target identification estimate was given.

Another way in which the graphical form of information presentation could be used to represent uncertainty is through the use of degraded or distorted images. Lind, Dershowitz, Chandra, and Bussolari (1995) provide evidence that the form of displayed information may affect the use of uncertain data. In a study to investigate the extent to which the graphic depiction of weather systems could be degraded (due to technical limitations) and still be acceptable to general aviation pilots, Lind et al. found that pilots' estimates of weather hazards increased as the graphical distortion increased. In this case, the distortion took the form of larger polygon/ellipse shaped depictions of weather patterns, in contrast to the non-distorted continuous, fine-grained representation. This increase in perceived risk might indicate a decrease in subjects' confidence of their understanding of the current specific weather patterns.

Thus, there is some indication that iconic representations based on degraded or distorted images may be used to convey the uncertainty associated with a decision aid estimate. In the following pilot study, we investigated properties of distorted and blended icon sets intended to convey uncertain information about an object's identity as either potentially hostile or friendly. Future experiments will investigate the impact of a subset of these icons, selected based on the pilot study results, on a decision-making task.

3.1 Pilot Study Method

3.1.1 Participants

Twenty participants, all undergraduate students, were paid $6.00 per hour for their participation in the pilot study.

3.1.2. Experimental Design

Five sets of pictures were chosen to represent the identity of an object as either hostile or friendly. These picture sets were classified as either abstract (without an obvious associated meaning), iconic (with an associated meaning), or both. Picture pairs were chosen in order to allow for the entire spectrum from friendly to hostile to be represented. Figure 1 shows the pictures used in the experiment.

In order to represent the probabilistic nature of the information graphically, a series of thirteen icons were created to represent a range of probabilities (i.e., from p(Hostile) – 0.0 to p(Hostile) = 1.0). The iconic and abstract picture pairs were distorted and blended using a pixelizing function found in Adobe PhotoShop 4.0. For example, the 50% friendly/50% hostile picture blended both of the pictures in a pair together. For the colored icons, the series of icons was created by coloring each pixel in the icon as either green or red based upon the probability desired. To illustrate how the pixelizing function works, the series of the distorted and blended pictures for picture pair (1) are shown in Figure 2.

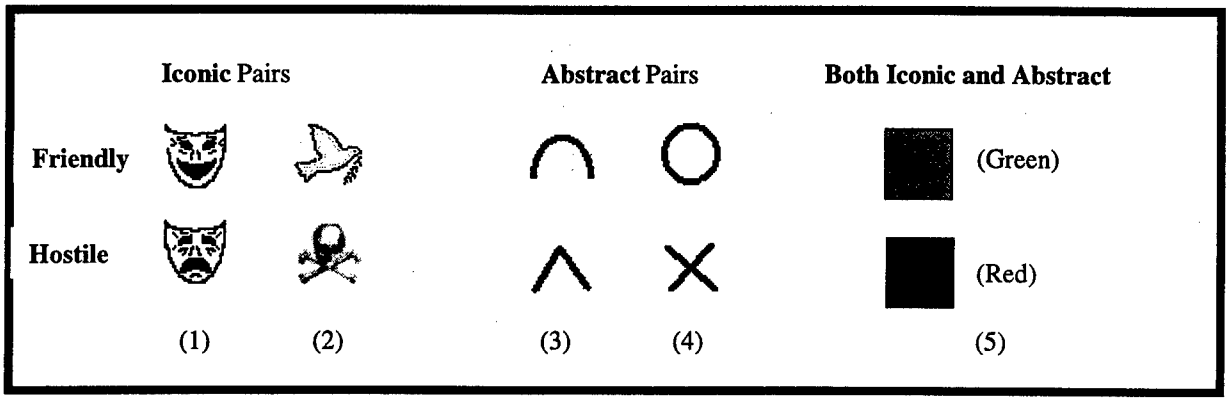


Figure 1. Five pairs of icons representing object identities as either hostile or friendly.

Each participant performed a series of tasks involving all five sets of icons. Ten participants performed the tasks under a "friendly" framing condition, and ten participants performed the tasks under a "hostile" framing condition. In the friendly framing condition, participants were given task instructions which described the icons as more or less friendly. In the hostile framing condition, icons were described as more or less hostile.

3.1.3 Procedure

The three experimental tasks were designed to measure whether the icons could be correctly sorted and assigned a probability rating according to the expected probabilities that the icons represented. Participants performed each of the tasks five times: once for each icon pair (see Figure 1).

In the first task, a timed sorting task, participants were asked to sort cards into piles according to the icon printed on the card. Participants were asked to create piles containing the same icon. There were five instances each of the 13 possible icons in a set, for a total of 65 cards. The time to sort the cards, and sorting errors, were collected.

In the second task, participants were asked to order the set of thirteen pictures from most to least friendly (or hostile), depending on the framing condition. They were not told which icons corresponded to the hostile or friendly ends of the scale (e.g., they were not told that a circle represented a most friendly, and an "x", least friendly). Participants performed this task using a Visual Basic computer program, through which they could drag and drop the icons into the desired order. The ordering of the icons was recorded automatically by the computer.

For the third task, participants were asked to rate each icon on continuous scale, with end points of least and most friendly (or hostile). Participants marked their rating along a line connecting the endpoints; this distance was later measured and scaled based on the length of the line, and used to identify their rating.

3.2 Pilot Study Results

3.2.1 Card Sorting

The times to sort cards based on the icon printed on the card did not differ significantly across picture pairs. Thus, the relative difficulty of identifying and sorting the thirteen icons did not appear to differ across sets.

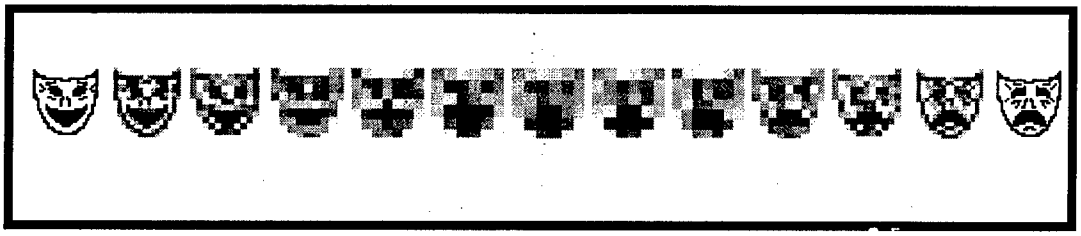


Figure 2. Series of 13 icons representing a range of probabilities that an object is hostile or friendly: from a probability of 100% friendly to 100% hostile.

3.2.2. Ordering

The order of the thirteen icons in each icon pair set was determined for each participant, for the hostile and friendly framing conditions, resulting in ten orders per icon pair for each framing condition. These orders were used to compute an average ranking for each icons for the five pairs, for both framing conditions. Ordering these average rankings resulted in an average order for each set, for both framing conditions (a total of 10 average orders). These average orders were correlated with the expected order (based on the way the icons were created), and a Spearman correlation coefficient was computed. These coefficients are shown in Table 1. All correlations were significant at the .01 level of significance, indicating that overall, participants were able to correctly order the sets of icons according to the intended levels of uncertainty.

Table 1. Spearman Correlation Coefficients comparing average rank orders to expected order for 5 Icon Pairs.

Icon Pair	Framing	
	Friendly	Hostile
Mask(1)	1.000	0.929
Dove(2)	1.000	0.984
Inverted V-U (3)	0.934	0.984
Circle-X(4)	1.000	0.951
Color(5)	1.000	0.890

Individual participant data was also examined: Spearman correlation coefficients were computed comparing each participant's order to the expected order, for both framing conditions. These correlations are indicated in Tables 2 and 3, corresponding to the Friendly and Hostile framing conditions, respectively. Correlations in bold are *insignificant* at the .05 level of significance. Inspection of Tables 2 and 3 shows that on a participant-by-participant basis, ordering was more consistent and correct in the friendly framing condition than the hostile framing condition. Note that negative correlations simply indicate that the participant reversed the hostile and friendly ends of the scale (they were not told which icons corresponded to which endpoints before the experiment). It is interesting to note that even for the two "abstract" icons, reversals happened at a rate less than chance, indicating that perhaps there was some meaning intrinsic to the abstract icons.

Table 2. Individual Correlation Coefficients for each participant (Friendly framing condition; bold correlations are *insignificant*).

P's	Mask (1)	Dove (2)	V-U (3)	Circle (4)	Color (5)
2	0.995	1.000	-1.000	1.000	0.995
4	0.989	0.995	1.000	1.000	1.000
6	1.000	1.000	0.995	1.000	1.000
8	1.000	1.000	-1.000	1,000	0.995
10	1.000	1.000	1.000	1.000	1.000
12	1.000	-1.000	-1.000	-1.000	-1.000
14	0.984	1.000	1.000	1.000	0.995
16	0.962	1.000	1.000	1.000	1.000
18	0.995	0.995	1.000	1.000	1.000
20	0.978	1.000	**-0.440**	**0.374**	1.000

Table 3. Individual Correlation Coefficients for each participant (Hostile framing condition; bold correlations are *insignificant*).

P's	Mask (1)	Dove (2)	V-U (3)	Circle (4)	Color (5)
1	**0.126**	**0.115**	**0.115**	0.115	**0.115**
3	1.000	1.000	1.000	1.000	0.995
5	0.566	**-0.038**	0.544	**-0.297**	0.665
7	**0.412**	**0.093**	**0.115**	**0.148**	**0.088**
9	**0.005**	0.714	**0.099**	**0.044**	**0.181**
11	0.978	1.000	1.000	1.000	**0.434**
13	**0.148**	1.000	1.000	0.995	**0.456**
15	0.978	1.000	0.978	0.995	0.989
17	0.995	0.995	1.000	1.000	1.000
19	**-0.165**	0.516	**0.280**	**0.440**	0.835

3.2.3 Rating

From the data collected on individual picture ratings an average rating was calculated for each picture within a picture pair category. These averages provided a range of estimates of the friendliness or hostility of each picture pair (Tables 4 and 5).

Table 4. Rating Spread for 5 icon pairs (Friendly Framing)

	Mask (1)	Dove (2)	V-U (3)	Circle (4)	Color (5)
High Rating	88.67	97.93	96.64	97.73	98.59
Low Rating	4.22	4.06	3.67	8.52	11.33

Note: Ratings for Dove, V_U, and Circle were corrected to account for obvious and consistent reversals between hostile and friendly endpoints.

Table 5. Rating Spread for 5 icon pairs (Hostile Framing)

	Mask (1)	Dove (2)	V-U (3)	Circle (4)	Color (5)
High Rating	66.56	72.27	62.11	74.06	63.69
Low Rating	24.22	17.97	38.20	21.48	14.30

3.2 Future Experiments

Future experiments will test the effect a subset of these icon pairs on decision-making in a dynamic identification task. Participants will be asked to identify objects as either friendly or not friendly, given a graphical icon of the object which depicts a decision-aid's probabilistic estimate of the object's identity. This icon will be based on either end-point icons with associated numeric probabilities, the full range of 13 icons, or the full range of 13 icons with associated numeric probabilities. Over time, estimates will tend (with some randomness) to become more certain; however, participants will be penalized for identification delays. The experiments will investigate the impact of information presentation on the point at which participants choose to identify objects. If graphical depictions (i.e., distorted icons) convey more information about the probabilistic nature of the identity estimate than numeric probabilities, then participants seeing the graphical depictions should choose to wait to make an identification until they are more certain.

4.0 Conclusions

Data fusion-based decision-aids can be implemented to provide support in a variety of situations. In order for those aids to provide effective support, the must provide information in a format that conveys important aspects of that information (e.g., its uncertain nature) and be trusted by the decision-maker. A framework for investigating trust in decision –aids, in adversarial decision-making situations, along with on-going experiments based on that framework, was discussed. Additionally, results from a pilot study investigating the utility of degraded and distorted images to convey levels of uncertainty were presented. Preliminary results indicated that sets of distorted icons could be appropriately ordered, and span a range of descriptive level, under particular framing conditions. Future experiments to investigate the effect of these representations on decision-making were described.

5.0 References

[1]. Andre, A. D., and Cutler, H. A. (1998). Displaying Uncertainty in Advanced Navigation Systems. In *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting* (pp.31-35).

[2] Banbury, S., Selcon, S., Endsley, M., Gorton, T., and Tatlock, K. (1998). Being Certain About Uncertainty: How the Representation of System Reliability Affects Pilot Decision Making. In *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting* (pp.36-39).

[3] Kirschenbaum, S. S., and Arruda, J. E. (1994). Effects of Graphic and Verbal Probability Information on Command Decision Making. *Human Factors*, 36(3), 406 – 418.

[4] Lee, J. D., Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, 35(10), 1243-1270.

[5]Lee, J. D. and Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies*, 40, 153-184.

[6] Lerch, F. J., and Prietula, M. J. (1989) How do we trust machine advice? In Salvendy, G. and Smith, M. J. (Eds.) *Designing and using human-computer interface and knowledge based systems.* Elsevier Science Publishers, North-Holland.

[7] Lind, A. T., Dershowitz, A., Chandra, D., & Bussolari, S. R. (1995). The effect of data link-provided graphical weather images on pilot decision making. *In IFAC proceedings – 1995.*

[8] Llinas, J. Bisantz, A. M., Drury, C. G., Seong, Y., Jian, J-Y. (1998). *Studies and analyses of aided adversarial decision-making. Phase 2: research on human trust in automation.* Center for Multi-Source Information Fusion Technical Report. State University of New York at Buffalo, Amherst, NY.

[9] Llinas, J., Drury, C., Bialas, W., Chen, A. C. (in press). *Studies and Analyses of Vulnerabilities in Aided Adversarial Decision Making.*(AFRL/HE-WP-TR-1998-0099) Wright-Patterson AFB, OH: Air Force Research Laboratory, Human Effectiveness Directorate, Crew System Interface Division.

[10] Muir, B. M., Moray, N. (1996). Trust in automation: Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics*, 39(3), 429-460.

[11]Parasuraman, R., Molloy, R., Singh, I. L. (1993). Performance consequences of automation-induced "complacency". *The International Journal of Aviation Psychology*, 3(1), 1-23.

[12] Rempel, J. K., Holmes, J. G., and Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, 49(1), 95-112.

[13] Sheridan, T. B. (1988). Trustworthiness of command and control systems. In *IFAC Man-Machine Systems*, 427-431, Oulu, Finland.

[14] Walts, E. and Llinas, J. *Multi-sensor Data Fusion.* Norwood, MA: Avtech House.

Literature Survey on Computer-Based Decision Support for Command and Control Systems

ELISA SHAHBAZIAN, Jean-Rémi Duquet, Marc-Alain Simard
Lockheed Martin Canada, 6111 Royalmount Ave., Montréal, Québec, H4P 1K6, Canada
tel: (514) 340-8310, extensions 8343, 8547, 8660, fax: (514) 340-8318
elisa.shahbazian@lmco.com, jean-remi.duquet@lmco.com, marc-alain.simard@lmco.com

Abstract - *Since 1991, the Research and Development (R&D) group at Lockheed Martin Canada (LM Canada) has been developing and demonstrating technologies which will provide Observe-Orient-Decide-Act (OODA) decision making capabilities/tools in Naval and airborne Command and Control (C2) for application on Canadian Patrol Frigates (CPF) and Canada's CP-140 (Aurora) fixed wing aircraft. Over the last three years LM Canada has also established a generic expert system infrastructure and has demonstrated that it is suitable for integrating these decision making technologies into real-time Command and Control System (CCS). However, before these technologies become integrated into the C2 of any operational platform it is important to understand how should these decision making tools function and be integrated into the CCS to ensure that the human operators trust, accept and use these tools successfully. To help understand such issues LM Canada performed a literature survey and collected and analyzed over 600 papers on this subject. This paper presents the results of this survey and some conclusions made for Naval C2.*

Keywords: Decision Support Systems, Blackboard, Testbed

1. Introduction

Canada's Halifax Class Canadian Patrol Frigates (CPF) and CP-140 (Aurora) fixed wing aircraft are planned to be upgraded within the next decade to be able to deal with far more demanding threat and mission environments of today and the future, than when these platforms were designed. The computer hardware and software capabilities of today permit the development of considerably more advanced decision support capabilities, compared with the capabilities existing on these platforms currently, helping them to deal with these new environments. Over the last 9 years the Research and Development (R&D) group at Lockheed Martin Canada (LM Canada) in close collaboration with Canada's research laboratories has been developing and demonstrating technologies which will provide Observe-Orient-Decide-Act (OODA) decision making capabilities/tools in Naval and airborne Command and Control (C2) for application on CPF and Aurora.

The research has been proceeding in a number of parallel activities including:

1. Algorithimc solutions for the decision support tools,
2. Testbed infrastructure for demonstrating these solutions,
3. Top-down systems analysis to understand the operational and mission requirements of these systems and the shortcomings of the existing systems.

The results of these research activities are incrementally being built into demonstration systems for the operators to observe and experiment with, and their feed-back is being used in the next iteration.

To ensure that these research activities are conducted in a systematic manner, a number of literature surveys have been conducted over the life of this program since 1991. The first was a survey into the technologies and algorithms for decision making tools, which started in 1991 as a contract from the Defence Research Establishment Valcartier (DREV) in 1991 and is still on-going. The second is the survey initiated in 1998 of the basic and applied literature on dynamic decision making and computer-based decision support in dynamic decision-making environments, to help understand how should the decision making tools function and be integrated into the CCS to ensure that the human operators trust, accept and use these tools successfully. This survey also was conducted as a contract from DREV.

This paper presents LM Canada's approach in applying the results of this recent survey for the development of the Decision Support System of the CPF.

2. The Current Infrastructure

Over the last three years LM Canada has established a generic expert system infrastructure and has demonstrated that it is suitable for integrating these decision making technologies into real-time Command and Control System (CCS). Figure 1 shows the CPF testbed established based on this architecture.

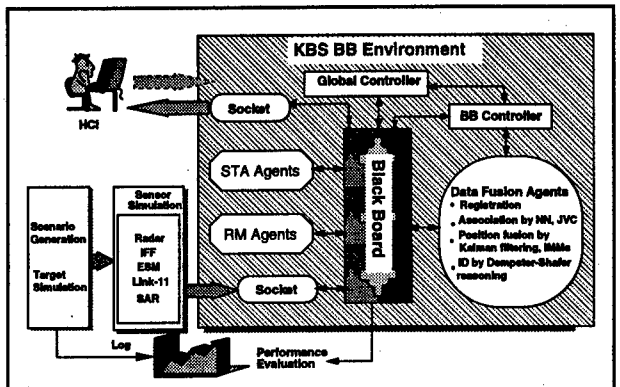


Figure 1: The CPF Decision Support System Demonstration testbed

This architecture was developed by LM Canada in collaboration with DREV and uses a Knowledge Based System (KBS) shell based on a Blackboard (BB)-based problem-solving paradigm. Details of this architecture have been published previously [1, 2, 3].

The major advantages of this architecture are:

1. It is able to support distributed real-time large applications,
2. Permits modular, incremental parallel and independent development, and
3. Permits implementation of numeric, mathematical and rule-based, heuristic algorithms within the same infrastructure.

The CPF testbed shown in Figure 1 has a closed loop simulation system, permitting the users to observe the decision support capabilities impact on the threat environment, as well as a very modular Human-Computer Interface (HCI), permitting the developers to experiment with various approaches for integrating and providing these decision support tools to the users and to apply the user feed-back.

The initial decision support capabilities that were implemented and demonstrated within this testbed were very close to the ones already existing within the current CPF Command and Control System (CCS). This was done to establish the initial baseline, ensuring that the users have a frame of reference. Next, based on a an internal fast review of the literature some additional decision support capabilities were added. Overall the currently available DSS capabilities include:

1. Multi-Source Data Fusion (MSDF):
   - Position estimation enhanced through:
     - Ellipsoidal gating including attribute data
     - Jonker, Volgenant and Castanon (JVC) for track/contact association
     - Adaptive Kalman Filter or IMM filters or 3 adaptive parallel filters for track estimation
     - Dissimilar data fusion (1D to 2D to 3D)
   - Target identification enhanced through automatic ID recommendadtions at all ranges based on any data available using:
     - Truncated Dempster Shafer for identity estimation capable to fuse any type of information
       - » Fuzzyfied kinematics
       - » ESM and IFF data
       - » Other misc sources of information

2. Situation and Threat Assessment (STA):
   - CPF-like Threat Ranking
   - Clustering
   - Rule based allegiance
   - Commercial corridor correlation
   - Maneuvering target detection
   - Track splitting detection
   - Fast incoming target criterion
   - Ownship Missile recognition
   - Mean Line of Advance

3. Resource Management (RM):
   - CPF-like Reactive Planning:
     - Point of Intercept
     - Point of first fire
     - Target Weapon Pairing
     - Weapon Designation
     - Resource Allocation
   - Deliberative Planning:
     - Decision tree (plan) creation
     - Plan evaluation/ optimization
     - Plan repair

At this point, before any further technological capabilities are developed, it is necessary to understand how these new tools should be validated and integrated with the CPF C2, and what approach should be adopted to develop the computer based DSS (CBDSS) of the future CPF. Hence a more systematic literature survey was initiated.

3. The Literature Survey

The survey included the basic and applied literature on dynamic decision making and computer-based decision support in dynamic decision-making environments. The review was divided in four (4) distinct tasks[4]. The tasks were:

Task I -- Identification of Tools and Information Sources

Task II -- Development of a Survey Methodology

Task III -- Literature Search and Classification

Task IV -- Results Analysis and Recommendations

Close to 600 references were found using the various channels identified during Task I.

The search was done using the keywords listed in **Table 1.** The first column of this table presents five themes that we felt would encompass all the topics of the literature search. The second column proposes topics that subdivide a theme into more specific subjects.

Table 1: Literature Search Topics

Themes	Sub-Themes (Phase I)
Computer-based decision support	Decision aids Decision Support Systems (DSS) Performance Support Systems (PSS) Trust in knowledge-based systems
Process control and computer-based aiding	Generic tasks, Work procedures Ecological interface design Human performance models Information visualization
Cognitive task analysis	Individual work Cooperative work
Cognition	Decision making, Situated cognition Distributed cognition, Mental models Socially shared cognition Human Performance, Mental workload Cognitive styles, Human expertise Human reliability and error
Human-Computer Interaction	Task analysis for HCI, Interaction styles User-centred system design methodologies On-line help and documentation Information Presentation Intelligent interfaces Usability engineering/ Evaluation

The literature found at this point was analyzed and further sorted based on their pertinence on the CPF CCS.

Based on the findings in the first three tasks it was concluded that the Results Analysis and Recommendations can take a number of different perspectives:

1. A theoretical analyses of the realization of a computer-based DSS for the future shipboard CCS that should be part of the integrated combat system. This includes summation from the Literature Survey of the concepts, models, methods, results, principles and guidelines for building dynamic systems that can help real-world decision-makers do their job more effectively and safely. This study addressed the new discipline of Cognitive Engineering (CE), the characteristics of dynamic and naturalistic environments and of the tactical combat environment, different levels of automation in computer-based systems and the place of a DSS, how to model a work domain or a complex system, the concept and characteristics of naturalistic decision making, different models of human behaviour and decision making, characteristics of naturalistic decision making, the question of how to aid the human operator at work, Ecological Interface Design (EID) framework and a proposal for improving it and presented several recommendations for building a well-engineered DSS within CCS.

2. A more practical, but generic approach for establishing a CBDSS within an existing large CCS. This approach recommended a complete spiral process for Human-Machine System Design that takes into account both human and technological aspects of system development in the specific context of the design and implementation of a CBDSS for the Halifax class ships. For each phase of the development process a set of potential human engineering tools have been described, in some cases a preferred approach was selected, in others the question was left open until the issues are better understood. The following tasks have been identified as the cornerstones of the CBDSS development process in each phase of the spiral:
    - System Analysis,
    - Task Allocation,
    - System Development and implementation,
    - HCI, developed using a prototyping approach,

- System Evaluation, with 3 distinct areas of interest: (1) Function Usability; (2) Operational Impact; and (3) User Fit.

3. The third perspective was to address the specific example of the technologies currently under development at LM Canada for the Halifax class ships and their impact on the functions on the frigate operators.

The next section focuses on the generic approach for establishing a CBDSS within an existing large CCS.

4. The CBDSS Development Approach

This approach is based on the theoretical analyses of the realization of a computer-based DSS for the future shipboard CCS that should be part of the integrated combat system, and the identified and recommended approaches for Cognitive Analysis in the surveyed literature. It tries to present these methods in a structured frame, from the perspective of System Design, taking into account various constraints that are often left aside in cognitive engineering literature (e.g., technological uncertainty, compatibility with accepted System Design Frameworks, feedback loop after system evaluation, etc.). It also introduces constraints and requirements driven by the scope and specific context of a decision support system for the Halifax Class ships.

The purpose of this section is therefore to present a global approach to Human-Machine Systems Design that takes into account both human and technological aspects, and that is suitable for the development of a CBDSS for the Halifax Class ships.

The current CCS of the Halifax Class was developed under a so-called "classical" System Design framework. Given the computer power available back then, which directly impacted the level of automation and the amount of information available to the operator, the CCS design emphasized primarily the automatic system; the so-called Threat Evaluation and Weapon Assignment (TEWA) system was (and still IS) performing mostly numerical and simple rule-based calculations, while most of the higher-level (cognitive) activities were left to a team of operators.

This approach to System Design aims to incorporate both the technology and the operator under a so-called Cognitive System Design Framework. In particular, we want to identify which of the approaches described in the literature is better suited for the development of a CBDSS in the specific context of the mid-life upgrade of the Halifax Class.

From the cognitive engineering and system design literature, a common trend in the way human factors should be included as part of the traditional system design approach can be identified. Figure 2 is drawn from a combination of a number of approaches to system design such as *human-machine system design "frameworks"* and *user-centred system design methods* (built from [5] and [6]).

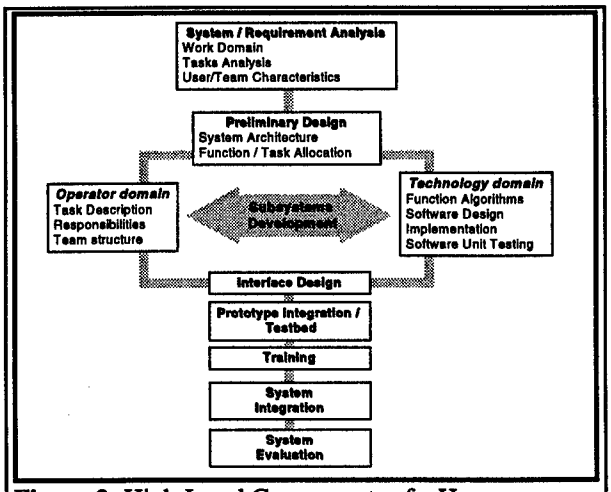


Figure 2: High-Level Components of a Human-Machine System Design Framework

This representation of the system design process seems quite "natural" to any experienced system or software designer, except that it allocates a larger place for concerns about the end user in the early stages of the design. The main difference is the dual nature of the subsystems development phase, for which the authors recommend a two-team approach, since typically the "human factors" experts generally will not be "technology" experts, and vice-versa. It is assumed that the "human factors" team will be heavily involved in the interface design and system testing phases.

This picture, even though it seems complete and coherent, lacks a major component, namely the sequence and feedback loops, both *between AND inside* each subphase.

Literature on system design presents at least three mature life-cycle methodologies that have been extensively used and documented in the past to develop large software systems: Waterfall, Prototyping and Spiral.

Before we select one of these System Design approaches, and make it compatible with cognitive

engineering guidelines and methodologies, we need to identify the various constraints and considerations that will drive our choice. The following constraints and points have been identified from the literature, as well as from our knowledge of the context of Halifax Class and DSS issues:

1. It is widely accepted that for the design of complex systems such as the current DSS, a Top-Down approach is recommended. This means that the system development should proceed from the general to the specific in terms of its components; for example, system analysis should first describe the global picture, then refine this picture in terms of subsystem, components, tasks, etc. until the system is defined well enough to allow design and implementation. This view is compatible with all the models presented above.

2. Task splitting, i.e., the allocation of functions between the operator and the automatic system, requires a good estimate of the "algorithmic" performance of the automatic part of the system. Unfortunately, in some systems these algorithms won't be developed and tested until after the task splitting activity. Many cognitive analysis papers fail to take technological developments into account, thereby implicitly assuming that little technological uncertainty remains at the design phase and that the project risk mostly lies in the task splitting and interface design activities. *These assumptions are incorrect in the case of a CBDSS for the* Halifax Class. The design framework that will be selected will comprise a System Evaluation phase which should validate some high-level concepts such as User Fit (Situation Awareness, Communication Effectiveness) and Operational Impact of the complete integrated system. Because of the scope and complexity of the project, because of unpredictable technological performance, and also because of some unavoidable "ad hoc" task allocation included in the initial design, it is ***very likely*** that at the system evaluation stage, some initial task allocation decisions are overturned, thereby impacting the whole design and implementation cycle. Therefore the selected approach should provide a feedback mechanism to properly address incorrect task allocation or performance prediction, from the results of the evaluation of the joint human-machine system. This strongly suggests a spiral approach to system design.

3. An important paradox exists throughout the cognitive engineering literature, when approaching the problem of selecting a "cognitively sound" system design framework. This paradox is well described in [6]:

   "(literature on human-system interactions) clearly establishes a pressing need to evaluate throughout the system development cycle, from concept formation to final acceptance and testing. (...) There is a balance to be achieved between conflicting needs. On one hand there is the need to accurately predict final system performance in the field with typical users working under realistic conditions. On the other hand, this prediction needs to be based on something less than the system itself. ***In particular, major decisions made at the concept level that misunderstand the nature of user needs or the operational environment, need to be caught before there has been a major investment in design or production."***

   This implies a "testbed", and the closer it is from the expected "final" system, the better the input to the system design. We are therefore caught in a situation of "deadlock", where we would need a working prototype of the system in order to properly design this system in the first place. In the absence of a prototype of the "final" system, the designer must rely on two inputs: an existing, incomplete system on which experiments and observations can be made according to cognitive engineering principles, and a set of "educated guesses" on the optimal "final" system. Clearly, the larger the gap between existing and final system, the larger the number of designer's "guesses", the bigger the risk of identifying major misallocations and design problems at the later stages, and the larger the cost of iterating on the design and implementation to correct them. The selected framework should therefore try to minimize - or to segment - the gap between the "initial" and "final" system. Again, this strongly points towards a spiral approach to system development.

4. A serious concern with large-scale projects such as a DSS for the Halifax Class is the risk that several system requirements change in the course of the project, or that new ones appear as a result of changing doctrine, main mission objectives, input sources or information needs. The scope and nature of the project also makes it very unlikely that all system requirements will be correctly identified and addressed up front at the beginning of the project (i.e., in the first few years). These

considerations call for a framework which allows incorporation of new requirements late in the system development cycle, something the waterfall approach does not permit in principle.

5. Another issue that follows directly from the previous consideration is the intended scope of the whole CBDSS design process. This will drive the important question as to *where to start the investigation, what constitutes an acceptable risk* and *what level of effort is realistic* in the context of the project. Sure enough, an ideal analysis would incorporate a complete redefinition and redesign of the control process on Halifax Class, relying on a complete, scientifically accurate, in-depth analysis of the work and task domains, and detailed models of the cognitive processes of the team of operators. Given the current context of the timelines, budgets and expectations, the affordability of selecting such Cognitive Analysis Frameworks (CAFs) is not obvious.

6. Along the same line, a point that should not be overlooked is that the Halifax Class ships are already operational, *and fully functional* given today's operational requirements and information sources. The known shortcomings/deficiencies of the existing Combat System *are not likely to be judged significant enough to justify any major redesign of the decision support systems available to* Halifax Class *operators*. Therefore it is probably unnecessary to aim for a complete redesign of the whole system, and it is likely that any new CBDSS for the Halifax Class will have to build up to a certain extent on the existing architecture and algorithms. As a consequence, the selected approach will probably need to accept *constraints dictated by the existing* Halifax Class *system* as a key input of the analysis.

All these concerns and issues directly impact the choice of a suitable high-level system design framework useable for the development of a CBDSS for the Halifax Class, and also affect (even though to a lesser extent) the recommendations we can make on specific Human-Machine system methodologies to be used in each phase of the System Design life cycle.

Considering the issues mentioned above, and considering the respective advantages and drawbacks of each proposed framework, the recommended approach to Human-Machine System Design is to follow a Spiral approach, as detailed below.

If the Waterfall approach was seen as a potential approach at the start, despite the scope of the project, the need to support potential design iterations, the number and complexity of initial system requirements, as well as the potential consequences of late discovery of requirement or task allocation problems are all serious concerns, which make the **waterfall approach extremely risky and impractical for the design and implementation of a CBDSS for the Halifax Class.**

The Prototyping approach is not suitable as a framework for the complete system development, first because of its less formal structure, and also because it suffers from some of the drawbacks of the Waterfall approach, namely the fact that the requirement analysis and task allocation are made up front, at the beginning of the project. However, the prototyping model is very appropriate for some of the components where a large amount of technological uncertainty remain and which involve a research component, for instance in the area of **software and algorithms development**. The **Testbed** and **HCI** development constitute other examples. Such a prototyping development of subcomponents is intrinsic to the spiral approach proposed for the complete system.

The Spiral model of System Design shown in Figure 3 allows an iterative sequence of requirements/ design/ development/ evaluation cycles, incorporating a prototyping approach to system development as a risk mitigation mechanism at each new cycle of the spiral. This framework allows the customer to reduce the risk by periodically reviewing the requirements and evaluating a "completed", although not exhaustive, working system. It also allows to naturally take into account technological uncertainty by using intermediate steps to reduce the gap between the current and the final CCS, each phase feeding the next with a better understanding of system requirements.

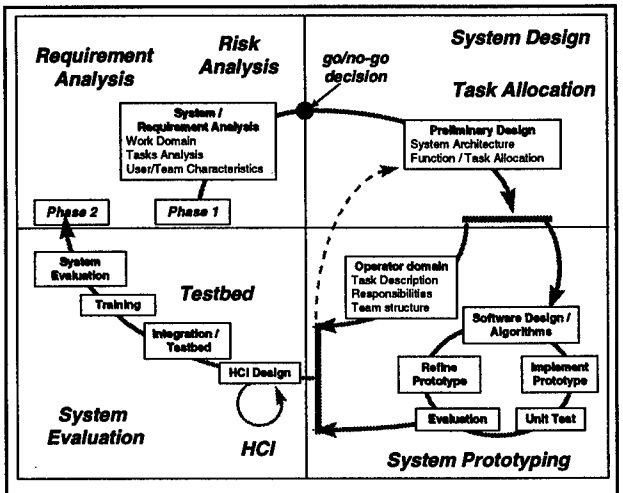


Figure 3: Refined Spiral Design Framework for a Computer-Based Decision-Support System

This *high-level* spiral development model describes the general activities to be performed and their sequence and feedback loops.

Each cycle of the spiral development starts with a requirement analysis, drawing from analyses of the work domain, system tasks and operator models, using analysis tools described further down. A risk analysis is then performed, followed by a "go/ no go" decision. It is assumed at this stage that *only a subset of the complete CBDSS requirements will be considered* at the initial cycle, with iterative additions made in subsequent phases. The same goes for the input analyses which will also increase in depth and breadth at successive iterations of the spiral process.

Based on the selected requirement system architecture will be defined, together with a rigorous function/task allocation activity. The results of this phase will feed a dual development phase: a first, so-called "cognitive" development team will investigate the structure and activities to be performed by the team of operators, while a second "technological" team will develop a prototype of the expected functionalities. Because of the expected level of uncertainty, this technological development will follow a *prototyping* approach, including implementation and unit testing of the components involved. This phase culminates in a quick validation of the initial task splitting and system design; in the case of a serious technological problem or task misallocation resulting in obvious performance degradation, it might be necessary to go back to the system design phase for a revision of task allocation (dotted line) before going to final system evaluation.

Finally, a HCI is developed from the previously identified tasks to be performed by the operator(s), following a prototyping approach. Testbed implementation follows, in order for a human factors team to evaluate the performance of the human/machine system. This evaluation results in a set of conclusions, which become system requirements for the next loop of the spiral development.

5. The Framework Application

The current CPF DSS Demonstration testbed architecture is excellently suited for application of the Refined Spiral Design framework for the CBDSS for Halifax Class, described above. Its modularity, independence of its components and flexibility in re-working/adding components will permit addressing the issues identified above. It will easily accommodate parallel "technological" and "cognitive" team investigations and any iterations they may require as a result of their analyses.

Based on the literature survey the tasks which should be included in the Halifax Class CBDSS development process include:

1. **System Analysis**, including:
   a) A **Cognitive Task Analysis**
   b) Skills-Rules-Knowledge (**SRK**) as a model of the decision-making process
   c) Work Domain functional analysis using the **abstraction hierarchy**
   d) A model of the **generic task** of the operator

2. **Functions and Tasks Allocation**, for which a few useable methodologies exist, but with no specific framework or methodology being particularly efficient or outstanding

3. **System Development and implementation,** using a 2-teamed Prototyping approach

4. **HCI,** developed using a prototyping approach and based on **EID**

5. **System Evaluation**, with 3 distinct areas of interest:
   a) **Function Usability** (e.g., ease of use), which is well understood and for which several methodologies exist
   b) **Operational Impact**, using pre-defined, numerical measures of performance
   c) **User Fit** (including Situation Awareness and mental workload), which is much less parametric and precise.

The testbed can be used to incrementally experiment with and evaluate various approached, with the aim to understand which of these methodologies can actually be implemented, whether they can be fully exploited, and to which level of detail they should be developed.

6. Conclusion

Based on a Literature survey on CBDSS for Command and Control this paper selected and described a complete Spiral approach to Human-Machine System Design that takes into account both human and technological aspects of system development, which we have presented and justified in the specific context of the design and implementation of a CBDSS for the HALIFAX Class ships.

A testbed architecture that can accommodate and facilitate such an approach was also described.

The discipline of CBDSS for future C2 in terms of both technological and cognitive aspects is quite young, and significant more effort should be applied in analyses and evaluations in testbed environments to ensure that the user trusts, accepts and uses DSS capabilities.

7. Acknowledgements

The authors would like to thank Dr. Bruce Chalmers of DREV for his support and participation in the CBDSS research al LM Canada over the last 5 years, and Dr. Jean-Marc Robert of École Polytechnique of Montreal for his significant contribution in the theoretical analyses of the results of the literature survey.

8. References

[1] Shahbazian E., Duquet J.-R., Macieszczak M., Valin P., A Generic Expert System Infrastructure for Fusion and Imaging Decision Aids, paper presented at EuroFusion'98 conference

[2] Bergeron P., Couture J., Duquet J.-R., Macieszczak M., and Mayrand M., A New Knowledge-Based System for the Study of Situation and Threat Assessment in the Context of Naval Warfare, regular paper presented at Fusion '98 conference, in session 13 "Naval Applications of Data and Information Fusion Techniques".

[3] Duquet J.-R., Bergeron P., Blodgett D.E., Couture J., Macieszczak M., andMayrand M., Analysis of the functional and real-time requirements of a Multi-Sensor Data Fusion (MSDF) / Situation and Threat Assessment (STA) / Resource Management (RM) system, in Sensor Fusion: Architectures, Algorithms, and Applications II, SPIE Aerosense '98, Orlando, 13-17 April 1998, Conf. 3376, pp. 198-209.

[4] LM Canada Report for Contract W8477-6-PH01/01-QE PWGSC FILE No. 061QE.W8477-6-PH01, Literature Survey on Computer-Based Decision Support for Command and Control Systems, in progress

[5] Taylor and Francis; Handbook of Control Room Design and Ergonomics; Ivergard, Chap 7. (1989).

[6] Matthews, M. L., Webb, R.D.G. and McCann, C.: A Framework For Evaluation of Military Command and Control Systems;__*Institution*: Defence and Civil Institute of Environmental Medicine; (1997).

Session RA4
Fusion for Fault Detection and Diagnosis
Chair: Mark Alford
Air Force Research Laboratory-Rome, NY, USA

Detection and Localization of Faults in System Dynamics by IMM Estimator

Ludmila Mihaylova [1] **Emil Semerdjiev**[1]

Bulgarian Academy of Sciences
Central Laboratory for Parallel Processing
'Acad. G. Bonchev' Str., Bl. 25-A, 1113 Sofia, Bulgaria
Phone: (359 2) 979 66 20; Fax: (359 2) 707 273
E-mail: lsm@bas.bg, signal@bas.bg

X. Rong Li *

University of New Orleans
Department of Electrical Engineering
New Orleans, LA 70148
Phone: 504-280-7416, Fax: 504-280-3950
E-mail: xli@uno.edu

Abstract—*The cost-effective Interacting Multiple-Model (IMM) algorithm is applied for rapid and reliable fault detection and localization in system dynamics. The paper also presents a new IMM approach for system parameter drift detection and estimation based on augmented state models.*

Key Words: multiple models, fault detection, IMM

1. Introduction

The multiple model (MM) approach is used for solving a very wide range of problems under uncertainty and abrupt changes in system identification [6,7,12,13], target tracking [1-2,4-6,8], control of industrial plants [12], etc. Among all existing MM state estimation algorithms, the IMM is one of the most popular and cost-effective [1,2,6]. A new and important application of the IMM algorithm is detection and diagnosis of failures in system dynamics. In the recent papers [3, 10, 14] it is shown that the IMM estimator is a more reliable fault detector in comparison with the fault detectors using a bank of "non-interacting" single-model-based filters running in parallel [9, 10]. In [14] detection and diagnosis of *sensor* and *actuator* failures by IMM is investigated, as well.

The purpose of the paper is to present how the IMM estimator can be applied when the faults cause *structural changes* in the system. A new IMM approach based on augmented state models is also proposed to recognize and overcome erroneous system behavior changes due to *parameter nonstationarity* known as a "drift." Test examples demonstrating the efficiency of the approach are given.

2. IMM Estimator for Fault Detection

Consider a system S composed of *independent* (*non-interacting*) subsystems $S_i, i=1,2,...,q$ connected in parallel (Fig. 1), with state vectors x_i, $i=1,2,...,q$. In the hard case these subsystems are *unobservable separately*. This system feature is taken into account in the measurement equation, where the output y_k is a sum of the outputs of all subsystems.

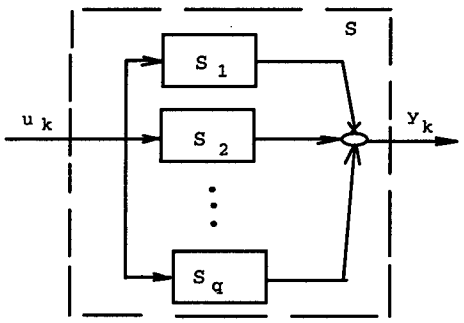


Fig. 1

The system behavior is described by the equations:

$$x_k = F_{k-1}(M_k)x_{k-1} + G_{u,k-1}(M_k)u_{k-1} + G_{v,k-1}(M_k)v_{k-1}(M_k), \quad (1)$$

$$z_k = H_k x_k + w_k, \quad (2)$$

[1] Partially supported by the Bulgarian National Science Fund - Contract No. I-808/98.
* Supported by ONR via Grant N00014-97-1-0570, NSF via Grant ECS-9734285, and LEQSF via Grant (1996-99)-RD-A-32.

where $x_k = \left(x_1^{'}, \cdots, x_q^{'}\right)^{'}$, $x_i \in R^{n_x}$ is the system state, $z \in R^{n_z}$ is the measurement vector; $u \in R^{n_u}$ is the control input vector; $v \in R^{n_v}$ and $w \in R^{n_z}$ are mutually uncorrelated, white zero mean noises with covariances Q_k and R_k.

The system mode M_k at time k is considered to be among l possible modes (models combinations), including the "zero" mode (all subsystems are active). The i-th combination of submodels in effect during the sampling period k of length T is denoted by $M_k = i$. The mode sequence is usually modeled as a Markov chain with known initial mode probabilities $\mu_i = P\{M_0 = i\}$ and transitional probabilities $p_{ij} = P\{M_k = j / M_{k-1} = i\}$, $i, j = 1,2,\cdots,l$. The IMM algorithm generates an overall state estimate as a weighted sum of state estimates $\hat{x}_i$, formed by a bank of Kalman filters, working in parallel [2, 6].

Consider the fault detection problem. *The main goal of IMM implementation is observing the system behavior to detect abrupt system structural changes caused by faults (or operators actions) or changes caused by a gradual system parameter drift.* In our study the main attention is paid on the *changes caused by faults in the system matrix F,* and/or *the matrices G_u and G_v, respectively,* as opposed to faults on the measurement matrix H.

Each *system structure change* is modeled as switching on/off of some subsystem(s). Here it is taken into account through an annihilation of appropriate column(s) in the matrix F, G_u and G_v. The remaining subsystems form a particular system substructure and a respective mode. A set of *mutually exclusive and exhaustive hypotheses*, describing all possible combinations of independent modes can be created in this manner. The probabilities of transitions between the different modes are usually known or preset according to the specific problem. The topical substructure is recognized as the one with the greatest mode probability.

Another important problem considered here is the *gradual system parameter drift.* An adaptive approach is proposed in the paper. It is supposed that a set of *initial* discrete values $m_{k,i}$ $(i = 1,2,\cdots,q)$ is known for each changing system parameter $m_k \in \Re^{n_M}$. The i-th system mode M_i is related to the hypothesis "*parameter m_i is changing while the others remain constant*". The "zero" system mode assumes "*no system parameter changes.*" For the i-th mode $(i = 1,2,\cdots,q)$ we introduce the augmented system vector $X_{k,i} = \left(x_{k,i}^{'} \quad \Delta m_{k,i}\right)^{'}$ for the system described by:

$$X_{k,i} = F_{k-1,i} X_{k-1,i} + G_{u,k-1,i} u_{k-1} + G_{v,k-1,i} v_{k-1,i} . \quad (3)$$

Model (3) allows us to detect and estimate the parameter deviation. The topical mode is recognized by the greatest IMM mode probability.

3. Simulation Results

A particular case of a linear tracking system consisting of three subsystems is considered in the examples below. Each subsystem is described by the discrete analogue of the transfer function $W(p) = 1/(T_i p + 1), i = 1,2,3$ obtained by sampling and a zero-order hold. The i-th time constant is denoted by T_i. The sampling period is 1 *s*. The *true* i-th discrete state space model has

$$F_i = e^{-1/T_i},\ G_{u,i} = G_{v,i} = 1 - e^{-1/T_i},\ H_i = 1.$$

The three *true* subsystems respectively have time constants $T_1 = 10\,s,\ T_2 = 2\,s,\ T_3 = 1\,s$ and

$$\begin{aligned} S_1&: F_1 = e^{-0.1},\ G_{u,1} = G_{v,1} = 1 - e^{-0.1},\ H_1 = 1, \\ S_2&: F_2 = e^{-0.5},\ G_{u,2} = G_{v,2} = 1 - e^{-0.5},\ H_2 = 1, \quad (4) \\ S_3&: F_3 = e^{-1},\ G_{u,3} = G_{v,3} = 1 - e^{-1},\ H_3 = 1. \end{aligned}$$

It is also denoted below $G = G_u = G_v$. The input control process u_k is a zero-mean stochastic one.

All presented results are based on 100 Monte Carlo runs.

Example 1.

The system is characterized by unknown changeable structure

$$S = \begin{cases} [S_1,\ S_2,\ S_3], & for\ \ 0 \le k < 100 \\ [S_1,\ S_2], & for\ \ 100 \le k < 300 \\ [S_1,\ S_2,\ S_3], & for\ \ 300 \le k < 400 \\ [S_2,\ S_3], & for\ \ 400 \le k \le 500 \end{cases}.$$

The matrices of the IMM models are:

$$F_1 = diag\left\{e^{-0.1} \quad e^{-0.5} \quad e^{-1}\right\}, G_1 = \left(1 - e^{-0.1} \quad 1 - e^{-0.5} \quad 1 - e^{-1}\right)^{'}$$

$$F_2 = diag\left\{e^{-0.1} \quad e^{-0.5} \quad 0\right\}, G_2 = \left(1 - e^{-0.1} \quad 1 - e^{-0.5} \quad 0\right)^{'}$$

$$F_3 = diag\left\{0 \quad e^{-0.5} \quad e^{-1}\right\}, G_3 = \left(0 \quad 1 - e^{-0.5} \quad e^{-1}\right)^{'}$$

$$F_4 = diag\left\{e^{-0.1} \quad 0 \quad e^{-1}\right\}, G_4 = \left(1 - e^{-0.1} \quad 0 \quad e^{-1}\right)^{'}$$

$H_i = (1\ 1\ 1), Q_i = 0.01^2, R_i = 0.2^2,\ i = \overline{1,4}$ and u_k is a white random process with a variance equal to 1. $diag\{.\}$ denotes a diagonal matrix.

The mode transition probability matrix Pr and the initial mode probability vector $\mu(0)$ are:

$$Pr(i,i)=0.97,\ Pr(i,j)=0.01;$$

$$\mu(0)=(0.94 \quad 0.02 \quad 0.02 \quad 0.02)'.$$

The computed average mode probabilities of the IMM fault detector are given in Fig. 2.

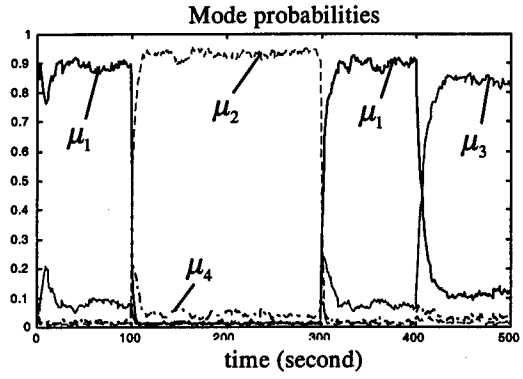


Fig. 2

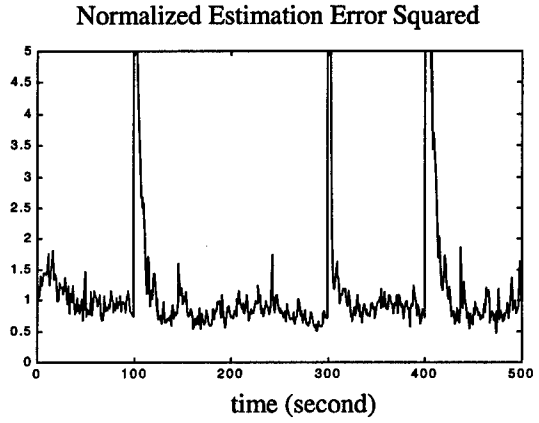


Fig. 3

It is obvious that the greatest IMM mode probability μ_{max} always provides the right decision:

- $1 \le k < 100$, all subsystems are active ($\mu_{max}=\mu_1$);
- $100 \le k < 300$, S_1 and S_2 are active ($\mu_{max}=\mu_2$);
- $300 \le k < 400$, all subsystems are active ($\mu_{max}=\mu_1$);
- $400 \le k \le 500$, S_2 and S_3 are active ($\mu_{max}=\mu_3$).

The Normalized Estimation Error Squared (NEES) [2] characterizes the filter consistency. It is shown in Fig. 3.

Example 2.

In the second test the true system dynamics is changed as follows:

$$S=\begin{cases}[S_2], & for \quad 0 \le k < 200\\ [S_2,S_3], & for \quad 200 \le k < 400\\ [S_1,S_2,S_3], & for \quad 400 \le k \le 500\end{cases}.$$

Five IMM models are used:

1) $F_1=diag\{0 \quad e^{-0.5} \quad 0\}, G_1=(0 \quad 1-e^{-0.5} \quad 0)'$;

2) Models *i=2,3,4* remain as in *Example 1*;

3) The *5*-th model coincides with the first model in *Example 1*.

The mode transition probability matrix and the initial mode probabilities for the IMM are:

$$Pr(i,i)=0.92,\ Pr(i,j)=0.02;$$

$$\mu(0)=(0.9 \quad 0.025 \quad 0.025 \quad 0.025 \quad 0.025)'.$$

The average mode probabilities are shown in Fig. 4 and the NEES is given in Fig. 5.

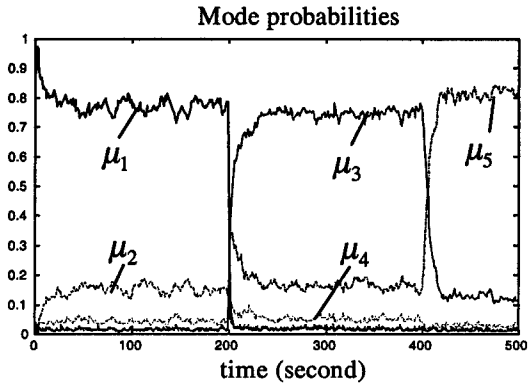


Fig. 4 Normalized Estimation Error Squared

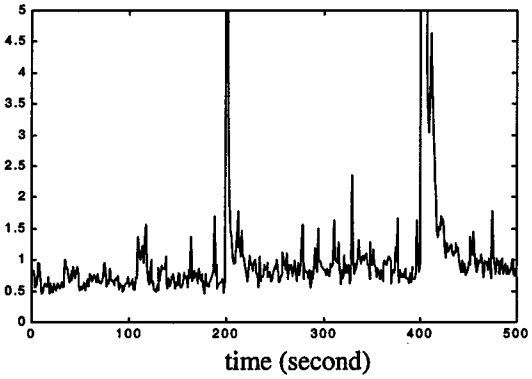


Fig. 5

Example 3.

This example illustrates the efficiency of the proposed approach in parameter drift detection and estimation. It is presupposed that the drift ΔT_i can occur in only one true subsystem transforming it into a nonlinear one. The problem is to detect where and when the drift appears, if there is any, and to estimate its direction and magnitude.

The different system modes are determined based on the time constants T_i, in which the drift can arise. Following the above idea, the system state vector has to be augmented by $\Delta T_i \geq 0$ (for convenience here it is chosen $\Delta m_i = -1/\Delta T_i$). An additional '*no drift*' hypotheses is introduced (*i= 0*), so the IMM is working by four EKFs running in parallel.

The *i*-th EKF equations have the form:

$$\hat{x}_{i,k/k} = \hat{x}_{i,k/k-1} + K_{i,k}\gamma_{i,k},$$

$$\hat{x}_{i,k/k-1} = f\left(\hat{x}_{i,k-1/k-1}, \Delta m_i\right),$$

$$\gamma_{i,k} = y_k - H_i\hat{x}_{i,k/k-1},$$

$$P_{i,k/k-1} = f_i^x P_{i,k-1/k-1}\left(f_i^x\right)' + G_k Q_{i,k} G_k',$$

$$S_{i,k} = H_i P_{i,k/k-1} H_i' + R,$$

$$K_{i,k} = P_{i,k/k-1} H_i' S_i^{-1},$$

$$P_{i,k/k} = P_{i,k/k-1} - K_{i,k} S_i K_{i,k}'.$$

where $\hat{x}_{i,k/k}$ and $\hat{x}_{i,k/k-1}$ are the filtered and predicted estimates of x_k; γ_i is the innovation process and S_i - its covariance matrix; K_i is the filter gain; P_i is the error covariance matrix, and $f_i^{\hat{x}} = \partial f_i/\partial\hat{x}_i$, $i = \overline{0,3}$.

The IMM is based on the following four hypotheses h_i:

- h_0: *there is no drift*:

$$F_0 = diag\left\{e^{-0.1} \quad e^{-0.5} \quad e^{-1} \quad 1\right\},$$

$$G_0 = diag\left\{1-e^{-0.1} \quad 1-e^{-0.5} \quad 1-e^{-1} \quad 0\right\},$$

$$Q_0 = diag\left\{0.01^2 \quad 0.01^2 \quad 0.01^2 \quad 0.6\right\};$$

- h_1: *a drift appears in subsystem* S_1:

$$F_{1,k} = diag\left\{e^{-0.1+\hat{x}_k(4)} \quad e^{-0.5} \quad e^{-1} \quad 1\right\},$$

$$G_{1,k} = diag\left\{1-e^{-0.1+\hat{x}_k(4)} \quad 1-e^{-0.5} \quad 1-e^{-1} \quad 0\right\}$$

$$, Q_1 = diag\left\{0.05 \quad 0.01^2 \quad 0.01^2 \quad 0.6\right\};$$

- h_2: *a drift appears in subsystem* S_2:

$$F_{2,k} = diag\left\{e^{-0.1} \quad e^{-0.5+\hat{x}_k(4)} \quad e^{-1} \quad 1\right\},$$

$$G_{2,k} = diag\left\{1-e^{-0.1} \quad 1-e^{-0.5+\hat{x}_k(4)} \quad 1-e^{-1} \quad 0\right\}$$

$$Q_2 = diag\left\{0.01^2 \quad 0.05 \quad 0.01^2 \quad 0.6\right\}.$$

- h_3: *a drift appears in subsystem* S_3:

$$F_{3,k} = diag\left\{e^{-0.1} \quad e^{-0.5} \quad e^{-1+\hat{x}_k(4)} \quad 1\right\},$$

$$G_{3,k} = diag\left\{1-e^{-0.1} \quad 1-e^{-0.5} \quad 1-e^{-1+\hat{x}_k(4)} \quad 0\right\}$$

$$Q_3 = diag\left\{0.01^2 \quad 0.01^2 \quad 0.05 \quad 0.6\right\}.$$

The measurement matrix is predefined as:

$$H_i = \begin{pmatrix}1 & 1 & 1 & 0\end{pmatrix}, \; i = \overline{0,3}.$$

The respective Jacobians $f_i^{\hat{x}}$ are:

- $f_0^{\hat{x}} = F_0$;
- $f_{1,k}^{\hat{x}} = F_{1,k}$, $f_{1,k}^{\hat{x}}(1,4) = \hat{x}_k(1)e^{-0.1+\hat{x}_k(4)}$;
- $f_{2,k}^{\hat{x}} = F_{2,k}$, $f_{2,k}^{\hat{x}}(2,4) = \hat{x}_k(2)e^{-0.5+\hat{x}_k(4)}$;
- $f_{3,k}^{\hat{x}} = F_{3,k}$, $f_{3,k}^{\hat{x}}(3,4) = \hat{x}_k(3)e^{-1+\hat{x}_k(4)}$.

The unknown true drift is modeled as a slow, moderate and fast (relatively to the time constants) change as follows:

$$x_k(4) = x_{k-1}(4) - 0.0075. \qquad (5)$$

At the beginning all subsystems are working without drift. Since it is assumed that the drift can be in only one subsystem, the transition between drifting subsystems is not considered; that is, $Pr(1,1) = Pr(2,2) = Pr(3,3) = 1$. So,

$$Pr = \begin{pmatrix}0.94 & 0.02 & 0.02 & 0.02\\ 0.00 & 1.00 & 0.00 & 0.00\\ 0.00 & 0.00 & 1.00 & 0.00\\ 0.00 & 0.00 & 0.00 & 1.00\end{pmatrix}, \; \mu(0) = \begin{pmatrix}0.97\\ 0.01\\ 0.01\\ 0.01\end{pmatrix}$$

To provide better overall performance the system covariance matrix Q is introduced. The description of the system input is modified for this purpose. It is a vector $\begin{pmatrix}u+v & u+v & u+v & 0\end{pmatrix}'$. The element $Q(4,4)$ is chosen to be much bigger than the other elements to provide a fast response to the parameter drift. A large value is also assigned to the element $Q(i,i)$, when the filter corresponding to the *i*-th hypothesis is running.

The estimated change can not be positive. For this reason an additional hard logic is used:

$\hat{x}(4) = 0$, when $\hat{x}(4) > 0$.

- *Test No. 1:* There is no drift. The average mode probabilities, the real value of $x(4)$ and its estimate $\hat{x}(4)$, as well as the NEES are given in Figs. 6 - 8.
- *Test No. 2:* A gradual change (5) has occurred in subsystem S_1. The average mode probabilities, the real value of $x(4)$ and its estimate $\hat{x}(4)$, as well as the NEES are shown in Figs. 9 - 11.

- *Test No. 3:* A gradual change (5) has occurred in the subsystem S_2. The respective plots are given in Figs. 12 - 14.
- *Test No. 4:* A gradual change (5) has occurred in S_3. The respective plots are given in Figs. 15 - 17.

As can be seen from the plots, the algorithm resolves the "competition" between its models after a period of drift accumulation.

It can be observed from the NEES plots that the consistency of the drift parameter estimate deteriorates when the drift brings the subsystems' parameters near one to another and a duplication of the subsystems appears. The drift in T_1 is the worst case from this point of view.

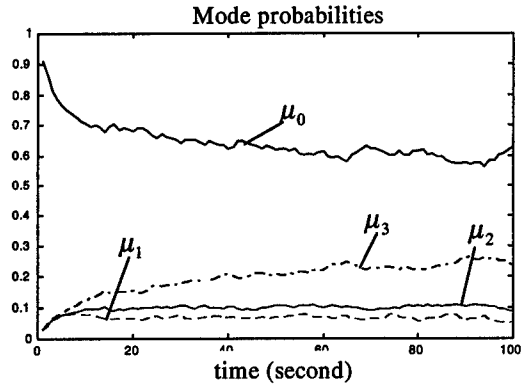


Fig. 6

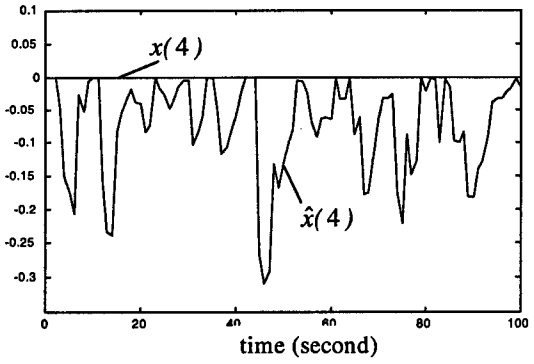


Fig. 7

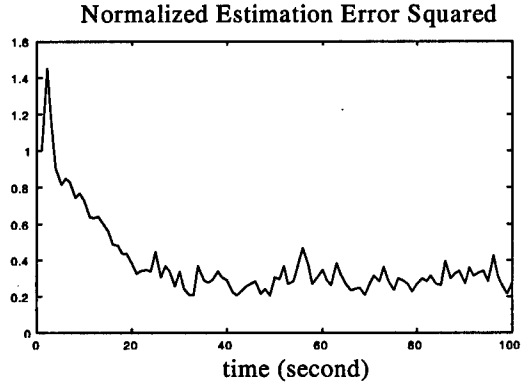


Fig. 8

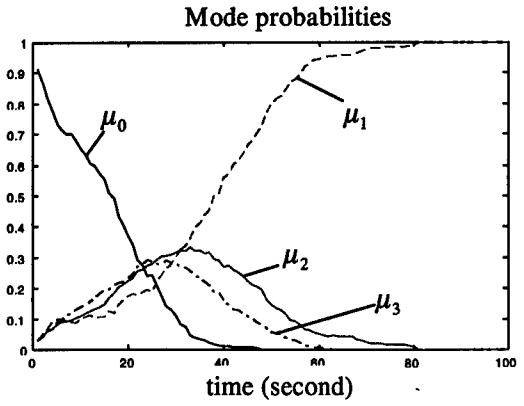


Fig. 9

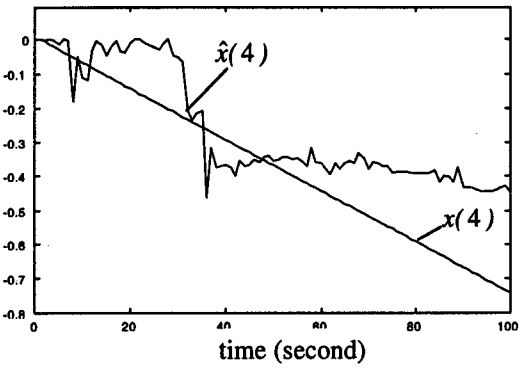


Fig. 10

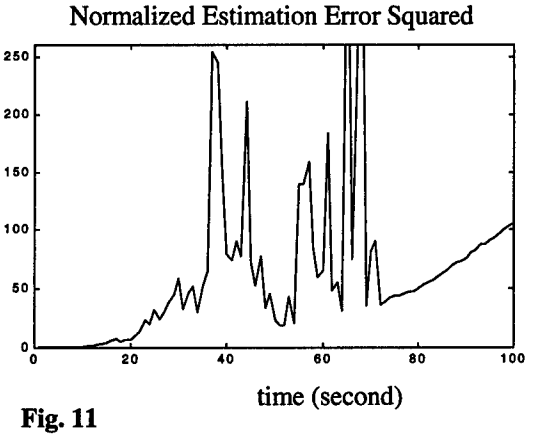


Fig. 11

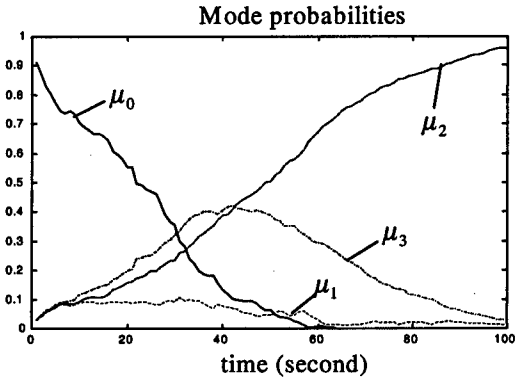


Fig. 12

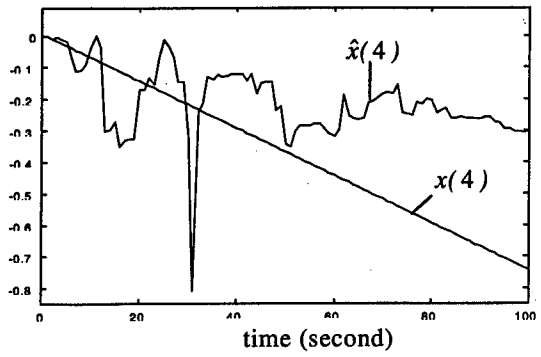


Fig. 13

Normalized Estimation Error Squared

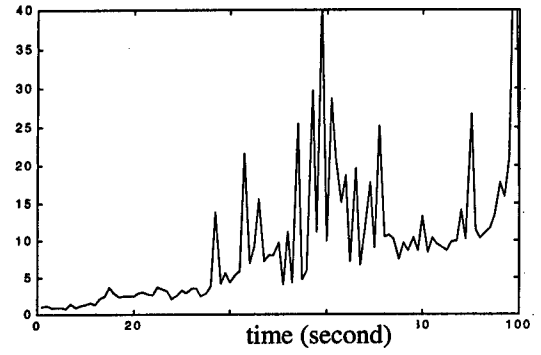


Fig. 14

Mode probabilities

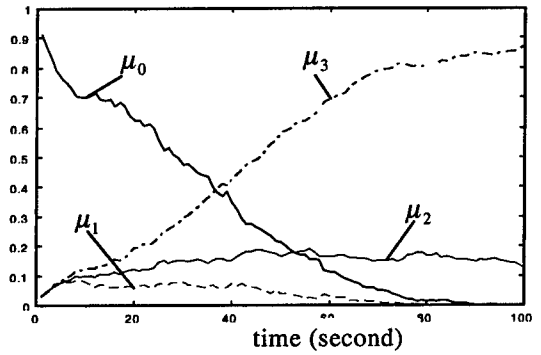


Fig. 15

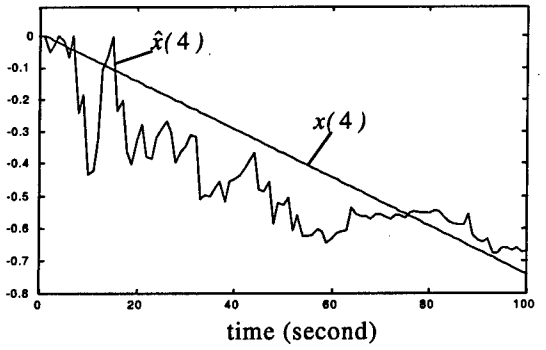


Fig. 16

Normalized Estimation Error Squared

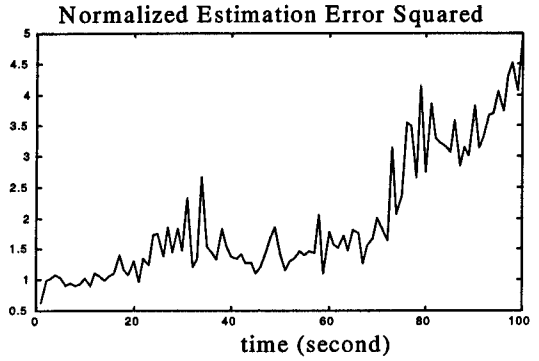


Fig. 17

Finally, it is noted that the effect of divergence in the Kalman filters due to numerical errors did not appeared in all the test examples, in spite of the fact that some of the matrices are sparse. This differs from the situations reported in [14].

4. Conclusions

The paper investigates the application of the IMM algorithm in fault detection and localization in the presence of abrupt changes in the system structure as well as in the presence of gradual parameter drift.

The dynamics changes are appropriately reflected in the system model and in the IMM transition probability matrix. Decisions are taken on the basis of the mode probabilities. It is shown that the IMM estimator can be used to detect system structure changes in general.

A new adaptive approach based on IMM estimator is proposed also to detect and estimate system parameter drifts. This gradual parameter change is represented as an additional component of the state vector. It is appropriately reflected in the system description. The algorithm's efficiency is confirmed by Monte Carlo simulations.

References

[1] Bar-Shalom, Y., Ed. *Multitarget-Multisensor Tracking: Applications and Advances*, vol. 2, Artech House, 1992.

[2] Bar-Shalom, Y., X. R. Li. *Multitarget-Multisensor Tracking: Principles and Techniques*, YBS Publishing, 1995.

[3] Efe, M., and D. P. Atherton. The IMM Approach to the Fault Detection Problem, in *Proc. of the 11th IFAC Symposium on System Identification*, Fukuoka, Japan, July 1997.

[4] Jilkov, V., P., Angelova and Tz. Semerdjiev. Design and Comparison of Mode-Set Adaptive IMM for Maneuvering Target Tracking, *IEEE*

Trans. Aerosp. and Electr. Systems, 35(1): 343-350, Jan. 1999.

[5] Jilkov, V. P., L.Mihaylova, X. R. Li. An Alternative IMM Solution to Benchmark Radar Tracking Problem, in Proc. 1998 Intern. Conf. on Information Fusion (Fusion'98), vol. 2, 924-929, Las Vegas, Nevada, USA, July 6-9, 1998.

[6] Li, X. R. *Hybrid Estimation Techniques.* In *Control and Dynamic Systems*. (Ed. C. T. Leondes), vol. 76, 213-287, Academic Press, 1996.

[7] Madjarov, N., Mihaylova, L. A Multiple Model Stochastic Estimator in the Presence of Structural Uncertainty in the System, in *Proc. of the National Conf. Automatics and Informatics '98,* 20-23 October 1998, Sofia, Bulgaria, 2: 12-15, 1998.

[8] Mazor, E., A. Averbuch, Y. Bar-Shalom, and J. Dayan. IMM Methods in Target Tracking: A Survey, *IEEE Trans. Aerosp. and Electr. Systems*, 34(1): 103-123, 1998.

[9] Maybeck, P. and G.Griffin, Jr. MMAE/MMAC Control for Bending with Multiple Uncertain Parameters, *IEEE Trans. Aerosp. and Electr. Systems*, 33(3): 903-912, 1997.

[10] Mehra, R., C. Rago, and S. Seereeram. Failure Detection and Identification Using a Nonlinear Interactive Multiple Model (IMM) Filtering Approach with Aerospace Applications, In *Proc. of the 11th IFAC Symposium on System Identification,* Fukuoka, Japan, July, 1997.

[11] Menke, T., and P. Maybeck. Sensor/Actuator Failure Detection in the Vista F-16 by Multiple Model Adaptive Estimation *IEEE Trans. Aerosp. and Electr. Systems,*31(4):1218-1229, 1995.

[12] Mihaylova, L., V. Jilkov, Tz. Semerdjiev. A Combination Algorithm for Noise Identification, In *Proc. of the ECCTD' 97*, Budapest, Hungary, 2, 848-852, 1997.

[13] Murray-Smith, R., T. Johansen (Eds.). *Multiple Model Approaches to Modelling and Control,* Taylor and Francis, 1997.

[14] Zhang, Y., X. R. Li. Detection and Diagnosis of Sensor and Actuator Failures Using IMM Estimator, *IEEE Trans. Aerosp. and Electr. Systems*, 34(4): 1293-1311, 1998.

Diagnostic Information Processing for Sensor-Rich Distributed Systems

Elmer S. Hung and Feng Zhao
Xerox Palo Alto Research Center
3333 Coyote Hill Rd.
Palo Alto, CA 94304
{eshung,zhao}@parc.xerox.com

Abstract

This paper describes a diagnostic system for processing high-bandwidth vibration data from distributed sensors for monitoring and diagnosis of electromechanical machines. The system employs time-frequency and principal component analysis techniques to extract and compress features and a Bayesian decision analysis to combine and classify data from multiple sources. Experimental multi-sensor diagnosis results are reported for classifying motor and solenoid vibration signatures from the paper drive plate of the Xerox DC265 digital copier.

Keywords: diagnostics, sensor fusion, time-frequency analysis, wavelets, STFT, Bayesian decision analysis

1 Introduction

Recent advances in batch-fabricated micromachined sensors and electronics have the potential to enable a new generation of condition-based monitoring and diagnostic systems for complex machinery. However, taking advantage of increased sensor and processing capabilities demands corresponding advances in computational techniques for analyzing massive amounts of data from distributed sensor systems.

The work in this paper is part of a larger effort at Xerox to identify and develop scalable processing architectures for interpreting data from many sensors that may be scattered or embedded inside a system. Specifically, in this paper, we present work on combining information from vibration sensors for in-situ diagnosis of component health on a paper drive plate from a Xerox Document Centre 265 (DC265) digital copier. The focus is on sensor data analysis, including the development of flexible time-frequency diagnostic filtering techniques for sensor-rich environments.

Although other researchers (e.g. [1]) have investigated time-frequency techniques for diagnostic vibration analysis, the emphasis in this paper is on the use of time-frequency filters to combine information from large distributed networks of sensors for machine diagnostics. We present a framework for analyzing time-frequency data from many sensors, where it is assumed that training data from lifetime tests is available, but it is too costly to explicitly model the dynamics of the physical environment around every sensor.

The diagnostic data processing system we investigate involves four major components: (1) signal feature extraction, (2) data clustering and compression of high-dimensional feature space information, (3) data aggregation of signal features from multiple sensors, and (4) signal classification and decision analysis. (See the block diagram in Fig. 1.) Data from lifetime tests of the system are used to train the diagnostics algorithm about normal and abnormal operating characteristics. The training is performed offline to generate the run-time diagnostic algorithm.

In the following sections, we first describe the experimental setup for the diagnostics testbed (Section 2). Section 3 describes time-frequency feature extraction techniques, the short-time Fourier transform (STFT) and wavelet analysis. Section 4 describes the use of principal component analysis (PCA) and other techniques to compress the resulting high-dimensional feature space onto a lower dimensional subspace based on the training data from lifetime tests.

The feature space data from different sensors or data analysis methods is aggregated using a statistical Bayesian analysis of probability density functions extracted from the training data (Section 5). A Bayesian discriminant function (Section 6) is then used to produce a simple "health index" that

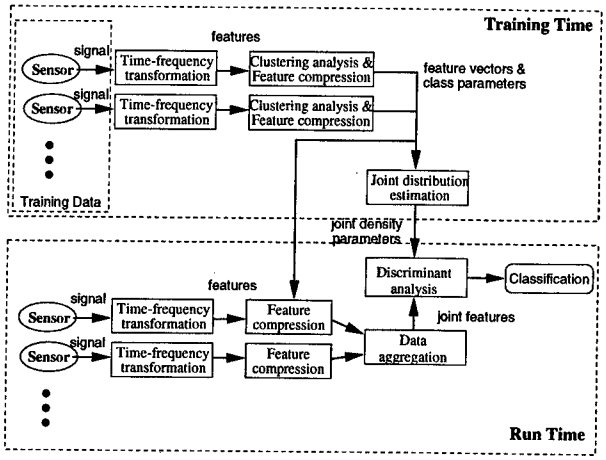


Figure 1: Block diagram of the diagnostic system.

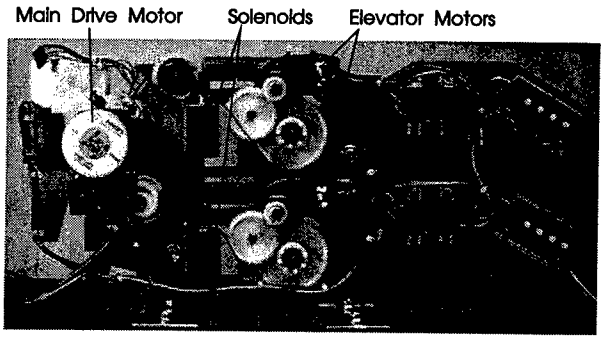


Figure 2: Diagnostic testbed: paper drive plate subsystem of Xerox DC265 copier.

measures how far the system is away from normal operation, or, more generally, an index that measures how close the current behavior is to identified modes of operation or fault conditions. These indices can then be used to generate decision classifications. Finally, we report on the results of applying this algorithm to experimental data (Section 7).

2 Experimental Setup

The drive plate subsystem of a Xerox DC265 copier (see Fig. 2) was chosen as a diagnostics testbed because it is a relatively complex electromechanical system involving different types of actuators with multiple modes of operation. The drive plate is responsible for acquiring paper from the paper tray and directing it to the paper path for xerographic processing. It contains a number of actuators including a main motor, two solenoids, and two elevator motors.

The testbed is instrumented with vibration sensors (PCB Piezotronics, model J352C65). Vibration sensors are used in the initial tests because they provide generic information that is useful for diagnosing different types of actuators, and because they require high data throughputs that may be important in future distributed sensor applications. The vibration sensors on the drive plate are sampled by an analog-to-digital converter (ADC) at 50 kHz with 12-bit resolution. The data is oversampled to provide better accuracy in the optimal bandwidth of the sensor (10 Hz-8 kHz).

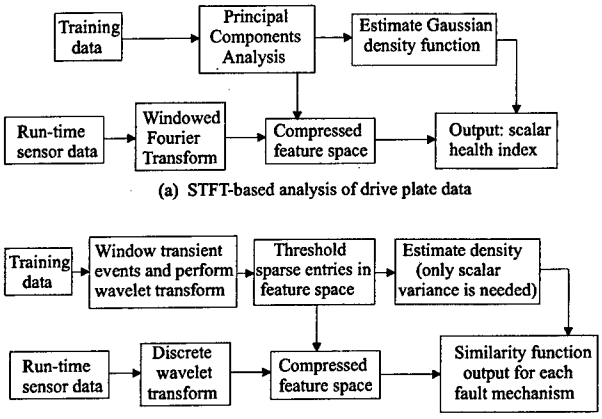
In order to test the diagnostic processing techniques, actuator behavior is purposely compromised in some experiments. In one experiment a washer is attached to the main motor to simulate unbalanced behavior. In another experiment a rubber plug is used to limit the plunger travel distance in one of the solenoids. The objective of these experiments is to see if it is possible to distinguish normal from compromised actuator behavior and to identify actuator operating modes by analyzing the vibration signatures from multiple sensors.

3 Time-Frequency Analysis

Two time-frequency-based techniques are used to analyze the vibration data: windowed short-time Fourier transforms (STFT) and wavelet analysis. For the diagnostic algorithms we have implemented, two training steps are performed offline based on data from lifetime tests in order to establish the normal and/or abnormal operating characteristics of the device. Sections 3 and 4 describe the first training step which involves the use of time-frequency analysis to generate a feature space that properly captures diagnostic information. The second training step (Section 5) is a Bayesian analysis of the data, which involves approximating a Gaussian density function in the product space of the feature spaces for all the sensors involved.

For the first training step, the STFT or wavelet analysis may be used directly to create the feature space, or, more likely, the training data may be used to compress the data onto a lower-dimensional feature space. For example, we use principal component analysis (PCA) to reduce feature space dimensionality.

For reference in the following discussion we present block diagrams of the two specific implementations we used in analyzing DC265 paper drive plate data (Fig. 3). One should note, however, that modules in the two approaches could be interchanged if desired. For example, one could use PCA on the wavelet-based feature space as illus-



(a) STFT-based analysis of drive plate data

(b) Wavelet-based analysis of drive plate data

Figure 3: (a) Block diagram showing STFT-based diagnosis algorithm implemented for DC265 drive plate vibration data. (b) Block diagram showing wavelet-based diagnosis algorithm implemented.

trated with the STFT analysis.

3.1 STFT

STFTs have been used previously in a wide variety of applications including speech, radar, and image processing [2]. For our diagnostic analysis, the strategy can be summed up as follows: During run time, the STFT method involves looking at a sliding window of vibration data in time, computing the Fourier transform of the data in the window, and using the transformed signal to classify the behavior based on its similarity to the training data. The question is how to make the most effective comparison between the incoming data and the training data given data from many sensors.

Note that spectral information from a discrete Fourier transform can be represented in terms of a high dimensional feature space: Suppose that each spectrum has N data points $\{(\omega_1, \alpha_1), (\omega_2, \alpha_2), \ldots, (\omega_N, \alpha_N)\}$, where the ω's are frequency samples and α's are magnitudes of the spectra. Each spectrum can be represented as a single point or vector $(\alpha_1, \alpha_2, \ldots, \alpha_N)$ in an N-dimensional feature space, S.

The first training step involves the collection of spectral data from numerous windows in time. Using principal component analysis, the sample spectra from the training data are used to collapse the high dimensional feature space S onto a lower dimensional subspace that captures the most important diagnostics information.

For example, typical spectra resulting from the main motor vibration signature for a window of

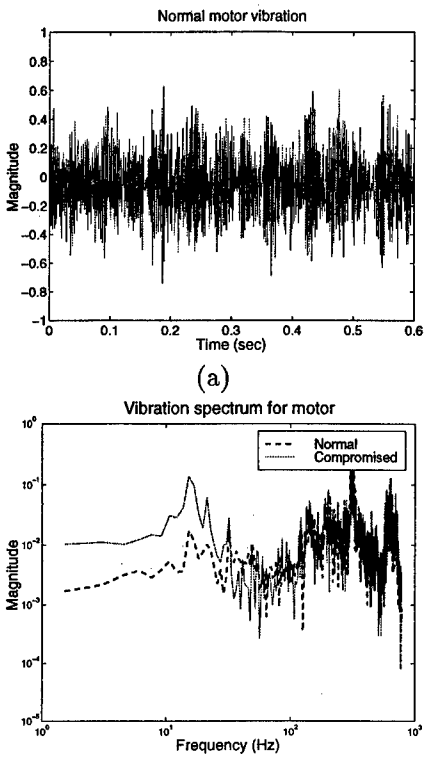


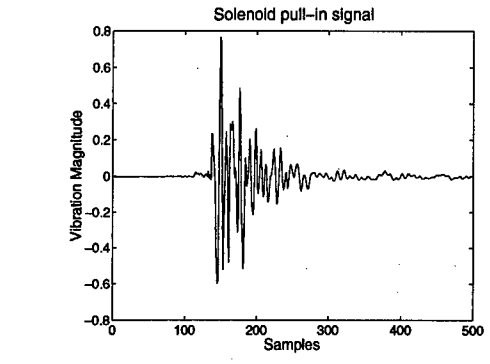
Figure 4: (a) Motor vibration time series (b) Sample frequency spectra for normal vs. compromised motor behavior from a time window of 655 ms.

length 655 ms is given in Fig. 4 along with the original time-series vibration data. The window length used should depend on the frequency bandwidth of interest. Fig. 4 (b) illustrates a detectable difference in the spectra for normal and compromised motor behavior. Note that for this graph and other vibration analysis in this paper, incoming signals are first decimated and low-pass filtered to reduce high frequency noise and aliasing effects from the sensors.

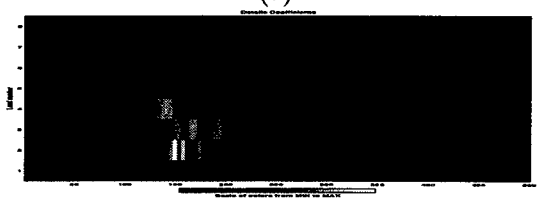
3.2 Wavelet analysis

A wavelet transformation of a signal produces a time and scale (which correlates to frequency) dependent expansion of the signal [3]. It is particularly useful for the analysis of non-stationary or transitory signals that do not have persistent statistical moments. Specifically, the wavelet analysis employs a family of wavelets, the so-called orthonormal basis functions $\{\psi_{u_n,s_j}\}_{(n,j)\in Z^2}$, where u_n and s_j are position and scale parameters and

$$\psi_{u_n,s_j}(x) = |s_j|^{\frac{1}{2}} \psi(\frac{x - u_n}{s_j}).$$



(a)



(b)

Figure 5: Wavelet analysis of transitory signals: (a) Solenoid pull-in signal; (b) Wavelet coefficients plotted as a two-dimensional intensity map: horizontal axis – time; vertical axis – scale. The coefficients are computed by a discrete wavelet transformation (DWT) with the Daubechies wavelet (db4-8).

Wavelet coefficients are obtained by convolving a wavelet with the signal $f(t)$:

$$a_{u,s} = \int_{-\infty}^{+\infty} f(t)\psi^*_{u,s}(t)dt.$$

Wavelet analysis has been applied by other researchers for fault detection; however most existing approaches aim at novelty or event detection against a steady-state background [4, 5], or are tuned to exploit system-specific features [6]. The wavelet-based discriminant analysis presented in this paper exploits not only event detection but also uses specific wavelet patterns to classify signals into different operating conditions.

For example, in Fig. 5 we take a time window of the vibration signal collected from the pull-in actuation of a solenoid on the testbed. The result is illustrated as an intensity map in the two-dimensional time-scale plane, the horizontal axis is the time, and vertical axis is the scale of the signal. Each tile records the amplitude of the wavelet expansion coefficient for the signal at scale level 2^j and position $2^j n$ in time.

In Fig. 5(b), the onset of the pull-in signal can be identified in time by a set of high-intensity tiles at levels 1-4. As features of the signal vary, the wavelet coefficients vary accordingly. Thus, the intensity variation in the distribution map of the wavelet coefficients could be used to pin-point when a change of interest occurs and what the change is in monitoring and diagnosis.

During the training phase, multiple windows $\{s_1, s_2, ..., s_n\}$ of transitory events (in our case solenoid firings) are extracted to form a training population for a given operating mode or condition. A feature vector $\vec{a_i} \in S$ for window s_i is formed by concatenating the coefficients of the wavelet transformation of s_i at all levels of interest. In practice, the number of levels to consider is a function of the signal energy distribution across the levels, also to be determined from the training samples.

4 Feature space compression

The feature space S produced by the STFT or wavelet transform is often high dimensional. This section describes techniques to compress this feature space into a more manageable lower-dimensional space. This is important in order to alleviate problems from overfitting. It also helps reduce communications and computational bandwidth requirements which is potentially critical for large sensor networks where data from many sensors must be aggregated.

4.1 PCA

For the STFT algorithm, principal component analysis (PCA) was used to project S onto a lower $k-$dimensional subspace, S_k. In the past, PCA has been used in a wide variety of settings including as a feature-space compression technique for a Bayesian classifier [7], and to produce reduced data characterizations of STFT's [8]. It seems particularly well-suited to the sensor fusion problems addressed in this paper because of the strong tendency for dimensionality explosion with increasing numbers of sensors.

The principal component analysis can be interpreted as follows: Let $\{x_1, x_2, \ldots, x_n\}$ be vectors in S, each representing a single windowed spectrum from the training data, and let $m = dim(S)$. For any $k < m$ the idea is to project the feature space onto a k-dimensional subspace, S_k, such that S_k minimizes the sum of the squares of the distances from each of the x_i's to S_k.

Computationally, this is accomplished as follows: Let X be the matrix whose columns are $\{x_1, x_2, \ldots, x_n\}$. For each $k < m$, S_k is the

space spanned by the principal component vectors, $\{u_1, u_2, \ldots, u_k\}$, which are given by the first k columns of the matrix U in the singular value decomposition (SVD) of X :

$$X = U\Sigma V^T$$

where U and V are orthogonal, and $\Sigma = diag(\sigma_1, \sigma_2, \ldots, \sigma_m)$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \geq 0$, and $p = min(m, n)$.

For $k << N$ the resulting spectral data representations on S_k contain far less data, but should capture the diagnostically useful information. Each x_i is now written in the new k-dimensional coordinate system based on the principal component basis vectors $\{u_1, u_2, \ldots, u_k\}$. For diagnostics purposes, we also keep one additional coordinate for each x_i containing the residual distance between x_i and S_k. This way the resulting feature space not only contains information that closely reconstructs the original feature space, but also provides information about how far away the original data was to the new representation.

4.2 Sparsity of feature space

Alternatively, the sparsity of the feature space can be exploited to compress the dimensions. Feature vectors computed from wavelet transform are typically sparse, with a large number of small coefficients which may be eliminated without losing diagnostically significant information. In other words, a signal can be adequately approximated using only a subset of feature dimensions that are tuned to record larger components of the signal. This technique was used to obtain a factor of two compression ratio for the wavelet analysis of the solenoid pull-in test case.

5 Bayesian Aggregation

Once a suitable feature space representation of the data is obtained, the question is how to combine information from multiple sensors and analysis algorithms, and how to make decisions regarding feature space output.

The first step involves combining features from multiple signals by considering a composite feature space consisting of the product of all the features spaces. In other words, suppose that the feature vectors from m sensors or data sources are given by $\{x_1, x_2, \ldots, x_m\}$. We then consider the composite feature vector $x = \prod_i^m x_i$.

Bayesian decision theory (e.g., [9]) can then be used to aggregate the data in the composite space. The probability that a hypothesis H_i is true given evidence x is given by:

$$P(H_i|x) = \frac{p(x|H_i)P(H_i)}{p(x)}$$

Omitting the normalizing constant $p(x)$ (independent of H_i), the heart of the analysis is to estimate the conditional probability density function $p(x|H_i)$ for each of the hypotheses, H_i.

Note that if data is not available from all the various fault conditions, it is also possible to use this framework to make decisions based on how far from normal the observed behavior is. In this case, the objective would be to simply estimate $p(x|H)$ from the data where H assumes a normal operating condition.

So how does one determine $p(x|H_i)$ from the training data information? In many interesting cases, it is feasible to assume that the sample data distribution $p(x|H_i)$ is close to Gaussian normal. In this situation, we may approximate the density function by simply estimating the mean and covariance of each sample cluster corresponding to a given hypothesis.

There are a few issues to point out. First of all, it may not be possible to approximate the density as Gaussian. In this situation, a number of techniques are available for approximating the density function, such as using a mixture of Gaussians [11], although, of course, the situation becomes more complex if the density function is not easily represented.

Second, if the dimensionality of the original feature space vectors has not been compressed, the dimensionality of the product space of feature vectors may be very large. There may not be enough data to generate a full representation of the density function. This is the case for the wavelet-based analysis (see the discussion in Section 6).

Note that if there are large numbers of sensors, this may also contribute to feature space dimension explosion. In this situation, it may be possible to hierarchically combine data from groups of sensors at a time. This increases the efficiency of the approach and reduces demand on data, but the drawback is that information about cross-correlations between sensor readings may be lost.

6 Discriminant Analysis

In this section, we assume that the density function $p(x|H_i)$ can be represented as a Gaussian. For sim-

plicity, it is then helpful to employ a discriminant function for classifying inputs [10].

We consider the discriminant function $g_i(x)$, a monotonic function of the *a posteriori* probability, $\log P(H_i|x)$:

$$g_i(x) = \log p(x|H_i) + \log P(H_i).$$

Since it is often difficult to estimate the prior $P(H_i)$, for our experiments we drop the $\log P(H_i)$ term and use the $g_i(x) = \log p(x|H_i)$. This gives the maximum likelihood classifier that selects H_i as the winning hypothesis if $\sup(g_i(x)) > \sup(g_j(x))$ for all $j \neq i$, where x may vary over the entire feature space.

For Gaussian density functions we have:

$$p(x|H_i) = \frac{\exp(-\frac{1}{2}(x-\mu_i)^t \Sigma_i^{-1}(x-\mu_i))}{(2\pi)^{d/2}|\Sigma_i|^{1/2}}$$

where μ_i and Σ_i are the mean and covariance of feature vectors of class i and d is the dimension of the feature space. Discarding terms in $g(x)$ that is independent of i, we have

$$g_i(x) = -\frac{1}{2}(x-\mu_i)^t \Sigma_i^{-1}(x-\mu_i) - \frac{1}{2}\log|\Sigma_i|. \quad (1)$$

When training data is only available for normal operation (hypothesis H), then $g(x) = p(x|H)$ represents the key statistic to measure. We define a "health index," $h(x)$ as follows:

$$h(x) = ((x-\mu)^t \Sigma^{-1}(x-\mu))^{1/2}$$

$h(x)$ is sometimes called the *Mahalanobis distance* [10], and measures how far x is away from normal behavior. This is the primary statistic used in the STFT analysis.

Note, however, that the covariance matrix Σ_i in Eq. (1) is an $N \times N$ matrix where N is the dimension of the composite feature space. Thus determining Σ_i requires the estimation of $(N^2+N)/2$ scalars. This can lead to overfitting of data if the feature space dimension is large, and the sample pool is small.

For feature spaces compressed with PCA, the full covariance matrix is estimated directly in order to compute $h(x)$. However, for the wavelet analysis, if the feature space is not compressed, then data overfitting is an issue. One way to circumvent this difficulty is to assume $\Sigma_i = \sigma^2 I$ where σ^2 is the Euclidean-norm variance of the feature vectors from the training data.

In this case, the discriminant function $g_i(x)$ simplifies to the following "similarity" measure:

$$s_i(x) = -\frac{|x-\mu_i|^2}{2\sigma^2} - \log\sigma \quad (2)$$

This function measures how close the input signal is to each training sample cluster. The surfaces of constant distance are hyperspheres as opposed to the hyperellipsoids measured by $h(x)$. In practice this means that each feature space or sensor reading is considered independently and cross-correlations are ignored when $s_i(x)$ is used; however the technique is robust for high dimensional spaces since only one variance parameter is estimated.

Alternatively, the approach in [12] does not assume feature independence and instead uses a probabilistic graph to model the dependencies among different levels at additional computational cost.

7 Results

This section describes results for using the diagnostics techniques on data from the DC265 paper drive plate testbed.

Fig. 6 shows the result of applying the STFT-based analysis outlined in Fig. 3(a) onto main motor vibration data from 3 sensors mounted on different parts of the drive plate. Training data is only used for the "normal" motor vibration, and the objective is to see if the resulting health index can distinguish between normal and compromised motor behavior.

A window of length 655 ms is used. Longer windows capture lower frequencies better but are more computationally intensive and take longer to respond. The first 2 principal component features and one residual component feature are kept during the PCA compression stage of the algorithm. Keeping more principal components keeps more information but encourages dimensionality growth. Both parameters may be chosen during training by looking at signal-to-noise ratios as described below.

Fig. 6(a) shows the individual time-series trace of the health index from each of the three sensors. There are six traces corresponding to the results for each sensor on two vibration data sets: one with a normal and one with a compromised motor running. Sensor #2 is the most sensitive to the behavior difference but all three sensors show a separation between the responses of normal and compromised motors (note that smaller values of the health index indicate input closer to normal). Fig. 6(b) shows the composite health index when data from all three sensors are aggregated together. It is clear that a simple thresholding would perform well for selecting normal from compromised behavior.

The composite index performs better than any of

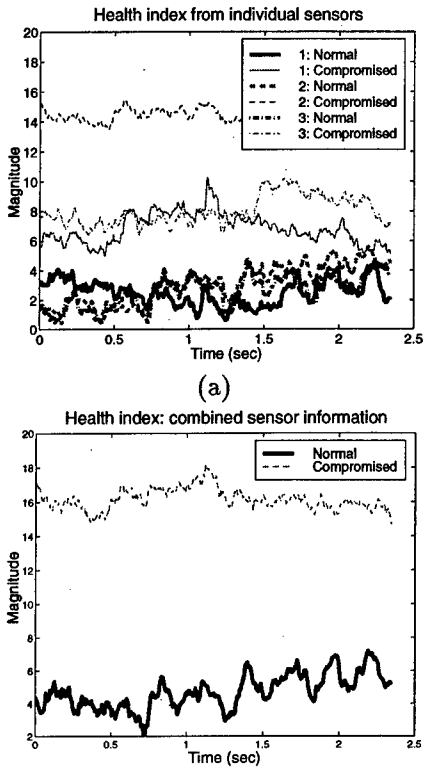


Figure 6: (a) Health index for three vibration sensors run on two different data sets: when the normal main motor and when the compromised main motor is running. (b) Composite health index of all three sensors for the same two data set.

the individual sensors. This can be shown by considering signal-to-noise (S/N) ratio statistics. For two hypotheses H_1 and H_2, we define the S/N ratio $r(x)$ of health index $x(t)$ as follows:

$$r(x) = \frac{|E(x|H_1) - E(x|H_2)|}{\sqrt{0.5(Var(x|H_1) + Var(x|H_2))}}$$

where $E(x|H_i)$ is the expected value of x given H_i and $Var(x|H_i)$ is the variance statistic of x given H_i. Table 1 details not only the S/N ratio contribution for each sensor, but also breaks down statistics for each principal component in the sensor feature spaces. Normally one would expect that additional information would tend to improve S/N ratio. Indeed, the composite S/N ratio is higher than that of any single sensor; however, because of the limited sample size of the data, this is not always the case as with the principal component breakdown for sensor #2.

Fig. 7 shows the time-series trace of the health statistic from the same three vibration sensors used in the previous experiment except this time the

	component			total
	1	2	3+	
sensor 1	1.79	0.81	4.00	4.69
sensor 2	4.88	0.75	13.60	11.48
sensor 3	2.93	0.74	5.01	5.38
composite				13.56

Table 1: Signal-to-noise ratio results from multi-sensor STFT analysis for distinguishing normal vs. compromised main motor vibration.

input consists of pull-in firings from a solenoid. Virtually the same algorithm is used for detecting solenoid behavior as for with the motor except that the window length used is a factor of 4 less (164 ms), since low frequencies are less important for solenoids, and 3 principal components are kept instead of 2.

Fig. 7(a) shows 6 time-series of health index traces. Three traces show individual sensor health index results from an experiment with a normal solenoid, three traces show results from a different experiment with an abnormal solenoid. Fig. 7(b) shows composite health index results for the two experiments. We see that the STFT method is flexible enough to detect normal vs. compromised behavior for solenoids as well as motors.

The final two figures show results from the wavelet-based analysis outlined in Fig. 3(b). In this case, we assume training data is available from all the various operating conditions, and the problem is to choose which operating condition or fault hypothesis is correct given new input data.

Wavelet coefficient features are extracted from the training data for solenoid and motor vibration signals and compressed using the thresholding technique. We compute the coefficients using the discrete wavelet transform with the Daubechies db5-3 basis functions on windowed signals of length 40 ms.

For each operating condition training data set, the resulting feature space data is used to generate the similarity measure given in Eq. (2). Figs. 8 and 9 plot the similarity measure for signals with respect to each of the known conditions or faults. Using the output of the similarity measure, the system classifies the signal at each time into one of the conditions or faults based on the classification of the dominant response.

In Fig 8, the objective is to classify vibration data from a solenoid pull-in event. The similarity measure exhibits a dominant peak response for the solenoid pull-in condition (solid line), indicat-

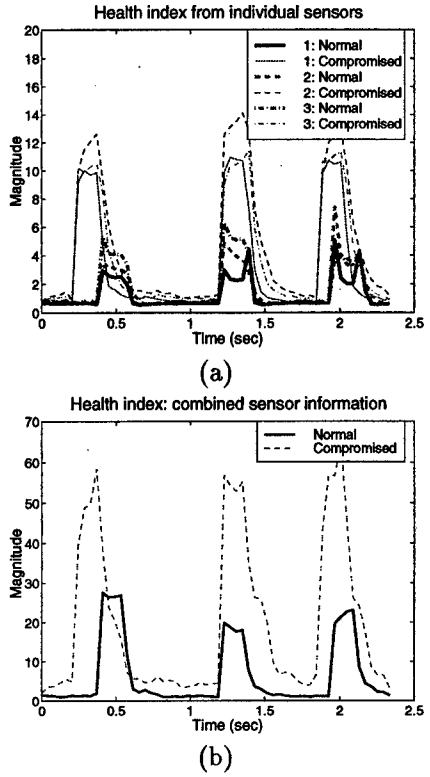


Figure 7: (a) Health index for three vibration sensors run on two different data sets: pull-in vibration data from a normal and compromised solenoid. (b) Composite health index of all three sensors for the same two data sets.

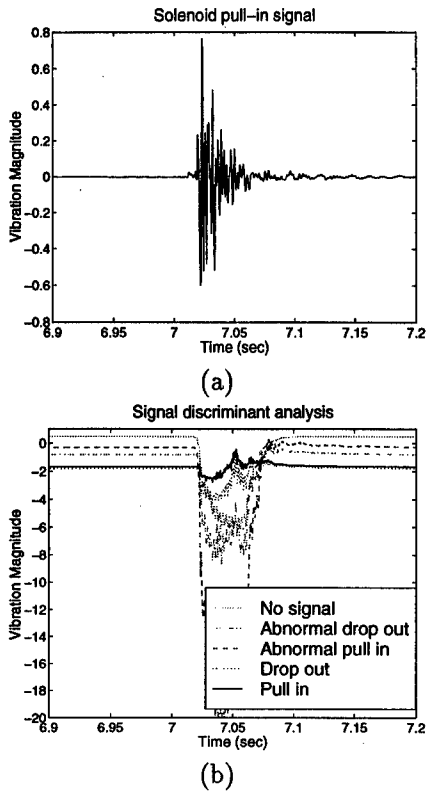


Figure 8: Wavelet-based signal discriminant analysis of solenoid signals: (a) Solenoid pull-in signal; (b) Similarity measures for five conditions: normal pull in, normal drop out, abnormal pull in, abnormal drop out, and no signal.

ing that a pull-in condition has occurred (the time for the event is determined by subtracting the moving window length from the time the peak response in the similarity measure occurs).

In Fig. 9, a solenoid pull-in signal is mixed with a motor signal. In this case, the similarity measure for the pull-in condition also responded with a dominating peak. Note that after the solenoid transient tapers off, the measure for the motor becomes dominant.

8 Conclusion

In this paper we demonstrate how time-frequency analysis of multiple sensors can be used to diagnose actuator behavior based on vibration data from a complex multi-mode electromechanical system, the Xerox DC265 paper drive plate.

The general objective of this work is to develop scalable processing techniques that are flexible enough to perform on a wide class of distributed sensor and actuator systems without high application-specific overhead. The time-frequency sensor data aggregation techniques in this paper appear to be promising tools, since multi-sensor diagnosis results are demonstrated using radically different solenoid and motor vibration signals with little adjustment in the algorithms. In the case of the wavelet technique, five different fault conditions are reliably classified.

9 Acknowledgment

The authors would like to thank Eric Manders for instrumenting the DC265 drive plate testbed. We would also like to thank Bob Siegel and Charles Coleman for providing the drive plates and helpful discussion about diagnostics techniques. In addition, we thank Eric Manders, Eric Jackson, Koenraad Van Schuylenbergh, Sriram Narasimhan, and John Gilbert for testbed setup assistance, data collection, and useful discussions about algorithms.

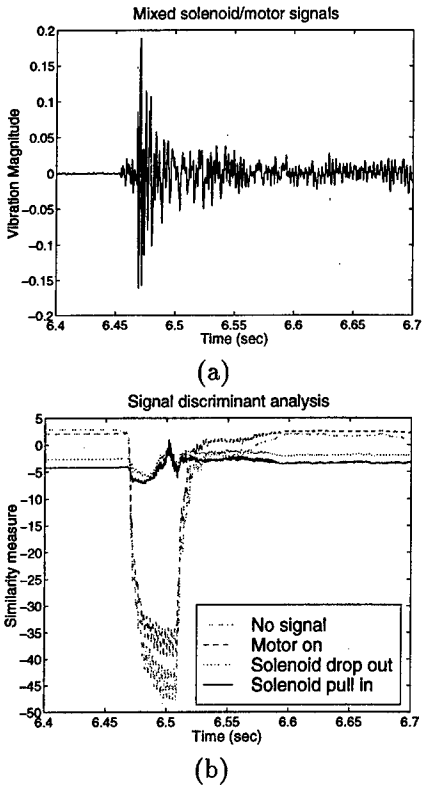


Figure 9: Wavelet-based signal discriminant analysis of mixed motor and solenoid signals: (a) Solenoid pull-in with motor turned on; (b) Similarity measures for four conditions: solenoid pull in, solenoid drop out, motor on, and no signal.

References

[1] B. Samimy and G. Rizzoni, "Mechanical signature analysis using time-frequency signal" *Proceedings of the IEEE*, 84(9), 1996, 1330-1343.

[2] A.V. Oppenheim and R.W. Schafer, *Discrete-Time Signal Processing*, Prentice Hall, 1989.

[3] S. Mallat, *A Wavelet Tour of Signal Processing*. Academic Press, 1998.

[4] R.J. Feriez and D.C. Lang, "Gear-tooth fault detection and tracking using the wavelet transform," *Proc. Prognosis of Residual Life of Machinery and Structures*, MFPT, 1998.

[5] Y. Wang, "Jump and sharp cusp detection by wavelets," *Biometrika*, 82(2):385-97, 1995.

[6] K. Bossley, R. Mckendrick, C.J. Harris, and C. Mercer, "Wavelet packet analysis in the condition monitoring of rotating machinery," *Proc. of Symp. on AI in Real-Time Control*, Pergman Press, 1998.

[7] C. Liu and H. Wechsler, "A unified Bayesian framework for face recognition," *Proc. ICIP '98*, v. 1, 1998, 151-155.

[8] D. Beyerback and H. Nawab, "Principal components analysis of the short-time Fourier transform," *Proc. ICASSP '91*, v. 3, 1991, 1725-1728.

[9] J. Manyika and H. Durrant-Whyte, *Data Fusion and Sensor Management*, Ellis Horwood, 1994.

[10] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis*, Wiley, 1973.

[11] R.A. Redner and H.F. Walker, "Mixture densities, maximum likelihood and the EM Algorithm," *SIAM Review*, 26(2), 1984, 195-239.

[12] M.S. Crouse, R.D. Nowak, and R.G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models." *IEEE Trans. Signal Processing*, 46(4), 1998.

Sensory Based Expert Monitoring and Control

Gary G. Yen **Haiming Lu**
Oklahoma State University
School of Electrical and Computer Engineering
Stillwater, OK 74078

Abstract - The gas-liquid two-phase flow is widely used in petro-chemical and energy industry. Accurate measurements of flow parameters such as flow regimes is the key to operating efficiency. However, due to the complexity of the characteristics in two-phase flow, it is very difficult to monitor and distinguish flow patterns on-line, *in-situ*. In this paper we proposed an efficient Acoustic Emission(AE)-based detection system combined with Fuzzy-Neural network to recognize four major patterns in air-water vertical two-phase column experimentally. Several crucial AE parameters are assessed and compared, and we found the density of AE Ring-Down Counts is an excellent indicator for the flow pattern recognition problem. A Fuzzy-Neural network is designated as a decision-maker to indicate an approximate transmission stage of a given two-phase flow.

Key Words: acoustic emission, fuzzy neural network, pattern classification

I. Introduction

The gas-liquid two-phase flow is defined as the flows of mixture of two homogeneous phases (i.e., gas and liquid) through a system. Since it often aid the description of heat and mass transfer mechanisms in a system, it plays a very important role and is popularly used in petrochemical and chemical process industries, energy and nuclear industry and biological engineering as well [1]. It has been proved that the operating efficiency of such a process is closely related to accurate measurement of flow parameters such as flow regimes and multiple flow velocities [2]. Generally speaking, flow patterns are classified as Bubbly [3], Slug [4], Churn [5] and Annular [6] regimes, these flow regimes typically have distinct flow characteristics and heat and mass transfer mechanisms, which are very useful for detail study in this field. Some detection techniques were applied to monitor and detect flow pattern based upon capture images or measurements of flow velocity. Xu [7] established a mathematical model (i.e., two-value (0/1) logical back-projection filtering algorithm) combining with a transmission-mode ultrasound computerized tomography system to reconstruct the image of a distribution of bubbles over a 2-D cross-section of a pipe, for both parallel beam scanning and fan-shape beam scanning geometry. Albusaidi and Lucas [8] proposed a technique, which consists of mounting an array of 64 axially separated conductivity sensors in a vertical pipe through which air/water mixture is flowing, to obtain the mean Cap bubble (or Taylor bubble) velocity and hence an estimation of the mean gas velocity by cross-correlation of the output signals. In thi spaper, we propose to use a NDT (Non-Destructive Test) method based on Acoustic Emission (AE) signals to examine gas-liquid two-phase flow phenomenon and classify the four major flow patterns as noted above.

Acoustic Emission is a term describing a class of phenomena whereby transient elastic waves are generated by the rapid release of energy from localized sources within a material. AE has developed rapidly over the last two decades as a nondestructive evaluation technique and a tool for material research. It's a high sensitivity technique for detecting active microscopic events in a material and has been successfully used in the field of monitoring the welding or crack in solid materials, such like metal, glass and ceramic under stress [9]. Some acoustic emission sensors were designed for monitoring the kinetics of chemical reaction [10]. In this paper, a system with AE method was applied to detect and classify four major regimes: bubbly, slug, churn and annular of vertical air-water two-phase flow.

II. Objective

The water flow rate characteristics of these four patterns are shown in Figure 1. Each of these four patterns has distinguished air/water density and flow speed ratio. The aim of our designed system is to

monitor the status of vertical air-water two-phase flow, distinguish four classes of major flow patterns and analyze some characteristics of those flow patterns.

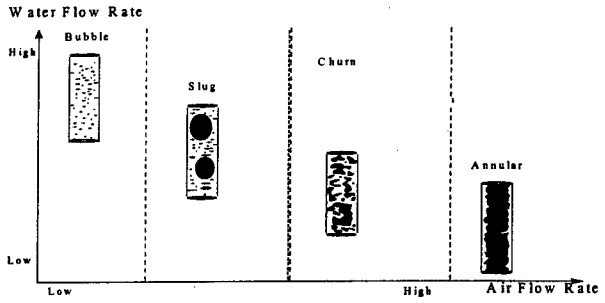


Figure 1 Water/air flow ratio of four major air-water two-phase flow patterns.

III. AE Parameters and Detection System

3.1 AE parameters

Figures 2a and 2b show the standard AE signal in time domain and frequency domain, respectively. Figure 2a shows one AE event (hit), which occurs as the amplitude value of sensor output exceeds a pre-specified threshold (e.g., 35dB). The period between hit start point and end point are called duration, the time series between duration are defined as AE Ring-Down Counts.

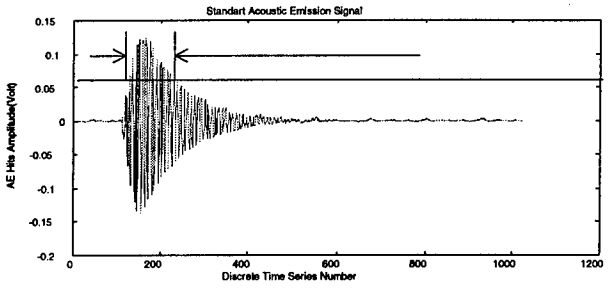


Figure 2a Standard AE event in time domain.

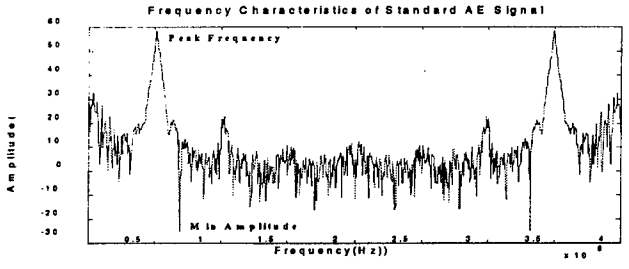


Figure 2b Standard AE event in frequency domain.

3.2 AE detection system

The experimental setup of the two-phase air-water column with AE sensor configuration is shown in Figure 3. The system includes: data acquisition part, signal processing part, data analysis and decision making part and data output part.

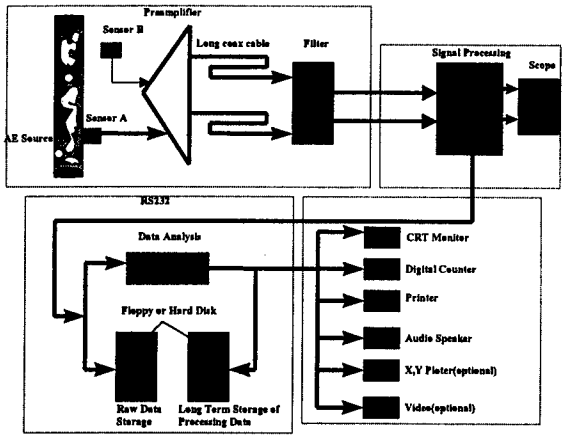


Figure 3 Configuration of experimental AE detection system.

In Figure 3, sensor A is an AE sensor which is glued with epoxy on the pipe for detecting AE signals occurs in flow, sensor B is another AE sensor placed near pipe and is dedicated for detecting background noise.

AE parameters is chosen as follows based upon some advice from field expert in our experiments:

- AE sensor resonant frequency: 150kHZ
- Sampling rate of DSP board : 4MHz
- Gain of amplifier: 40dB& 60dB
- Time window of each AE Hit: 1024 point (256us)
- Threshold voltage: 0.0586V(35dB)
- Data acquisition time: 90second

IV. Experimental Results

4.1 Time series data characteristics

Different parameters of AE signals were compared to see if these four flow regimes can be distinguished linearly or nonlinearly.

a. For the Peak Amplitude of AE hits, Probability Distribution Figure of Amplitude of AE Hits as shown in Figure 4 was analyzed.

b. For number of AE hits, Number of AE Hits Occurred in Each Second as shown in Figure 5 was analyzed.

Result 1

From Figure 4, pattern of ***Bubbly*** can be easily distinguished with the other three patterns from the value of Peak amplitude of AE hits or probability distribution of amplitude of AE hits. But it's very difficult to distinguish ***Slug, Churn*** and ***Annular*** pattern when we only use Peak Amplitude parameter.

From Figure 5, patterns of ***Bubbly, Slug*** and ***Churn*** can be classified separately from the number of AE hits occur in the given time range (1 second). But it's difficult to recognize the difference between ***Churn*** and ***Annular*** only by AE hits number. So other parameters must be introduced to solve this problem.

c. AE Ring Down Counts definition:

Since AE signals in our experiment are of relatively short duration (less than 1msec), reach maximum amplitude early in the signal (always assume 0), and decay nearly in exponentially—as shown in Figure 1, we can calculate the sensor output as:

$$V(t) = V_0\, e^{-rt} \sin wt \qquad (1)$$

where:

$V(t)$ =output voltage of sensor

V_0 =initial signal amplitude

r =decay constant(>0)

t =time

w =signal frequency

Since Threshold voltage V^* has been setup, we can count the number of times the sensor voltage exceeds it---this technique is known as ring down counting. For the signal represented by Eq. (1), the number of counts(N) to the nearest integer is given by:

$$\mathrm{N}=\frac{t^*}{2\pi / w} = \frac{w}{2\pi\gamma} \ln \frac{V_0}{V^*} \qquad (2)$$

Where $V^* = V_0 e^{-rt^*}$

$$\text{And} \quad t^* = \frac{1}{\gamma} \ln \frac{V_0}{V^*} \qquad (3)$$

Thus we can analyze the number of AE Counts Occurring in Each Second (Figure.6*)*, the maximum and minimum density of AE counts for each of four patterns is shown in Table 1

Table 1 Density range of four major two-phase flow patterns

	Bubbly	*Slug*	*Churn*	*Annular*
Maximum number of AE counts occurs in each second	27	749	7196	22401
Maximum number of AE counts occurs in each second	0	86	2703	12478

Result 2

From Figure 6 and Table 1, the difference of density of ring down counting for AE signal of these four patterns is obvious, and it can be an excellent indicator for this pattern recognition problem. Combine with the result of AE hits, such conclusion can be drawn that the number of AE counts of each AE hit is different. This can be shown in Figure 7. It also can be deduced that the areas of AE signal for these four patterns are different. Energy can be calculated to present these AE events as following equations:

$$E = \frac{1}{R}\int_0^\infty V^2(t)dt \qquad (4)$$

Here, R is resistance, Equations 1 and 4 can be combined to yield (assuming the signal decays to background level (V^*) after a time t

$$E = \frac{1}{R}\int_0^\infty V_0 e^{-2rt} \sin^2 wtdt$$

$$= \frac{V_0^2}{4R(w^2+r^2)}\left[\frac{w^2}{r^2}(1-e^{-2rt^*}) - 2e^{-2rt^*} \sin wt^* \right. \qquad (5)$$

$$\left. (r \sin wt^* + w\cos wt^*)\right]$$

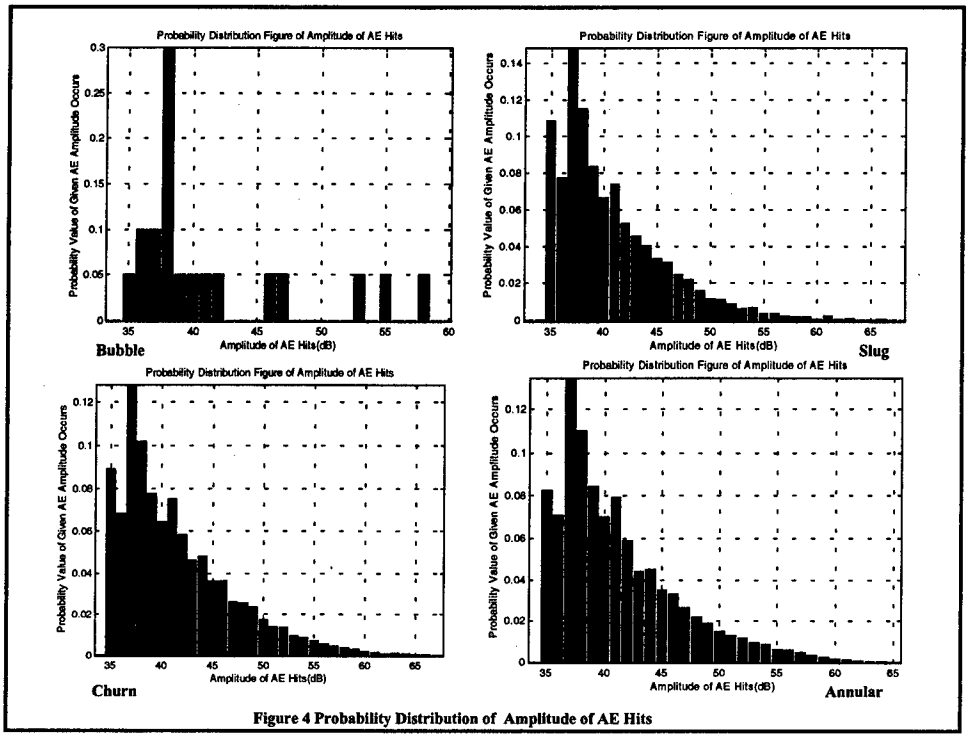


Figure 4 Probability Distribution of Amplitude of AE Hits

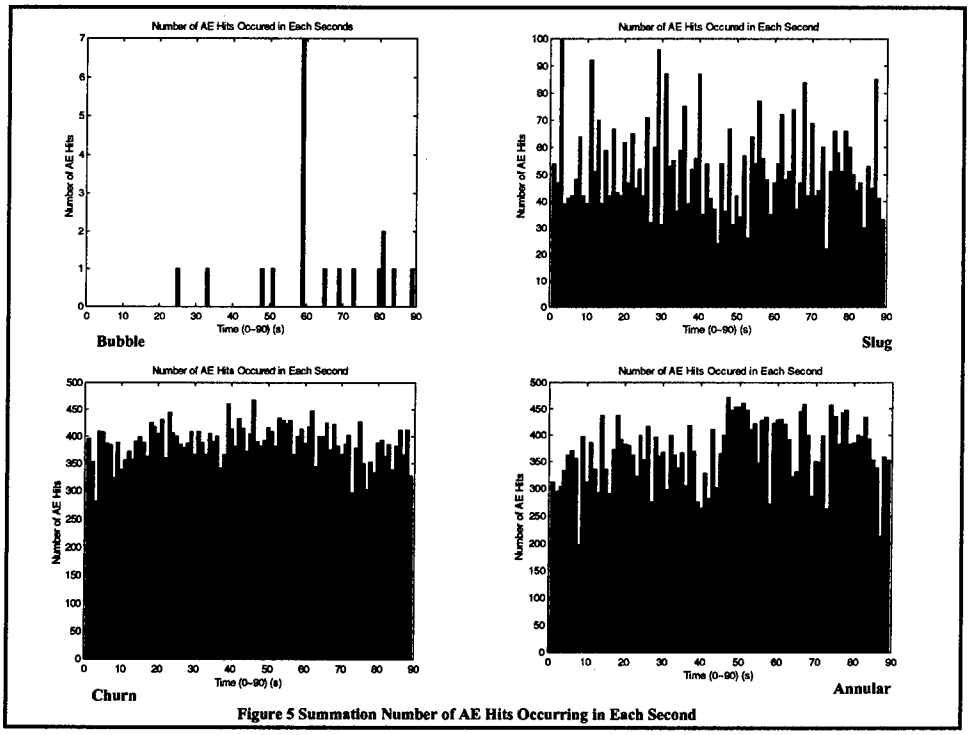


Figure 5 Summation Number of AE Hits Occurring in Each Second

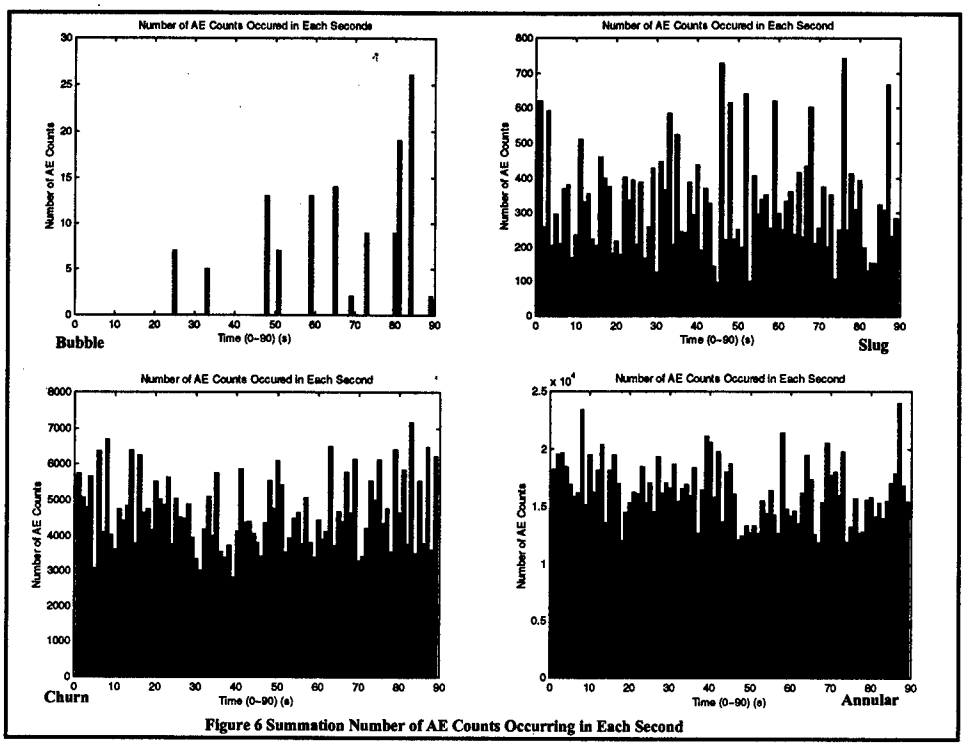


Figure 6 Summation Number of AE Counts Occurring in Each Second

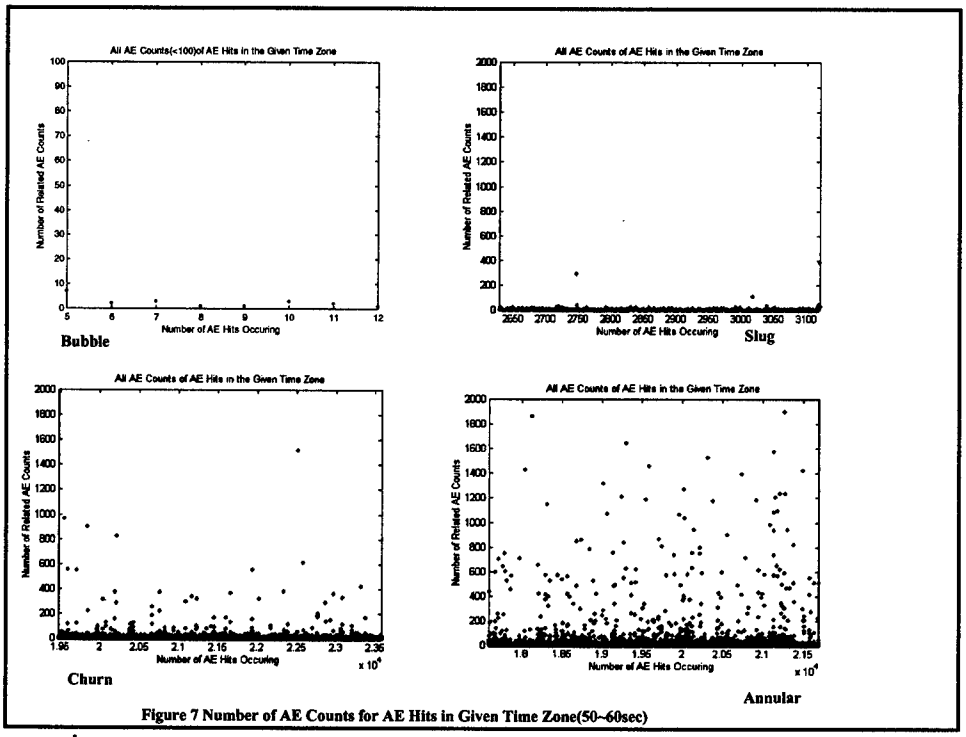


Figure 7 Number of AE Counts for AE Hits in Given Time Zone(50~60sec)

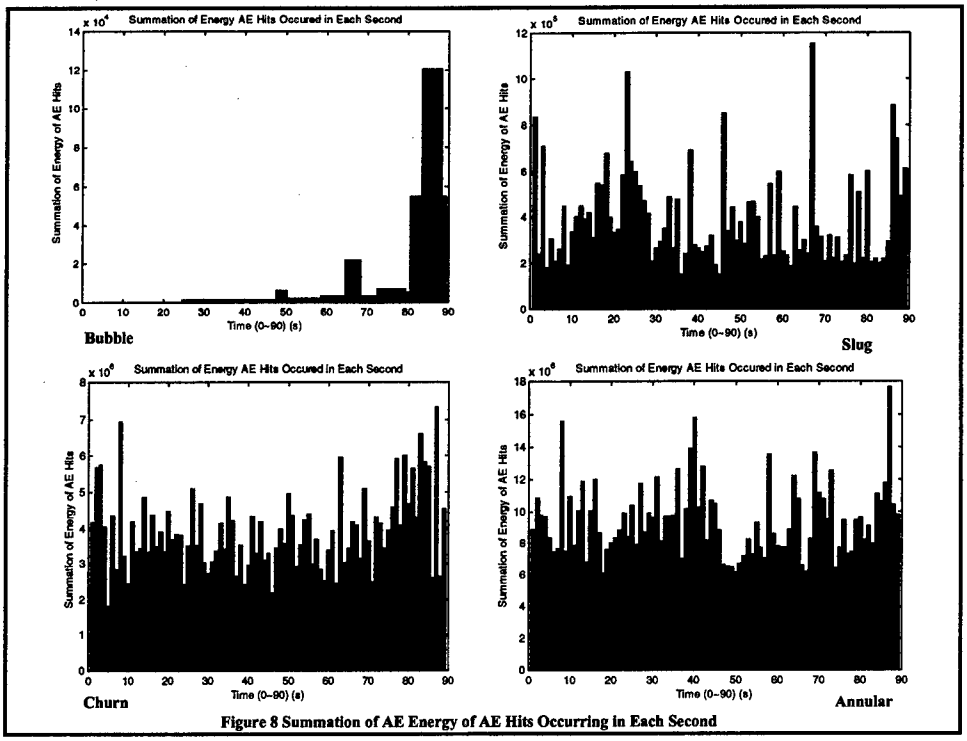


Figure 8 Summation of AE Energy of AE Hits Occurring in Each Second

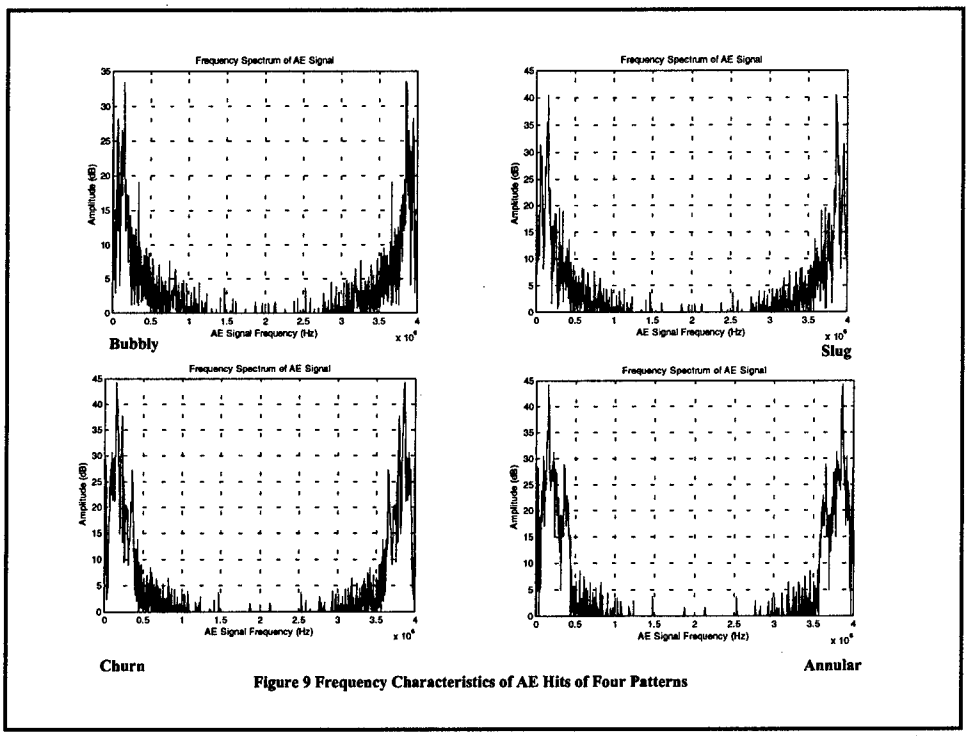


Figure 9 Frequency Characteristics of AE Hits of Four Patterns

If t is assumed that $e^{-2rt^*} \ll 1$, the Equation 5 can be written as

$$E \cong \frac{V^2 w^2}{4Rr(r^2+w^2)} \quad (6)$$

Figure 8 shows the result of Equation 6.

4.2 Frequency Domain and Noise Cancellation

In Frequency domain, FFT of AE signals for these four patterns was shown in Figure 9 and it is obvious that to identify the patterns in frequency domain is more difficult than in time domain. This is because the resonant frequency of AE sensor is almost fixed (150KHz), which makes frequency characteristics more non-intuitive in most spectrum range except resonant frequency. In our experiment environment, background noise is not overwhelmed and since frequency characteristics is not crucial for experiment result, noise effect can be ignored here.

4.3 Decision Maker

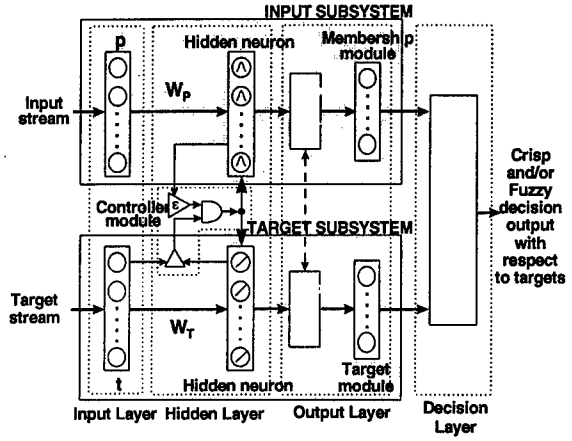


Figure 10 Structure of Decision-maker Fuzzy-Neural network.

Figure 10 shows decision-maker network of the system [11]. This is a Fuzzy-Neural network. Input stream includes two parameters, i.e., AE hit and AE count density (number ouccured in one second)—of four flow patterns. Target stream includes ideal values of these two parameters for the same flow pattern. Neural network is trained to make correct decision to described air-water two-phase flow stage. This decision-maker not only recognizes those four major patterns mentioned above accurately, but tell the pattern transmission stage approximately through fuzzy decision output.

V. Conclusion

Application of AE and Fuzzy-Neural Network on air-water two-phase flow pattern recognition problem was proposed and discussed. In this study, several AE parameters were extracted from four major two-phase flows pattern signals and the results were discussed. AE events and Ring-Down Counts density can be combined as a stable and excellent indicator to describe flow patterns accurately. They form the input stream of Fuzzy-Neural network and after training the network, system output can tell the continuous flow stage (includes four major patterns and transition part) on line in real-time.

References

[1] G.Hewitt, *Measurement of Two Phase Flow Parameters,* London: Academic, 1978.

[2] W.Galegar, W.Stovall and R.Huntington, "More Data on Two-phase Vertical Flow," *Pet. Ref.*, 33,208, 1954.

[3] E.Rhodes, "Gas-Liquid Flow," *Heat Transfer & Fluid Flow Services, Atomic Energy Research Establishment*, Harwell, England,1982.

[4] R.Fernandes, R.Semiat and A.Dukler, "Hydrodynamic Model for Gas-Liquid Slug Flow in Vertical Tubes," *AIChEJ.*, 29, 981, 1983.

[5] Y.Taitel, D.Bornea and A.Dukler, "Modelling Flow Pattern Transmissions for Steady Upward Gas-Liquid Flow in Vertical Tubes," *AIChEJ.*, 26, 345, 1980.

[6] G.Hewitt, "Disturbance Waves in Annular Two-Phase Flow", *Proc. Inst. Mech. Engrs.*, 184, 142, 1969.

[7] L.Xu and L.Xu, "Ultrasound tomography system used for monitoring bubbly gas/liquid two-phase flow", *IEEE Transaction on ultrosonics, ferroelectrics, and frequency control*, 44,1, 67-74, 1997.

[8] K.Albusaidi and G.Lucas, "Measurement of multiple velocities in multiphase flow", *IEEE Colloquium on advances in sensors for fluid flow measurement,* pp12/1-12/4.

[9] M.Bassim, M.Dudar, R.Rifaat and R.Roller, "Application of acoustic emission for non destructive evaluation of utility inductive reactors", *IEEE Transaction on Power Delivery,* 8, 1, pp281-284, 1993 .

[10] S.Bang, R.Lec, J.Genco and J.Ransdell, "Acoustic emission chemical sensor" *IEEE Proceedings of Ultrasonic Symposium, 1993. vol.1*, pp439-443, 1993.

[11] G.Yen and P.Meesad, "An effective neuro-fuzzy paradigm for machinery condition health monitoring," Proceedings of International Joint Conference on Neural Network, July 1999, to appear.

Diagnosing Valve Clearance Fault using Multi-Parameter Fusion

Lixiang Shen and **Francis E.H. Tay**

Department of Mechanical and Production Engineering
National University of Singapore, Singapore, 119260
{engp8633, mpetayeh}@nus.edu.sg

Abstract - *In this paper, the Dempster-Shafer evidence reasoning is used in a new domain - Fault Diagnosis of Reciprocating Machinery. Other than the new application for the Demspter-Shafer approach, this paper describes a new way of implementation - multi-parameter fusion - which requires selected parameters extracted from every sensor to be fused. This is in contrast with existing methods that need the eigenvectors to be fused. By selecting the relevant parameters as compared to generating the eigenvectors, this method is much easier to implement. Through the implementation, it is shown that this method decreases the uncertainty in the diagnosis systems. In addition, after the preprocessing of the information extracted from each sensor, it can reduce the computing time at the fusion stage.*

Keywords: Dempster-Shafer evidential reasoning, fault diagnosis, multi-parameter fusion

1 Introduction

Sensor systems have been improving rapidly and ancillary data are increasingly available. As a result, interest in extracting the higher level information contained in all kinds of sensing contexts has led to extensive demand for computer-based, automated methods for the analysis of multi-source data. This requirement gives rise to many information integration and fusion methods. The potential advantages in integrating and fusing information from multiple sensors are that the information can be obtained more accurately and in less time and at a less cost. Features that are impossible to perceive with individual sensor can now be obtained through the use of multiple sensors. Multiple sensors, together with information integration and fusion, have been used in many areas, such as navigation, target identification, robot control and multi-target tracking.

In this paper, the application of information integration and fusion has been broadened to fault diagnosis in machinery. Up to now, the conventional method used to diagnose fault in machinery is to observe the change of FFT (Fast Fourier Transformation) spectrum of a single sensor. For large rotational machinery, because its signal from displacement sensor is very simple, like pure sine curve, the information presented in the spectrum is enough to be used for diagnosis. However for reciprocating machinery, such as reciprocating compressor and diesel engine, due to the complex structure and multi-excite sources existing in diesel engine, the vibration signals collected from the engine surface have the following characteristics:

- Presence of a number of self-exciting vibration and forced vibration in the diesel engine that is running. Therefore the width of spectrum in frequency domain is very large.
- The vibration signals in the time domain are more complex compared to a large-scale rotational machinery, which is a pure sine curve.
- In a diesel engine, such as 4135 engine, the stroke cycles are fixed. Therefore the time series appear periodical. However in every period, there exist many other periodical vibrations within the stroke cycle.

So using single sensor's information is not enough to diagnose fault types. In this paper, several acceleration sensors are used to sample the vibration signals from the surface of a 4135 diesel engine. A new fusion method, multi-parameter fusion, is used to get the final judgment.

The remainder of this paper is organized as follows. In the second section, the fundamental knowledge of evidence reasoning is introduced. The parameters used for

information fusion are studied in the third section of this paper. In the fourth section, the above selected parameters are fused and the final judgment is given. In the same section, some questions related to fusion are discussed. And the conclusion is given at the end of this paper.

2 Preliminary of Evidence Reasoning

There are many algorithms, which can be used to integrate multiple sensors' information. Among them, distributed Kalman filtering and the Bayesian approach are well known. Bayesian approach offers a highly formalized and rigorous way to assign and propagate confidence. However, these algorithms require substantial a prior information, such as initial values and initial covariance matrices for distributed Kalman filtering and prior probabilities for the Bayesian approach. In many cases, prior information is either unavailable or not known precisely. Another weakness is that they cannot represent uncertainty in the systems very well. These inadequacies give rise to the Evidence Reasoning.

Evidence Reasoning, also called the 'Dempster-Shafer theory' or the 'belief function theory', has been found useful in dealing with uncertainty in many domains, for instance, diagnostic system and radar surveillance system.

The basic notions of Evidence Reasoning were presented by G. Shafer (1976). The following concepts are the fundamental concepts of D-S theory: frame of discernment, basic probability assignment (BPA), and belief function and plausibility function. They are introduced as follows.

A frame of discernment Θ is a finite nonempty set.

The basic probability assignment (BPA) on Θ is a function

$$m: P(\Theta) \rightarrow R_+ \tag{1}$$

where $P(\Theta)$ is the powerset of Θ and R_+ is the set of nonnegative reals, satisfying the following conditions:

$$1.\ m(\varnothing) = 0 \tag{2}$$

$$2.\ \sum_{\Delta \subseteq \Theta} m(\Delta) = 1 \tag{3}$$

For a given basic probability assignment m two functions are defined.

• A function Bel: $P(\Theta) \rightarrow R_+$ is called the belief function over Θ (generated by m) iff for any $\theta \subseteq \Theta$,

$$Bel(\theta) = \sum_{\Delta \subseteq \theta} m(\Delta) \tag{4}$$

• A function Pl: $P(\Theta) \rightarrow R_+$ is called the plausibility function over Θ (generated by m) iff for any $\theta \subseteq \Theta$,

$$Pl(\theta) = \sum_{\Delta \cap \theta \neq \varnothing} m(\Delta) \tag{5}$$

The plausibility function Pl can be definable by the belief function Bel:

$$Pl(\theta) = 1 - Bel(\Theta - \theta),\ for\ \theta \subseteq \Theta \tag{6}$$

From a given belief function, a basic probability assignment can be reconstructed:

$$m(\theta) = \sum_{\Delta \subseteq \theta} (-1)^{|\theta - \Delta|} Bel(\Delta),\ for\ \theta \subseteq \Theta \tag{7}$$

The union of all subsets $\theta \subseteq \Theta$ that are *focals* is called the *core* of Θ.

A belief function Bel is called *Bayesian belief function* iff $Bel(\theta) + Bel(\Theta - \theta) = 1$ *for* $\theta \subseteq \Theta$. The following conditions are equivalent.

1. Bel is *Bayesian*.
2. $Bel(\theta \cup \Delta) = Bel(\theta) + Bel(\Delta)$, *for* $\theta, \Delta \subseteq \Theta$ *and* $\theta \cap \Theta = \varnothing$; (8)
3. $Bel = Pl$;
4. All focal elements are singletons.

Dempster's Rule of Evidence Combination:

Evidence obtained on the same subject from two probabilistically independent sources can be combined into joint evidence of the subject. For instance, two pieces of evidence expressed by two basic-probability-assignments $m_1(A)$ and $m_2(B)$ can be combined into a signal piece of joint evidence by

$$m_{12} = \begin{cases} \dfrac{\sum_{C=A \cap B} m_1(A) \cdot m_2(B)}{1 - K}, & if\ C \neq \varnothing \\ 0, & if\ C = \varnothing \end{cases} \tag{9}$$

where the constant K is

$$K = \sum_{A \cap B = \varnothing} m_1(A) \cdot m_2(B) \qquad (10)$$

which represents conflicting information in these two pieces of evidence. In (9), combined information is normalized after the conflicting information is removed. Dempster's rule reduces to Bayesian approach when the belief function is the same as the plausibility function.

3 Establishing Parameter Field

In this paper, a new fusion method is proposed – multi-parameter fusion. Multi-parameter fusion because the parameters, which represent the information contained in sampled signals, are extracted and these parameters are used in the fusion framework, instead of the eigenvectors.

The case study used in the research is a 4135 diesel engine. The parameters are:

Rated Engine Power: 80 horsepower

Rated Engine Speed: 1500rpm

Four states are simulated on this diesel engine. They are

- Normal
- Intake valve clearance is too small
- Intake valve clearance is too large
- Exhaust valve clearance is too large

Among these four states, three fault types were simulated in the intake valve and exhaust valve on the second cylinder head. Three points are selected to collect vibration signals. They are the first cylinder head, the second cylinder head and another one, which is in the middle point of piston stroke, on the surface of cylinder block.

Six parameters are extracted from the vibration signal of each sampling point. These six parameters can be divided into two categories, frequency domain and time domain. They are introduced as follows:

(1) Frequency domain parameters:

a. *IF* - Waveform Complexity in frequency domain

$$IF = -\sum_{i=1}^{N/2} X(i) \log X(i) \qquad (10)$$

where $X(i)$ - the FFT spectrum

From the equation (10), it can be seen that *IF* is a frequency domain entropy, reflecting the complexity of FFT spectrum.

b. *CG* - the center frequency of spectrum

$$CG(X) = \sum_{K=1}^{N/2} \frac{K}{N/2} \mu(X(K)) \qquad (11)$$

where $\mu(X(K)) = X(K) \Big/ \sum_{j=1}^{N/2} X(j)$

$X(K)$ - the FFT spectrum

$k=1,2,\ldots,N$

(2) Time domain parameters:

a. *IT* - Waveform Complexity in time domain

$$IT = -\sum_{i=1}^{m} \lambda_i \log \lambda_i \qquad (12)$$

where λ_i - the singular value of a time series according to its period

m - the numbers of periods in a time series

The physical significance of *IT* is to reflect the complexity of time series. It is time domain entropy.

b. σ - Nonperiod complexity

$$\sigma = \frac{m}{m-1} \sum_{i=2}^{m} \lambda^2{}_i \Big/ \sum_{i=1}^{m} \lambda^2{}_i \qquad (13)$$

where λ_i - the singular value of a time series according to its period.

c. D_x - the variance of time series

$$D_x = \frac{1}{n} \sum_{i=1}^{n} \left[x(t_i) - \bar{x} \right]^2 \qquad (14)$$

where n - the length of a time series

$\bar{x}$ - the mean value of whole series

$x(t_i)$ - the time series

d. α_4 - the kurtosis of time series

$$\alpha_4 = \frac{1}{n} \sum_{i=1}^{n} \left[x(t_i) \right]^4 \qquad (15)$$

The above six parameters reflect the information contained in vibration signals both from the frequency domain and time domain. For *IT* and σ, they reflect the time series' periodical characteristic because the single fault type shows the periodicity in time domain and the energy will increase in a certain frequency in spectrum, which is reflected by parameters, *IF* and *CG*. Variance D_x and Kurtosis α_4 are the measures of the data distribution.

4 Using Multi-parameter Fusion to Diagnose Valve Fault of a 4135 Diesel Engine

In this section, the above extracted parameters are fused using D-S theory. For simplification, the belief function is only calculated as the final results.

4.1 Defining the Basic Probability Assignment – Mass Function

There are many fusion methods to be used in different domain. In this paper, in light of the characteristics of vibration signal collected from the surface of 4135 diesel engine, the basic probability assignment (mass function) is defined as follows.

$$m_i(A_j)=\frac{w_i C_i(A_j)}{\sum_j w_i C_i(A_j)+N_s(1-R_i)(1-\alpha_i\beta_i)} \quad (16)$$

$$m_i(\theta)=\frac{N_s(1-R_i)(1-\alpha_i\beta_i)}{\sum_j w_i C_i(A_j)+N_s(1-R_i)(1-\alpha_i\beta_i)} \quad (17)$$

$i=1, 2, \ldots, N_s, j=1, 2, \ldots, N_c$

where $C_i(A_j)=1/d_{i,j}(X_i,Y_j)$ – Relation Coefficient (18)

$d_{i,j}(X_i,Y_j)=\sum_{k=1}^{n}|x_{ik}-y_{jk}|$ –Manhattan Distance (19)

$\alpha_i=\max_j\{C_i(A_j)\}=1/\min_j\{d_{i,j}(X_i,Y_j)\}$ – The maximum relation coefficient (20)

$$\beta_i=\left(\frac{N_c\alpha_i}{\sum_j C_i(A_j)}-1\right)\Big/(N_s-1) \quad (21)$$

– comprehensive effect coefficient, which includes the global factors affecting the diagnosis results.

$$R_i=\frac{\alpha_i\beta_j}{\sum_k \alpha_k\beta_k} \quad (22)$$

– balance coefficient

N_c – the number of fault types, $N_c=4$ here.

f_0 – normal

f_1 – small intake valve clearance

f_2 – large intake valve clearance

f_3 – large exhaust valve clearance.

N_s – the number of sensors, $N_s=3$ here.

w_i – weight coefficient, which is determined according to practical experience.

4.2 Applying the Multi-parameter Fusion

In this fusion framework as shown in Figure 1, some assumptions are proposed to process the multi-sensor and multi-parameter fusion. First, different sensors are independent from others. Second, different fault types are independent from others. That is to say, no two fault types can coexist in the engine simultaneously. Table 1 presents the fusion results.

There are 19 cases tested in this paper. Out of the 19 final results, only two cases are wrongly categorised. The verification degree – ratio of correct diagnosis over the total number of cases – is 17/19. This shows that the method is effective in its diagnosis. For illustration, only four cases are listed in the Table 1.

From the results listed in Table 1, it can be seen that using single sensor, some types of fault cannot be determined. These are denoted as 'Uncertain'. After fusing the parameters extracted from every sensor, the verification degree increased while the uncertainty decreases.

In order to reduce computing time, the user can select simple parameters. This is unlike the use of eigenvectors which are fixed as the eignevectors correspond to their respective fault types. In so doing, this multi-parameter fusion method overcomes the shortcomings of the D-S fusion algorithm by simplication and thereby reducing the complexity of the problem.

5 Conclusions

In this paper, a new fusion method – multi-parameter fusion, has proposed and implemented to diagnose the fault types of diesel engine. Through the analysis of fusion results, the following conclusions can be drawn:

- Multi-parameter fusion is a feasible method to be used in fault diagnosis.
- This method has many advantages, such as decreasing uncertainty in the fusion and presents high verification probability as compared to the single sensor. It can reduce the computing complexity when compared with using eigenvector fusion.
- A new diagnosis method using D-S theory has been presented.

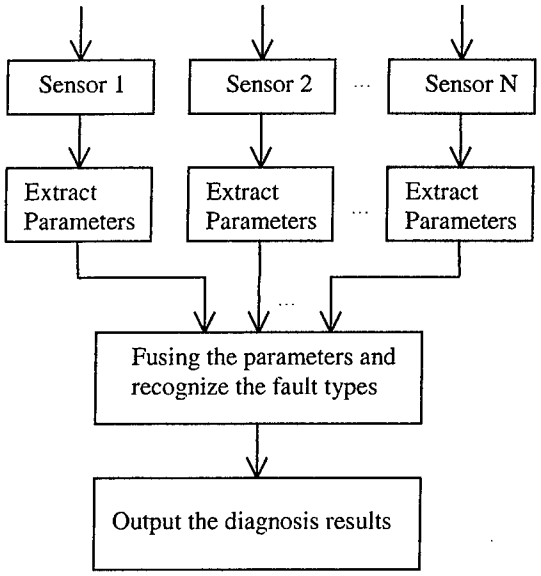


Figure 1. The flow chart of diagnosing system using D-S theory

References

[1] Hong, L. and Lynch, A.: Recursive Temporal-Spatial Information Fusion with Application to Target Identification. IEEE Transactions on Aerospace and Electronic Systems, Vol. 29, No.2: 435-445, 1993.

[2] Luo, R. C. and Kay, M. G.: Multisensor Integration and Fusion in Intelligent Systems, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 19, No. 5: 901-931, 1989.

[3] Bogler, P. L.: Shafer-Dempster Reasoning with Applications to Multisensor Target Identification Systems, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 19, No. 5: 968-977, 1987.

[4] Larson, T. D., Anderson, N. A. and Ravn, O.: Sensor Management for Identity Fusion on a Mobile Robot, The proceedings of the Fifth International Conference on Control, Automation, Robotics and Vision (ICARCV '98), Singapore, 246-250, 9-11, Dec. 1998.

[5] Kim, H. and Swain, P. H.: Evidential Reasoning Approach to Multisource-Data Classification in Remote Sensing, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 25, No. 8: 1257-1265, 1995.

[6] Shen, L.: Analysis of Uncertain Information in Fault Diagnosis, Master Thesis, Xi'an Jiaotong University, P. R. China, 1997.

[7] Li, H.: Information Fusion Technology, National Defense Industry Press, 1997.

[8] Wang, J: A Stuty of Fault Diagnosis Approach for the Fuel Injection System of A Diesel Engine Based on Multisensor Data Fusion Technique, Xi'an Jiaotong University, Master Thesis, P. R. China, 1996.

[9] Dagel, J. F. and Brady, R. N.: Diesel Engine & Fuel System Repair, 4th Ed., Prentice Hall Press, 1998.

[10] Meng, J.: Some Advanced Techniques in Fault Feature Extraction for Large Rotating Machinery, Ph.D Thesis, Xi'an Jiaotong University, P. R. China, 1996.

[11] Liu, H.: On the Integration and Extraction of diagnostic Information, Ph.D Thesis, Xi'an Jioatong University, P. R. China, 1997.

Table 1. The multi-sensor and multi-parameter fusion result

Fault types	Sensor	Mass function $m(f_i)$/belief function $Bel(f_i)$				
		$m(\theta)$	$m(f_0)$	$m(f_1)$	$m(f_2)$	$m(f_3)$
f_0 (Normal)	Sensor1	0.4962	0.3704	0.0073	0.0146	0.1115
	Sensor2	0.4433	0.4857	0.0695	0.0007	0.0009
	Sensor3	0.3825	0.6036	0.0043	0.0035	0.0061
	Fusion	0.2097	0.7225	0.0261	0.0058	0.0360
f_1 (Intake Valve Clearance is too small)	Sensor1	0.3566	0.0046	0.5921	0.0420	0.0047
	Sensor2	0.5167	0.1980	0.2126	0.0320	0.0407
	Sensor3	0.4400	0.0032	0.4377	0.0030	0.1162
	Fusion	0.2234	0.0654	0.6272	0.0274	0.0565
f_2 (Intake Valve Clearance is too Large)	Sensor1	0.5178	0.0262	0.1925	0.2394	0.0242
	Sensor2	0.4821	0.0030	0.0039	0.3562	0.1548
	Sensor3	0.3048	0.0015	0.0010	0.6867	0.0024
	Fusion	0.2181	0.0097	0.0615	0.6506	0.0601
f_3 (Exhaust Valve Clearance is too Large)	Sensor1	0.4132	0.0532	0.0053	0.0115	0.5168
	Sensor2	0.5172	0.0122	0.0149	0.2771	0.1785
	Sensor3	0.4155	0.0062	0.0408	0.0251	0.5125
	Fusion	0.2352	0.0261	0.0220	0.1074	0.6093

Application of Neural Fusion to Accident Forecast in Hydropower Station

XU Lingyu **ZhAO Hai** **XIANG Xin** **DU Qingdong**

School of Information Science and Engineering

Northeastern University

Shenyang 110006

P.R .China

Abstract-This paper deals with a new application of neural fusion to accident forecast of large transformer in hydropower station. The main idea is to diagnose a facility by collecting its disparate classes of information. Here, three classes of sensors are used to collect temperature, gas and load. We assume that only one sensor in each class and the observations from every sensor are independent. Data has been processed with fuzzy rules before they are sent to the fusion center. The fusion center will compare these data with the transcendental knowledge and then make a decision. It will give a real-time, complete evaluation of the possibility of the series accident. Then, it sends the decision to the performance system. Because there is no formulation to calculate with, a neural network is used and trained with groups of experience data until it becomes stable. The study is imbursed by national natural science fund and its production will be applied to Fengmang hydropower station of China.

The conclusion that this application can make the facility safer is confirmed by experiments given in the paper. It forecasts accidents accurately with fewer virtual alarms or damages.

Keywords-neural fusion, disparate classes of information, fuzzy

1.Introduction

In hydropower station, a very important thing is to give a real-time, all-around evaluation of a transformer. When the transformer is in the state of general running, both safety and danger do exist. General information such as temperature, load predicts the facility is in order, but they also tell the danger at the same time. In real-time monitoring and forecasting systems. In traditional methods these classes of information are often considered respectively. In fact, danger is often combined with several factors. Traditional methods give uncompleted decision. Virtual alarms or accidents happen sometimes. Human being experts can avert these errors successfully. They can fuse all real-time information and get external conclusions. There are two reasons: The first is that the human experts consider more than one factors; The second is their experience. They compare the situation with their transcendental knowledge.

If we want to evaluate the situation completely, multi-source of information is necessary. For example, many mercury-thermometers are used to get the whole temperature. They are the same kind of sensors: mercury-thermometers. In distance detection, maybe both vidicon and infared-ray instruments are used. They are

different kinds of sensors, but they collect the same class of information—distance. If we want to forecast accidents of a facility, maybe more than one class of information is needed. Maybe a group of sensors to detect temperature, another to detect load, another gas etc. This is called disparate classes of information collection. According to these classes of information, We can get more comprehensive status of a facility. The decision will be more accurate.

Fusion is one of the best methods to simulate human experts. The method here is to collect disparate classes of information firstly. After comparing with the transcendental knowledge coming from human experts. It will give a real-time, complete evaluation, tell the possibility that series accident will happen and then send the decision to the performance system. In this paper, the architecture using fuzzy process and neural fusion is given in section 2. Three classes of information: temperature , gas and load are collected from transformer. Fuzzy rules are given to process the information before it is sent to fusion center in section 3. Three levels BP network used as the fusion center is in section 4. Experiments in section 5. The conclusions in section 6. Reference follows section 6.

2.Architecture

The first problem is that we can't fuse these collected data directly because they represent disparate classes information. So they must be pretreated before they are sent into the fusion center; the second is how we make use of human experts' knowledge.

The architecture is given in Fig1. Three input variables collect the three classes of information continually. We set one sensor to collect temperature, another to collect gas and another load. We assume that the observations from every sensor are independent. T— temperature, G—gas, L—load are collected and processed with fuzzy rules respectively. The semifinished results T', G', L' represent their own modulus of danger. Danger is a fuzzy concept. It is the estimation which human experts give after comparing input data with their experience. So IF/THEN fuzzy rules are better here. T', G', L' belong to [0,1]. Two questions are solved in the proceeding: The first is that information is translated into danger modulus according human experts experience. The second is that their outputs can be accepted by neural network. There is no formula to calculate with, but we have groups of experience data instead. So neural network is more employable as the fusion center. Few input variables; Low complexity; Facility works stable; The output is 0 or 1. "0" represents "in order" and another is accident. The factors above decide problem can be considered as a simple clarification question. In most cases neural network will be convergent. The network is trained until it is stable. Then it is applied to fuse input data and make decisions in time.

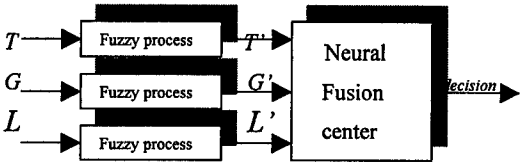
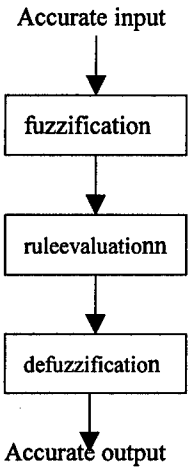


Fig 1. architecture

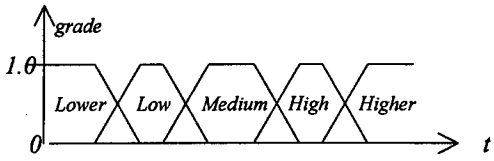
3.Fuzzy process

This process consists of next three steps.

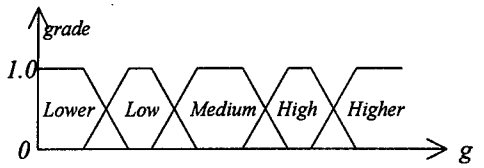


3.1 Fuzzification

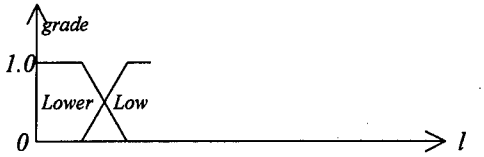
Five ranges are divided in order to describe Temperature, Gas, Load and Danger modulus with fuzzy value. Danger modulus is in [0,1]. Four membership functions: G1(t), G2(g), G3(l),G4(d), are defined in Fig 2.



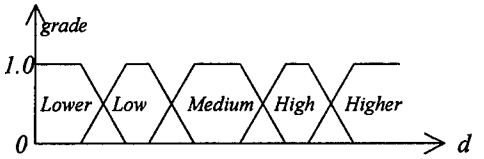
.(a) Membership functions of temperature



.(b) Membership functions of gas



.(c) Membership functions of load



(d) Membership functions of danger

Fig 2. Membership functions

3.2 Rule Evaluation

The most important step is rule evaluation. In this paper, Mamdani method is used to produce semifinished results.

X=T, G, L

If X is lower then D is smaller;... (1)

If X is low then D is small;... (2)

If X is middle then D is medium;... (3)

If X is high then D is big;... (4)

If X is higher then D is bigger;... (5)

For example X=x1; x1 belongs to both "lower" and "low".

a).For x1 belongs to "lower". A grade can be got, according to this grade and rule(1) and the shadow can be defined in Fig 3(b).

b).For x1 belongs to "low". A grade also can be got, according to this grade and rule(2) .The shadow can be defined in Fig 3(d).

c).Put two shadows into one reference frame and get a new shadow in Fig 3 (e).

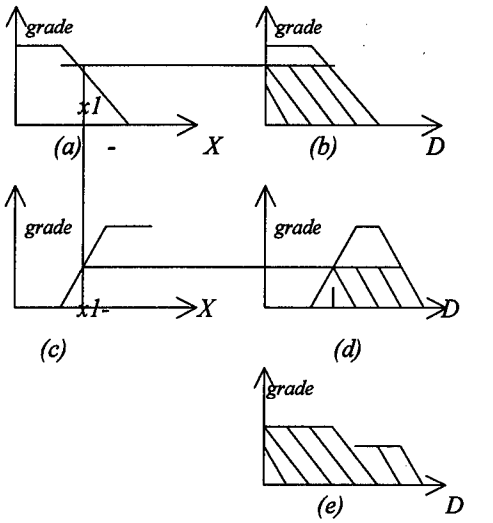


Fig. 3 Mamdani method sketch map

3.3 Defuzzification

The last step is defuzzification. It gives the accurate output. According to Mamdani method, the centroid of Fig 3 (e) is the output result. If we use G4(y) to expresses membership functions, y expresses danger modulus, and this location may be found with the following formula:

$$D^* = \frac{\int G4(y) * y \, dy}{\int G4(y) \, dy} \qquad (6)$$

4.Neural Fusion

Fusion center incepts three input variables, gives ---complete, real-time and accurate evaluation. It is well known that a forward neural network using BP algorithm solves some estimation or classification problems successfully. In this application, we use a three layers network. In the

input layer, there are three input variables which represent danger modulus of a transformer---T' for danger of temperature, G' for danger of gas and L' for danger of load. The number of middle layer's neural units can be changed in the process of learning in order to get a right result. It can be testified that a 3-layered BP neural network can simulate any continuous function's output if it has arbitrary number of units in the middle layer. In experiments, we use a network with five to nine units in the middle layer. In the output layer, only one unit exists. It can divide the input data into two groups. If output is close to zero, we can judge that the transformer is in order. If it is close to one, the transformer is in danger. All the neural units apply one form of state function like SIGMOID function. The following are structure of neural network Fig4 and formula of BP algorithm (7):

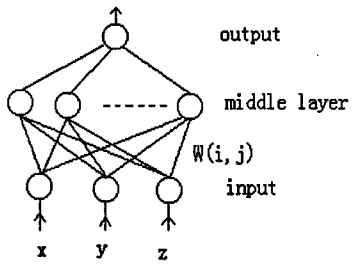


Fig4 Structure of the BP Neural Network

$$OUT_j^n = F(\sum_{i=0}^{N} W_{ij}^{n-1} \bullet a_i^{n-1} + B_j^n)$$

$$E = OUT^N(1-OUT^N)(D-OUT^N)$$

$$e_i^n = OUT_i^n(1-OUT_i^n)W_{ij}^n \bullet E$$

$$\Delta W_i^n = A \bullet OUT_i^n \bullet e_i^n \qquad (7)$$

$$\Delta B_i^n = A \bullet e_i^n$$

$$F(x) = \frac{1}{1+e^{-x}}$$

where OUT_j^n is output of jth unit of nth layer, E is error of the network, D is destination output, e_i^n is error of ith unit of nth layer, A is learning actor, B is bias of unit, F is state function.

Firstly, we must get groups of samples including input and output. Results are known. There are about tens or hundreds of samples for the neural network to learn. Secondly, after the network has finished learning, we input the other groups of samples into the network. If the outputs errors are between 0~+/-0.2, It can be deduced that the neural network can be used as a "judge" of a transformer.

5.Experiments

The method is applied to the hydropower station emulator. This emulator simulates two generator groups of Fengman hydropower station. This method is applied to the protection system of the 1# generator transformer which transforms power from 1# generator to power network. The protection monitoring system of 2# generator transformer has the same sensors , but when it makes decisions, it consider these classes information respectively, not synthetically. The system of 1# is trained with 70 groups (35 groups safety data, 35 groups danger data) experience data (Fig5) and neural network became convergent.

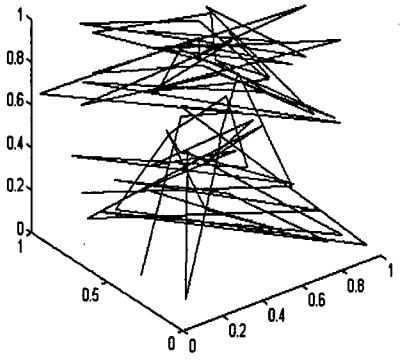


Fig5. Sample data space

The other 70 groups of data are tested. The errors of outputs are in Fig 6 :'+' :1-danger output; 'o': 0-in order output.

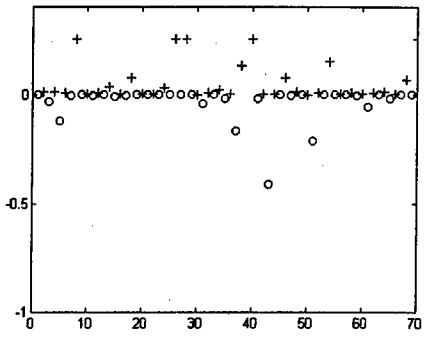
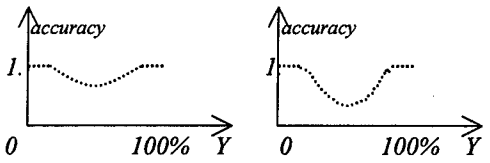


Fig 6 Outputs errors testing

After a period of testing, the accuracy of both systems is in Fig 7(Y is the danger that human experts gives).



(a)1#transformer *(b)2#transformer*

Fig 7 Accuracy sketch map

6.Conclusions.

In hydropower station, using neural fusion to forecast accident of generator transformer. The system applied this method evaluates situation more completely, more like human experts. Compared with the system which have the same sensors but give decision by consider every class of information respectively, in cases of great safe or great danger, both systems can give right evaluation. But if it is not very clear whether the accident happen or not. Neural fusion method is better. It gives fewer virtual alarms and fewer accidents happen.

For few input variables with one output, if the question is simple clarification, BP network is better. It is convergent and needs less training data. Fuzzy process is necessary as pretreament. The system will give real-time decisions.

Reference

[1]Soheil Shams "Neural Network Optimization for Multi-Target Multi-Sensor Passive Tracking" Proceeding of IEEE vol. 84.No.10.Oct 1996

[2]Hakan Delic "Fundamental Structures and Asymptotic Performance Criteria in Decentralized Binary Hypothesis Testing" IEEE Transactions on Communications. Vol. 43.No.2.Jan 1995.

[3] Liu Jinkun "A Study of blast furnace status forecasting system" Information and Control of China. Vol. 26, No.4 Aug. 1997.

Le Fan "Neural network and Neural Computer" Xian Traffic University Publishing Company of China. 1991

[4] Kong Yaohong "Data Fusion Theory and Application". Xian Electronic Science and Technology University Publishing Company of China 1997

[5] Dai Ruwei "Intelligent Control: Fundamentals and Application" National Defence Industry Publishing Company of China 1998

Session RB1
Management and Business Information Fusion
Chair: Nianyi Chen
Chinese Academy of Sciences, Shanghai, China

Aggregate Set-utility Fusion for Multi-Demand – Multi-Supply Systems

Erik P. Blasch
BEAR Consulting
2393 Fieldstone Cir, Fairborn, OH 45324
erik.blasch@sensors.wpafb.af.mil

Abstract – *Microeconomic theory develops demand and supply curves to determine the market equilibrium for commodity exchange. The demand and supply functions are the result of consumers utilities and producers production functions for different product combinations. The interaction is a game-theoretic approach to determine the quantity and prices with which goods and services are traded. Economic theory works with the long run equilibrium concept, yet with constant alteration of information, decisions are made in the short run. People, with changing preferences, shift their fused-aggregate utility function for a set of preferences rather than a single commodity. The paper investigates a set-utility function based on a fused perception of the dynamic changes of the corporate supply and consumer demand curves for various products.*

1. Introduction

The interaction between supply and demand policies of households and corporations is dependent on prices and quantities [1,2]. The interaction between these variables model market events such as the clearing price in exchange. Analyzing these policies is difficult when people's preferences vary in time, substitutes and competing goods change, and the value of money is altered by other markets. For instance, if price information is coming from a variety of sources, it might have different reported ranges dependent on the source. However, these price resolutions can be fused to form a composite set of information which allows a consumer or a producer to make decisions on how to determine a fair price based on how much a corporation wants to produce and how much a consumer wants to spend. A corporation uses prices to alter the exchange quantity of goods, which imparts changes to people's spending behavior.

The purpose of the paper is to address the different resolutions of measurement microeconomic data that drives corporation's production policies and is similar to the macroeconomic model from Blasch [3]. This paper is organized in the following fashion. Section 2 presents the economic model for demand and supply functions and discusses time-delay errors that corrupt these measurements. Section 3 presents the multiresolution technique for fusing, propagating, and updating measured price states that result from dynamic quantity changes in supply and demand. Section 4 formulates the problem and section 5 presents simulated results. Finally, Section 6 discusses some concluding remarks.

2. Microeconomic Model

Microeconomic theory seeks to model the economy as a function of demand and supply functions. The *Demand Function* (Q_d), is the relationship of quantity demanded to product prices and consumer income. The *Supply Function* (Q_S) is relationship of quantity supplied to production costs of wage rates and capital inputs [2]. The functional equilibrium determines the price of goods.

$$Q_d = Q_s \tag{1}$$

A dynamical equilibrium exists between prices and quantities and is cyclical between households and businesses through the goods and factors markets, shown in Figure 1.

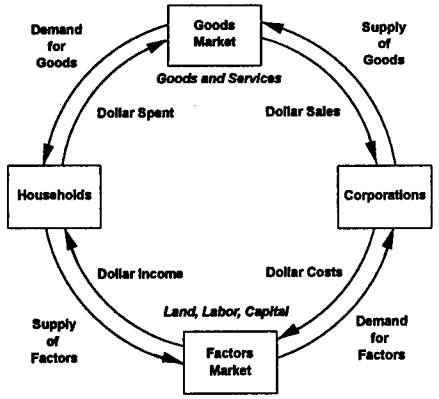


Figure 1. Exchange of Quantity and Prices [1].

The price-quantity model assumes that prices represent the value of goods. The goods market equilibrium shows a set of good's price and people's income from labor where:

$$Q_d = f(p_1, ..., p_n, m) \quad (2)$$

where m is the amount of consumers income [2]. From *utility theory*, income is the wealth constraint for quantity demanded,

$$m = Q_1 p_1 + ... + Q_n p_n \quad (3)$$

which shows that all decisions are based on a set of products a consumer can purchase.

Rearranging, we have:

$$Q_1 = \frac{m}{p_1} - ... - \frac{p_n}{p_1} Q_n \quad (4)$$

By determining the indifference point at which a consumer will equally value products, a marginal rate of substitution (MRS), (p_n/p_1), is determined for each good in a set.

A price consumption curve can be drawn for equilibrium sets of goods as one price changes, keeping other prices and income fixed. For each good, we can draw the inverse relationship between price and quantity demanded called the *law of demand* [1].

The quantity of goods supplied is a function of the corporation's production function. In the long run, all inputs are variable; however, in the short run, inputs of capital, K, and labor, L, are fixed. The relationship for capital and labor Q = f(K,L) is:

$$Q_s = \Delta L(MP_L) + \Delta K(MP_K) \quad (5)$$

where MP_L is the marginal product of labor, MP_K is the marginal product of capital, and $\Delta K/\Delta L = -MP_L/MP_K$ at zero output is termed the marginal rate of technical substitution (MRTS). The MRTS can also be related to the wage rate, w, and rental price of capital, r, by:

$$\frac{MP_L}{MP_K} = \frac{-\Delta K}{\Delta L} = \frac{w}{r} \quad (6)$$

Thus, we have a relation between the quantity supplied and the income consumers receive. The price equilibrium is shown as a relationship between quantity supplied and quantity demanded as shown in the Figure 2.

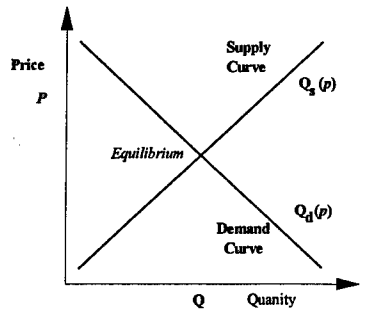


Figure 2. Price-Quanitity Equlibrium.

Using the models for quantity as a function of price, a state and measurement model is formed. Quantities and prices are variables and by inverting the demand function, we have 2 simultaneous equations:

$$\begin{bmatrix} \dot{Q}_{supply} \\ \dot{P}_{demand} \end{bmatrix} = \begin{bmatrix} C_s(wL,rk) \\ B_d(m,p_n) \end{bmatrix} \begin{bmatrix} Q \\ P \end{bmatrix} + \begin{bmatrix} w_s \\ w_d \end{bmatrix} \quad (7)$$

$$\begin{bmatrix} z_s \\ z_d \end{bmatrix} = [H_s \; H_d] \begin{bmatrix} Q \\ P \end{bmatrix} + \begin{bmatrix} v_s \\ v_p \end{bmatrix} \quad (8)$$

where $C(mL, rK)$ is the cost of the producer and $B(m, p)$ is the budget constraint of the consumer. By including uncertainty in the models, $v(t)$ and $w(t)$ are zero-mean mutually independent white Gaussian noise sequences with known covariances $Q(t)$ and $R(t)$, respectively.

The monitoring of economic variables is dependent on availability, time of measurement, and reporting confidence. If the reporting producer and consumer have time to fuse many perceived estimated values and quantities, the confidence is high, but requires delays in the updating of the information. The reporting time and confidence can be formulated as a multiresolution fusion problem, where multiple consumers and producers update knowledge of information at different time intervals.

3. Multi-Demand/Supply Relationships

The multiresolutional approach [4,5] propagates state values given sequential measurements. To develop the system equations for this approach, each point in time is expressed based upon the starting point of the block of time values. Figure 3 illustrates the

decomposition and fusion that is described by the following equations. The basic state equation is:

$$\underline{x}_{k+1} = A_k \underline{x}_k + B_k \underline{w}_k \quad (9)$$

which may represent multiple demand-supply functions.

The second time point in time, based upon the *current state*, is expressed by

$$\underline{x}_{k+2} = A_{k+1} \underline{x}_{k+1} + B_{k+1} \underline{w}_{k+1} \quad (10)$$
$$= A_{k+1} A_k \underline{x}_k + A_{k+1} B_k \underline{w}_k + B_{k+1} \underline{w}_{k+1}$$

The initial condition for the *first propagation time* state x and covariance P, a measure of uncertainty, may be expressed as

$$\hat{x}_{0|0}(k_N) = \begin{bmatrix} A \\ A \end{bmatrix} \underline{x}_0, \quad (11)$$

$$P_{0|0}(k_N) = \begin{bmatrix} A \\ A \end{bmatrix} \underline{P}_0 \begin{bmatrix} A \\ A \end{bmatrix}^T + B_0 Q_0 B_0^T, \quad (12)$$

where $B_0 = \begin{bmatrix} B & 0 \\ A & B \end{bmatrix}$, and

$$Q_0 = diag \left\{ \begin{bmatrix} Q(0) & 0 \\ 0 & Q(0) \end{bmatrix} \begin{bmatrix} Q(1) & 0 \\ 0 & Q(1) \end{bmatrix} \right\}$$

The equations for a *blocked-time system* may be written as:

$$\underline{x}_{m+1} = \bar{A}_m \underline{x}_m + \bar{B}_m \bar{w}_k, \quad (13)$$

where $\underline{x}_m = [\underline{x}_k^T, \underline{x}_{k+1}^T]^T$ and

$$A_m = diag\,[A_{k+1}, A_k]$$

Based upon the first observation, time k_4, the *estimate* is propagated at the highest resolution (N = 4):

$$\tilde{\underline{x}}(k_4) = A \hat{\underline{x}}, \quad (14)$$

$$\tilde{P}(k_4 + 1|k_4) = A_{k_4} P(k_4|k_4) A_{k_4}^T + B_{k_4} Q_{k_4} B_{k_4}^T \quad (15)$$

Using the measurement matrix:

$$z_{k_4} = H_{k_4} x(k_4) + v(k_4) \quad (16)$$

the *update* covariance is immediately computed:

$$x_{k_4+1|k_4+1} = \tilde{\underline{x}}(k_4) + K_{k_4}[z_{k_4} - H_{k_4} \tilde{\underline{x}}(k_4)]$$

$$P_{k_4+1|k_4+1} = [I - K_{k_4} H_{k_4}] \tilde{P}(k_4) \quad (17)$$

where K_{k_4} is the Kalman Gain.

Now, the generalized equations are derived using a wavelet approach to propagate Kalman-filtered updated states in time.

3.1 Discrete Wavelet Transform

For a given sequence of signals $x(i,n) \in L^2(Z)$, $n \in Z$ at resolution level i, the *lower* resolution signal can be derived by:

$$x(i-1, n) = \sum_k h(2n-k)x(i,k) \quad (18)$$

The *added detail* is given by:

$$y(i-1, n) = \sum_k g(2n-k)x(i,k) \quad (19)$$

The original signal $x(i,k)$ can be recovered from two filtered and sub-sampled signals $x(i-1, n)$ and $y(i-1, n)$.

$$x(i,n) = \sum_k h(2k-n)x(i-1,k) + \sum_k g(2k-n)y(i-1,k) \quad (20)$$

The *lowpass* filter $h(n)$ is the impulse response of a Quadrature Mirror Filter (QMF) and $g(n)$ and $h(n)$ form a conjugate mirror filter pair:

$$g(L-1-n) = (-1)^n h(n) \quad (21)$$

where L is the filter length. The derivation here is similar to [4], with implementation coming from [5] where the *Daubechies' Filter* [6] is used for processing information at various resolutions. A more rigorous approach of wavelet filters can be found in Strang [7].

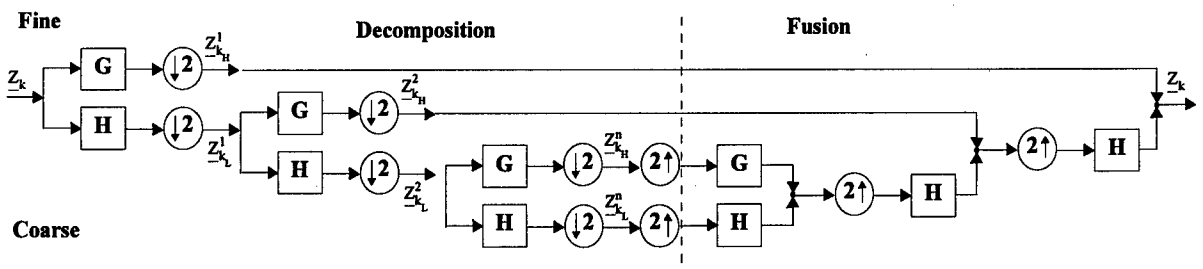


Figure 3. Control Flow for Distributed Multiresolutional Filtering.

Consider a *finite* sequence of *n*-dimensional *random vectors* at resolution level *i* with a length of a data-block:

$$\underline{X}(k_i) = [\underline{x}^T(k_i), \underline{x}^T(k_i + 1),...,\underline{x}^T(k_i + 2^{(i-1)} - 1)]^T \quad (22)$$

To change $\underline{X}(k_i)$ to the form required by the wavelet transform, a linear transformation is introduced [5]:

$$\underline{X}'(k_i) = \mathbf{L}_i \underline{X}(k_i) \quad (23)$$

where L_i is a matrix of 1's and 0's which transforms the data order, but not the magnitude of the data.

The wavelet transform vector form is:

$$\underline{X}(k_{i-1}) = \mathbf{L}_{i-1}^T \cdot diag\{\mathbf{H}_{i-1}, ...,\mathbf{H}_{i-1}\} \cdot \mathbf{L}_i \cdot \underline{X}(k_i) \quad (24)$$
$$\underline{Y}(k_{i-1}) = \mathbf{L}_{i-1}^T \cdot diag\{\mathbf{G}_{i-1}, ...,\mathbf{G}_{i-1}\} \cdot \mathbf{L}_i \cdot \underline{Y}(k_i)$$

where $\mathbf{H}_{i-1}$ and $\mathbf{G}_{i-1}$ are scaling and wavelet operators. Similarly, mapping from level $(i - 1)$ to level (i) can also be written as:

$$\underline{X}(k_i) = \mathbf{L}_i^T \cdot diag\{\mathbf{H}_{i-1}^T, ..., \mathbf{H}_{i-1}^T\} \cdot \mathbf{L}_{i-1} \cdot \underline{X}(k_{i-1}) + \mathbf{L}_i^T \cdot diag\{\mathbf{G}_{i-1}^T, ..., \mathbf{G}_{i-1}^T\} \cdot \mathbf{L}_{i-1} \cdot \underline{Y}(k_{i-1}) \quad (25)$$

Since G_{i-1} is a *highpass* filter operator and the sequence $\underline{X}(k_i)$ is a noise driven one, $\underline{Y}(k_{i-1})$ is a sequence of "noise-like" signals. However, the sequence $\underline{Y}(k_{i-1})$ is *not white* and is correlated with $\underline{X}(k_{i-1})$ - *lowpass* filtered.

3.2 Distributed Multiresolution Filtering

The equations for the distributed multiresolution filtering are presented in [8]. The general methodology is performed by:

1. Propagating from *m* to *m* +1, where m is the money price -quantity (*p*, *q*) value.
2. Transmit (*p*, *q*)-1 estimate to (*p*, *q*)-2 update
3. Perform (*p*, *q*)-2 measurement updates
3a. Transmit (*p*, *q*)-1 predicted values to (*p*, *q*)-4,
4. Transmit (*p*, *q*)-1 updates to the (*p*, *q*)-4 site
5. Estimate fusion of (*p*, *q*)-1 and (*p*, *q*)-4 results
6. Propagate the (*p*, *q*)-4 update

Note that there is a time multiresolution fusion of market data at the equilibrium points and a spatial fusion of demand and supply curves which is similar to a multirate-multiresolutional filtering problem.

4. Problem Formulation

The system being investigated is an market model with four prices and quantities for a product. Since each consumer/producer has only partial information about the market (due to the uncertainties of data collection), it is naturally desired that four sources of measurements, from four observations, be fused to achieve a higher confidence about the state of the market.

Since the prior information about the market is nearly linear, the *dynamics* are approximated by the linear relationships plus a modeling error given by:

$$x(k+1) = x(k) + 1.5y(k) + w_x(k) \ , w_x(k) \sim N(0,\sigma)$$
$$y(k+1) = x(k) - 1.5\ y(k) + w_y(k) \ , w_y(k) \sim N(0,\sigma)$$

or, $[\underline{x}_{k+1}] = [\Phi] \cdot [\underline{x}_k] + [\underline{w}_k], \underline{w}_k \sim N(0,\mathbf{Q})$

where *k* is time, the modeling error covariance matrix **Q** is given by **Q** = *diag* {10, 10}, and $w_x(k)$ and $w_y(k)$ are uncorrelated. The initial values are $\mathbf{P}_0 = [\mathbf{I}]$ and $\underline{x}_{k_4} = [300,0]^T$, assuming that the prices are in the range {200,400}. The *measuring process* for producers and consumers is described by the following measurement models, each of which is represented at their own timely and economic perspectives:

$$[\underline{z}_k^i] = [\mathbf{H}^i] \cdot [\underline{x}_k^i] + [\underline{v}_k^i], \underline{v}_k^i \sim N(0, \mathbf{R}^i) \quad (26)$$

where the measurement matrices $\mathbf{H}_i$, i = 1,...,4 are identity matrices and the measurement error covariance matrices $\mathbf{R}_i$ are $\mathbf{R}_1$= *diag*{10,10}, $\mathbf{R}_2$ = *diag*{20,20}, $\mathbf{R}_3$ = *diag*{30,30}, and $\mathbf{R}_4$ = *diag*{40,40},

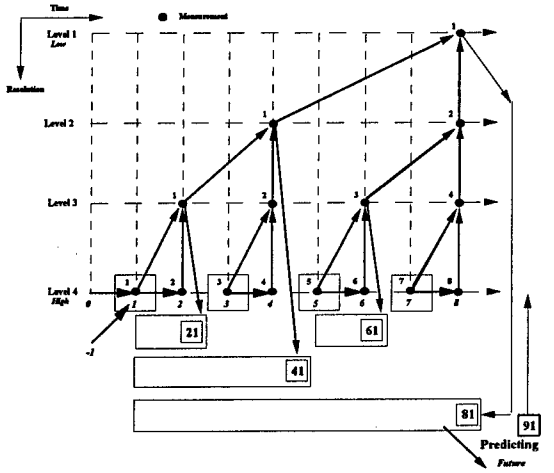


Figure 4. Real-time Multiresolution.

proportional to the resolution where timely measurements are assumed to have less accuracy than blocked updates.

Simulation runs are completed for 352 measurements, which is approximately a business year. The measurements are combined into 4, 2, and 1 measurements periods. Figure 4 shows how real-time values are propagated in time. Likewise, semi-real-time value updates are shown in Figure 5.

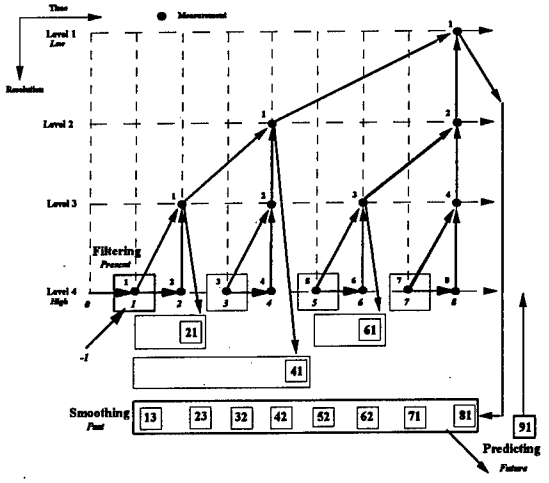


Figure 5. Semi-Real Time Multiresolution.

5. Simulation Results

A MATLAB program using the Daubechies' filter, the wavelet-multiresolution technique, simulates the multiresolution Kalman filter's performance for a set of time and quantity reports. Block time processing consists of measuring multirate states, processing the information at various resolutions, and fusing the results. In addition, the prediction function at the end of each time-block update predicts the time-associated next measurement. Level 4 is the real-time approach with 8 measurements used in the fusion process. Level 1, 2, and 3 are the semi-real-time approaches where measurements are processed, fused, and compared at various levels to the system (truth) model. Note that a *real-time* multiresolutional sensor fusion method is used to estimate the state equilibrium by fusing the information, sometimes from a single observer, since only the highest resolution is desired during the analysis.

5.1 Economic Measured Inputs

Input data is the result of measuring the market at different resolutions. Figures 5-8 show the four resolution of inputs, where it is assumed one

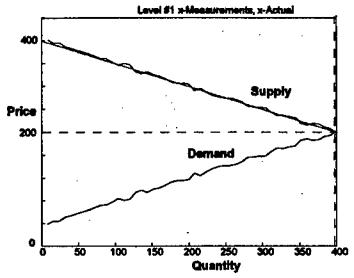


Figure 6. Spatio-Temporal Resolution Level 1.

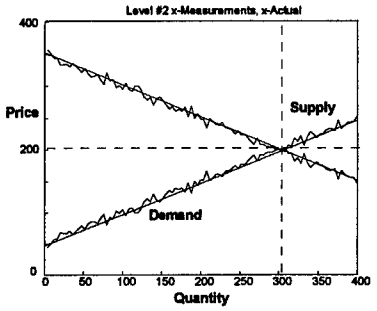


Figure 7. Spatio-Temporal Resolution Level 2.

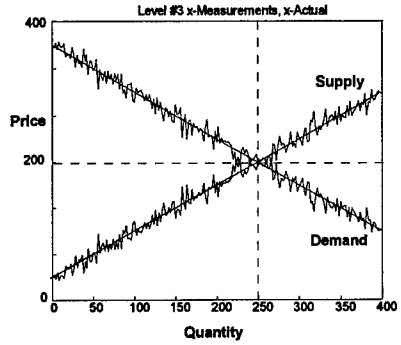


Figure 8. Spatio-Temporal Resolution Level 3.

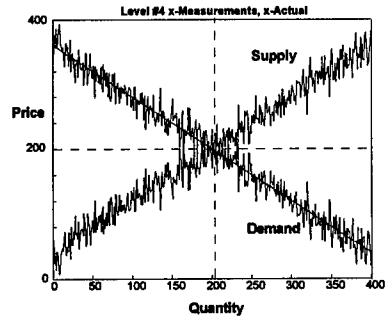


Figure 9. Spatio-Temporal Resolution Level 4.

demand/supply function update has highest resolution, but the largest variance.

5.2 Economic Estimated Outputs

For each set of value of a level corresponds to the consumer/producer resolution. By waiting, the consumer/producer would have a better estimated market value for variables in the demand/supply functions. From the fused result, we see that if we fuse curves and resolutions, we have a better estimate of aggregate prices and quantities.

6. Discussion and Conclusions

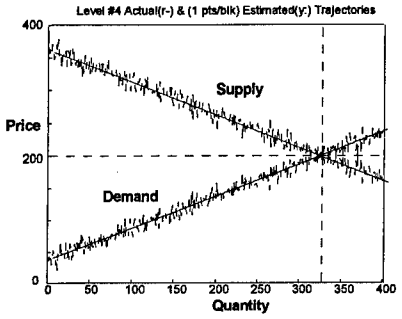


Figure 10. Fused Result at Highest Resolution.

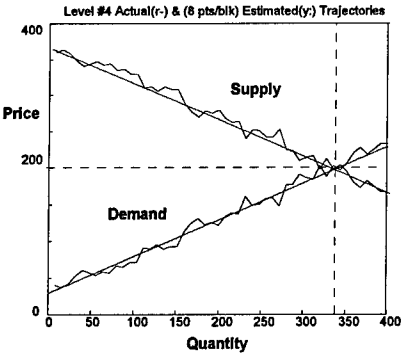


Figure 11. Fused Result at Coarsest Resolution.

The results show that estimation by the multiresolution technique allows for a variety of time fused updates dependent upon data variability and measurement confidence. Typically, measuring market data is the aggregate average perceived value. Since information available from different markets is reported at a variety of times, the methodology would be appropriate to incorporate data from a multiple set of observations. The difficulty with the analysis is that prices and quantities are typically observed values lagged in time, as shown in Figure 12, where the curve shifts to the right, but we only have the information from the measured curves, shown as dashed in Figure 12.

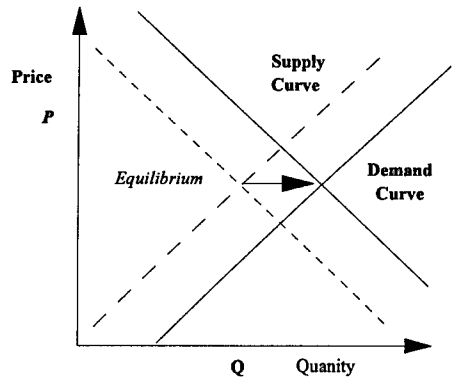


Figure 12. Policy Changes [1].

The multiresolution technique, used for sensor fusion models, is appropriate for assessing the time-delay updates associated with microeconomic system models. These results show that the model developed is applicable to updating consumers and producers with timely fused-estimates of variables for demand and supply functions.

REFERENCES

[1] R. Ruffin, *Intermediate Microeconomics*, Harper Collins, NY, 1992.

[2] R. Pindyck and D. Rubinfeld, *Microeconomics*, McMillian Publishing Co., NY, 1989.

[3] E. Blasch, "Decision Making in Multi-Fiscal and Mulit-Monetary Policy Measurements," *Fusion98,* Las Vegas, NV, pg. 285-292, 1998.

[4] L. Hong, "Multiresolutional Distributed Filtering," *IEEE Transactions on Automatic Control*, 9, April 1994, pp. 853 - 856.

[5] L. Hong, "Adaptive Distributed Filtering in multi-coordinated systems," *IEEE Transactions on Aerospace Electronic Systems*, 27, pp. 715 - 724, July 1991.

[6] L. Daubechies, "Orthonormal bases of compact supported wavelets," *Com. Pure Appl. Math*, XLI, pp.909-996, 1988.

[7] G. Strang and T. Hguyen, *Wavelets and Filter Banks*, Ch. 6, Wellesley-Cambridge Press, Wellesley, MA, 1996.

[8] E. Blasch, "Multiresolutional Distributed Filtering and Prediction for Sensor Integration," *Proceedings of SPIE, Wavelet Applications*, Orlando, FL, pg. 158-167, April, 1998.

Knowledge Discovery Applied to Agriculture Economy Planning*

Bing-ru Yang and Shao-jun Huang
Information Engineering School
University of Science and Technology, Beijing, China, 100083
Email: bingru.yang@bj.col.com.cn

Abstract: The paper presents double-base cooperating mechanism by studying the knowledge discovery based on database (KDD) which changes the structure, running process and the mechanism of KDD. Then a new Knowledge discovery based on database is established as KDD*. Applied to agriculture economy planning, the KDD* provides scientific decision for instructing agricultural production.

Key Words: Knowledge Discovery, Agriculture Economy, Decision.

1. Introduction

In the agriculture research, management and its basic level department, a large amount of data, examples, knowledge and experiences have been accumulated. In the field of agricultural crop the data are not made full use. The accumulated data on seedling, soil, fertilization, water, harmful insect of all kinds of crops as well as weather and calamities are saved as archives. That is to say, the phenomenon of plentiful data and poor knowledge is more serious in agriculture than other. So the demands for knowledge discovery are more eager. If some new rules which are produced by dynamic changed factors can be found through finding interrelations of the factors from the plentiful data, examples, common experiences and knowledge, the economical and social benefits will be very great.

The agriculture is a large and complex system. The types of soil in the world are enormous. The kinds of crops are complex. The calamities of harmful insects appears frequently and their symptom changes constantly. The interrelations and its effects among fertilizer, water, density and weather haven't been recognized. This is also the same with in the livestock, birds, fish and forestry. The relative database and knowledge base are characterized as large, multi-dimension, dynamic, incomplete and uncertain.

In recent years the market information didn't flow smoothly in many places, especially the crop production planning isn't instructed by the large dynamic market information. It causes blindness in the production planning and great fluctuation in the price which greatly affect agricultural market economy. How to collect the information in realization and find valuable and regular knowledge so as to effectively forecast and take measures in time will play an important role to the agricultural production.

Knowledge acquisition is always regarded as a bottleneck in the realization of intellectual system. Knowledge discovery partly solved the problem of knowledge acquisition. At present the development in knowledge discovery is mainly the traditional knowledge discovery based on database (KDD). Some intellectual methods, such as fuzzy logic, neural network, rough set and chaotic theory, are used in the KDD. But the KDD lacks means used by existing knowledge which helps to focus. The hypotheses and rules produced by KDD are directly evaluated. They are set into the knowledge base if passing the evaluation. Then the following defects are formed: first, many meaningless hypotheses are produced. It increases the burdens of evaluation and check on consistency and redundancy. It is close in the process of knowledge discovery; second, the KDD mines according to the need and interest of person, which lacks creative thought of computer itself to mine heuristically and directly. third, at present there are many experimental verification and original system but few practical system and tools.

In accordance with the above question, we first present the double base cooperating mechanism which is used to make basic knowledge base limit and drive KDD. This will lead to an open system of KDD: KDD* which is based on double base cooperating mechanism. KDD* breaks through the closeness of KDD. It makes database cooperate with knowledge base through interruptive and heuristic coordinator to find new knowledge.

2. The Introduction of KDD*

★subsidized by the emphasis item of National Natural Science Fund (69835001)

2.1 General Frame of KDD*

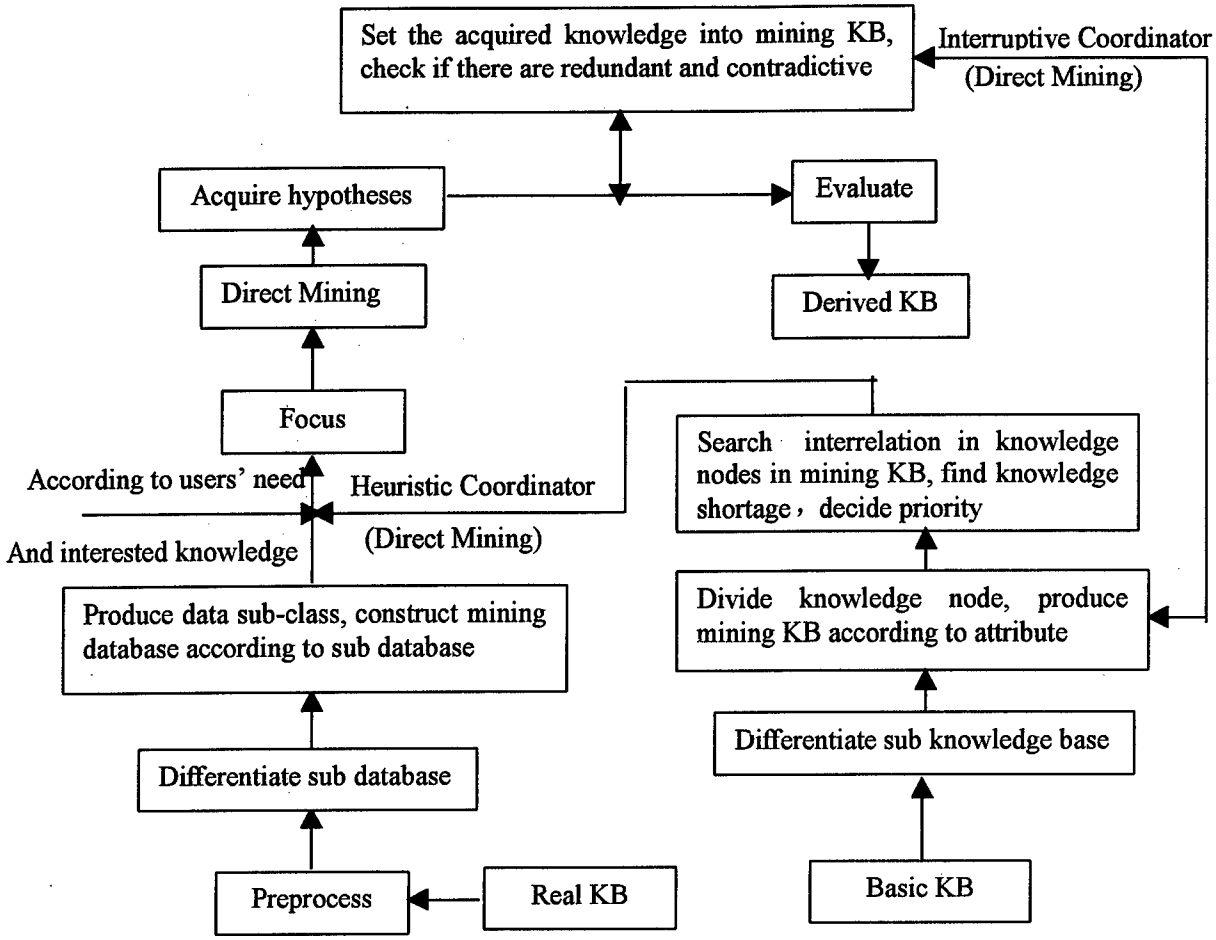


Fig.1 General Frame of KDD*

KB—Knowledge base

This figure shows the logic structure of the system and the relations between all the parts. From this figure we can see that the modules can be divided into the following parts:

Pre-processing: To process the original data by purifying the data, specific changing, etc. and create the DMDB which is used in the process of data mining and knowledge discovery.

Hypothesis rules: It is the core process of KDD. It uncommonly abstracts the hidden, unknown and potential valuable information in database which has the character of large amount of data, incompleteness, uncertainty, structure and causality

Focusing: namely to chose data from data mining. The main method in focusing is clustering analysis and detecting analysis. The method to direct the focusing are: (i) the expert, through man-machine interaction, inputs the knowledge in which he is interested and direct the direction of the data mining. (ii) Data directional mining by using heuristic coordinator.

sparseness. In the system the abstracted information is causality relation rule. Thus the basic knowledge base will be further improved. The mining methods that are used are statistics induction reasoning and qualitative reasoning. The former method will be discussed in 2.2.

Double base cooperating mechanism: to process the acquired rules by using interruptive coordinator and heuristic coordinator, and to exciting the data focusing for data mining by using relative strength. This will be discussed in 3.2.

Evaluation: this process is mainly used to evaluate the acquired rules in order to decide whether they will be stored into the derived knowledge base. The main methods are: (i) relative strength sets up a threshold value and be realized by computer; (ii) experts evaluate through man-machine interaction interface and also evaluates all kinds of figures and analysis materials provided by visual tools. Experts

evaluate mainly by using experiences and the relative strength of acquired rules. The rules are stored as new knowledge into the derivative knowledge base after passing evaluation.

2.2 The Reasoning Algorithm of Causality Statistics Induction

This algorithm uses incomplete induction approximate inference in statistics and credibility theory in uncertainty theory, by counting the examples in database and using the property with a large amount of examples as module, and gets a set of rules by credibility theory.

Possessed conditions: data focusing has been completed i.e. is ready to mine the two language variable A, B (e.g. the kinds of crop and its production). The mining process is as follows:

2.2.1 The Computer Decide the Relativity of the Corresponding Language Value through Statistic Analysis.

Divide A, B as $A(A_1, A_2, \ldots, A_m)$, $B(B_1, B_2, \ldots, B_n)$ according to their language values. If A and B are both single variable then we have $A(A_1, A_2, A_3, A_4, A_5)$, $B(B_1, B_2, B_3, B_4, B_5)$. Given A is the intersection of ml variables, $m=5^{ml}$. Given B is the intersection of nl variables, $n=5^{nl}$. Thus there are altogether $m \times n$ kinds of combination$\langle A_i, B_j \rangle$ i=1, 2....m j=1,2,...,n. To the possibility factor Pk=Cnk/N k=1,2,..., $m \times n$ corresponding to each computation, P=0.5 is the highest possibility. If Pk>0.5, $\langle A_i, B_j \rangle$is selected, otherwise it is eliminated and these two are considered to have no relativity.

2.2.2 Analyze A and B through Visual Tools

Experts can use visual tools, such as a distribution figure to decide the combination of the selected or eliminated areas. The areas here have one to one mapping relation with the language value mentioned above, i.e. the language value and the corresponding radius equals the corresponding area. The acquired area combination must be changed into corresponding language combination which is to be used in the later computation. Get the two highly relative properties e.g. A_i and B_j, and draw the corresponding values e.g. statistic value N, statistic value Cn(A_i, B_j)appearing both in Ai and Bj, statistic value Cn(A_i)appearing in A_i, and statistic value Cn(B_j)appearing in B_j to decide which variable have causal relation.

2.2.3 Get Weight of the Premise in the Hypothesis Rule ($A_i \rightarrow B_j$)

Given Ai is single premise, its weight is 1; given A_i is the interaction of many premises, i.e. rule R: $A_i \rightarrow B_j$ is:

$R: (P_1, p_1) \wedge (P_2, p_2) \wedge \ldots \wedge (P_n, p_n) \rightarrow (Q, q)$

Then the corresponding ri in (P_i, p_i) (Pi,pi) and (Q , q) can be gotten from the following formula. The weight in its rule can be gotten according ri.

$$r_i = \frac{m\sum_{j=1}^{m} q_j \cdot (p_{ij} - (p_i - q)) - (\sum_{j=1}^{m} (p_{ij} - (p_i - q)))(\sum_{j=1}^{m} q_j)}{\sqrt{m\sum_{j=1}^{m} (p_{ij} - (p_i - q))^2 - (\sum_{j=1}^{m} (p_{ij} - (p_i - q)))^2 (m\sum_{j=1}^{m} q_j^2 - (\sum_{j=1}^{m} q_j)^2)}}$$

Causality statistics induction reasoning algorithm flow is shown as following:

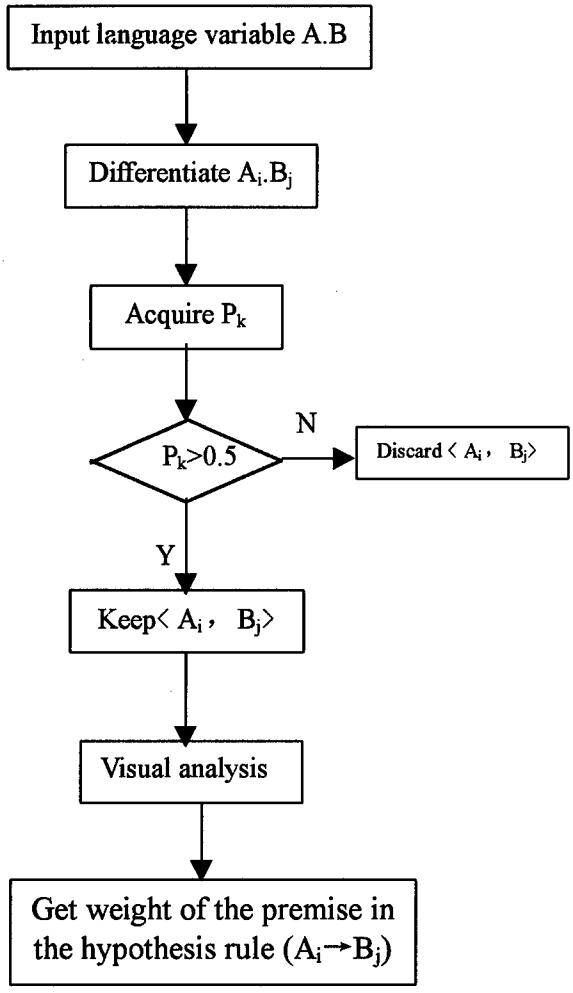


Fig.2 Causality Statistics Induction Reasoning Algorithm Flow

3. Double Bases Cooperating Mechanism:

3.1 Basic Theory

The technological realization of double-base cooperating mechanism is to construct interruptive and heuristic coordinators. To realize them there are some requirements: The large (basic) knowledge

base is divided into several correlative sub-knowledge bases according to each domain; Meanwhile, the real database is divided into correlative sub-databases according to each domain. Thus the layers between knowledge nodes in mining knowledge base and data sub-class (structure) in mining database make a one to one mapping. The basic theory which is proposed by us is pan-homotopy conception and the following structure mapping theorem: (Details can be found in reference [1][2])

Theorem (Structure Mapping Theorem): Aiming at X, in the sub-database corresponding to sub-knowledge nodes, <E, F> of knowledge nodes and <F, D> of data sub-class (structure) are identical pan-homotopic type spaces.

This theorem presents the mapping of layers between knowledge nodes in the sub-knowledge base and data sub-class in corresponding sub-database, shown in fig.3.

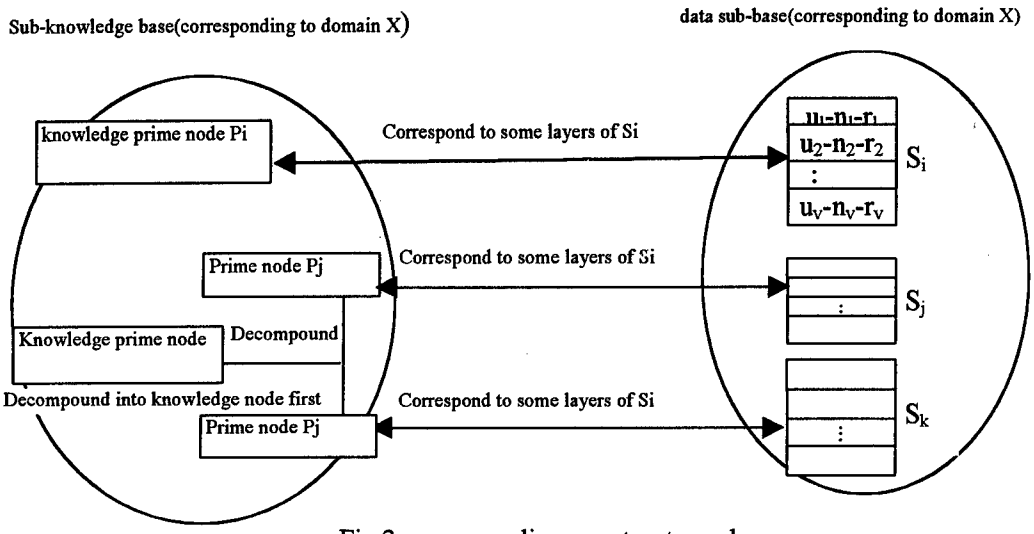


Fig.3 corresponding construct graph

On the basis of the research above, we can see that in the knowledge discovery system mathematical structure of database and knowledge base can be essentially come down to pan-homotopy category. Namely database is pan-homotopy category combined with data sub-type (structure) set and "mining path", which is called data mining category; and knowledge base is pan-homotopy category combined with knowledge nodes set and "reasoning arc", which is called knowledge reasoning category. Moreover some results about the isomorphy and restricting mechanism of knowledge reasoning category $C_R(E)$ in <E, F >and data mining category $C_D(F)$ in <F, D > are got, and "directional searching" and "directional mining process" are solved.

3.2 The Technological Realization

3.2.1 Interruptive Coordinator

The main function of the interruptive coordinator is, when the rules (knowledge) have been created from the focusing of the data in the real base, to "interrupt" the process of the KDD and to search whether there is a repetition of the created rule in the corresponding position of the knowledge base. If so, cancel this created rule and return to the beginning of the KDD. There need some special technology and methods to process contradiction. If not, continue the process of the KDD i.e. evaluate and store the result.

Because the interruptive coordinator is introduced into KDD, the inconsistent and redundant knowledge can be canceled earlier. Only those who are possibly accepted as new knowledge are evaluated and the evaluation work is greatly reduced. At the same time redundancy is processed in real time. This avoids complication of problem accumulated in a long time. In practical expert system, the amount of rules which finally become new knowledge are rather small compared with the original knowledge (it is difficult to find new knowledge), and a great number of rules are repetitive and redundant, so the introduction of interruptive type coordinator into KDD enhance the efficiency.

3.2.2 Heuristic Coordinator:

The function of heuristic coordinator

is to search irrelevant state of knowledge nodes in knowledge base under the principle of property on which knowledge base is established. Knowledge shortage is found. Data sub-class corresponding to real database uses heuristics and is activated to produce "directional mining process". To find the knowledge shortage in knowledge base especially in rule base, one of the methods is to compute the causality rule strength in each possible knowledge node in the whole causality network.

The causality rule strength consists of a group of three factors which can be expressed as

$$\pi(H, E)=<\alpha,\beta,\gamma> \quad \alpha=CF(E)*P(E)$$
$$\beta=CF(H, E) \quad \gamma=CF(H)*P(H)$$

Among it CF(E) is the reliability of premise, P(E)is pre-probability, CF(H, E) is the reliability of rule, CF(H)is the reliability of conclusion, P(H) is pre-probability. CF(E) and CF(H, E)are known. It consists of the whole random and fuzzy uncertain information of the rule. According to the causality rule strength the priority of directional mining can be determined and those can not be mined will be excluded.

4. Properties of KDD*

Compared with KDD, KDD* is a new structure of knowledge discovery which blends KDD and double base cooperating mechanism. It has the following characters:

1) KDD* organically make new knowledge found by KDD* communicate and merge with the knowledge in knowledge base and become one organism .

2) In the process of knowledge discovery, KDD* processes those redundant, repetitive and incompatible information in real time. This effectively decreases the complication of problem caused by a accumulated process. At the same time the preconditions are given for the merge and fusion of new and old knowledge.

3) KDD* changes and optimizes the process, structure and running mechanism of knowledge discovery.

4) From cognition KDD* strengthens and provides intellectual degree of knowledge discovery and enhances the cognition of computer itself. This is the direction for a long team.

5) Double base cooperating mechanism, the core technology of KDD*, shows the mapping between sub-knowledge base and data sub-class under a certain principle of establishing base. It provides a valid technology to decrease search space and improve mining efficiency.

5. Knowledge Discovery in Agricultural Economy Planning

In agriculture system there are abundant data which form all kinds of database such as relation database, time-spatial database, object-oriented database and multimedia database. But the data in these database are not made full use and hold plenty of storing space. Therefore it is necessary to mine.

In order to find knowledge from a database, it is necessary to process the database and establish corresponding basic knowledge base. Then All kinds of methods are used to mine the data in database. For example, Selenium (simplified as Se) is a necessary microelement for human and animal. It has many biological functions. Lack of Se is the main reason for many diseases, such as cataract, mastitis, cancer, large bone disease and so on. Rice is one of the main foods in the world. The content of Se is related to nutrition of Se in the human body. But most rice production areas are short of Se or have low content of Se. Therefore if we can find the dynamic changing rules under which rice sorbs Se, it will play an important role to instruct agriculture production and improve human health. Now there are some processed agricultural data which are shown in the following tables.

KDD* is applied to analyze the data in the table and finds that the accumulation of dry material isn't at the same speed with that of Se in the rice. The peak of former is in the middle of growing period, the latter in the late period. This is a rule that will be stored in Knowledge base. According to the rule we should fertilize Se again before the period of filling starch in rice. On the other hand rice has certain ability to sorb Se. So fertilizing Se in those areas that lack Se or have low content of Se can greatly enhance the content of Se in rice and improve its nutrition quality. Doing so on one hand can instruct us to fertilize reasonably, on the other hand can instruct manufacturer of fertilizer to add different microelement in different stage so as to meet the demand of agriculture production. Other data of agricultural crop can be treated so.

Table 1 Dry material accumulation of rice in the whole bearing period

Bearing period	Growing time (d)	Accumulate speed ($\mu g \cdot pot^{-1} \cdot d^{-1}$)	Stage accumulation ($\mu g \cdot pot^{-1}$)	Stage comparative accumulation (%)	accumulation ($\mu g \cdot pot^{-1}$)	comparative accumulation (%)
Seedling period	30	0.755	22.65	6.70	22.65	6.70
Spic period	60	5.564	166.93	49.38	189.58	56.09
Filling starch period	80	3.818	76.35	22.59	265.93	76.01
Ripe period	100	3.605	72.09	21.33	338.02	100

Table 2 Se accumulation of rice in the whole bearing period

Bearing period	Growing time (d)	Accumulate speed ($\mu g \cdot pot^{-1} \cdot d^{-1}$)	Stage accumulation ($\mu g \cdot pot^{-1}$)	Stage comparative accumulation (%)	accumulation ($\mu g \cdot pot^{-1}$)	comparative accumulation (%)
Seedlin g period	30	24.30	729.00	5.82	729.00	5.82
Spic period	60	142.02	4260.68	34.03	4989.68	39.85
Filling starch period	80	151.61	3320.19	24.22	8021.78	64.07
Ripe period	100	224.95	4498.99	35.93	12520.77	100

6. Conclusion

Agriculture production is an important thing to a country and its people. Reasonable planning for agriculture production will take great effect on a country. The article provides a new method of scientific decision for agriculture economy planning. It decreases the loss caused by planless production and will be instructive to the development of agriculture.

On the basis of KDD, double base cooperating mechanism can be applied to mine knowledge automatically and directionally. It can also process repetitive, contradiction and redundant rules. This will greatly improve the mining effeciency. The two kinds of coordinator can be independent system and install any existing KDD software to communicate with original knowledge base. It expands the function of original KDD greatly.

Reference

[1] Bingru Yang, KD(D&K) and Double-Bases Cooperating Mechanism, Journal of System Engineering and Electronics, Vol.10, No.1, 1999.

[2] Bingru Yang, Double-Base Cooperating Mechanism in KDD, International Symposium on Computer, 149-152 (1998).

[3] FIM and CASE for Evaluation of Hazavd level Based on Fuzzy language Field, Fuzzy Sets and System, North Holland, Vol.95, No.2, 83-89 (1997).

[4] S.S. Anand,D.A.Bell,J.G.Hughs,EDM:A General Framework for Data Mining Based on Evidence Theory,Data &Knowledge Eng., 18,189-223 (1996).

[5] R.H.Lchiang,etal.,A Framework for the Design and Evaluation of Reverse Engineering Methods for Relational Databases,Data &Knowledge Engineering,21,57-77 (1997).

[6] H.-L.ong and H.-Y.Lee,A New Visualization Technique for Knowledge Discovery in OLAP,Proceedings of the First Pacific-Asia Conference on Knowledge Discovery and Data Mining,279-286 (1997).

[7] C.Li and G.Biswas,Unsupervised Clustering With Mixed Numeric and Nominal Data-A New Similarity Based Agglomerative System,Proceedings of the First Pacific-Asia Conference on Knowledge Discovery and Data Mining,35-48 (1997).

Fusion of Neural Classifiers for Financial Market Prediction

Trish Keaton

Dept. of Electrical Engineering (136-93)
California Institute of Technology
Pasadena, CA 91125

Information Sciences Laboratory (RL 69)
HRL Laboratories, LLC
Malibu, CA 90265

ABSTRACT: *Forecasting financial currency markets is an extremely challenging problem because of the complex and highly chaotic nature of such markets. Motivated by the substantial profits that could be gained by having a system that could accurately predict large trends in the market, financial institutions are looking on advances in machine learning, neural networks, and statistics to provide them with another analysis tool. Researchers are investigating the use of back-propagation neural networks for financial time series prediction, due to their success on other pattern recognition problems such as machine & handwritten character recognition. However, to date their performance has been considerably lower than that achieved on the character recognition problem domain. This is due in large part to the tremendous amount of noise inherent in the data, which hinders the learning of good mapping functions. We believe that redundant forecasting through the synergistic use of multiple neural network predictors in combination with an intelligent decision aggregation scheme, may be the key to increasing the success rate of computer-aided forecasting systems. In this paper, we conduct an empirical and comparative study on the use of alternative methods for data preprocessing, fitness evaluation, and decision fusion. We demonstrate the advantage of our multiple classifier approach in predicting changes in the foreign exchange rate of the U.S. Dollar versus the German Mark over 250 days of trading.*

Keywords: financial market analysis, time series prediction, classifier fusion, evidential reasoning, neural networks

1. Introduction

Recently, the idea of combining multiple neural networks has become an area of great interest amongst pattern recognition researchers [3], [4], [5]. The rationale behind this current direction is that often real-world problems are far too complex for any single method to generate the best results for all possible types of inputs. Instead, an ensemble consisting of multiple models is learned, and then the classification decision is made by combining the classifications of the individual models. This approach has led to improved recognition rates over any of the individual constituents of the ensemble. However, the amount of improvement in accuracy has been found to be directly related to the "error independence" of the individual classifiers. Hence, this scheme has typically been applied to the fusion of complementary or orthogonal feature sets, such as strokes and cavities for character recognition, since classifiers based on very different feature sets often make errors in an uncorrelated manner.

In the future, we intend to explore the aggregation of forecasting models based on multiple feature sets such as wavelet and Fourier coefficients. However, in this paper we study the benefits of combining single layered feed-forward neural networks trained by back-propagation on an identical data set. In this case, network diversity was achieved by the inherent randomness associated with the back-propagation algorithm's initialization of a network's weights. Pattern classifiers trained in this manner can be viewed as approximations from different directions to the same goal, somewhat like reaching the peak of a mountain from different starting conditions. Hence, each classifier may behave differently with each individual input pattern, however in the long run their error rates will be nearly the same. Under these circumstances, the fusion of *redundant* classifiers

can potentially improve the overall performance of the system by reducing the uncertainty associated with the classification, just as in everyday life we often consult more than one expert before making an important decision.

This paper is organized as follows. In section 2, we provide an overview of our multi-classifier system for predicting financial markets. Section 3 discusses the design of the Artificial Neural Network (ANN) classifier used for time-series prediction, including the alternative cost or error functions utilized during the back-propagation learning of the network parameters. Section 4 describes the fusion methodologies explored for the aggregation of the prediction decisions. The latter sections present experimental results conducted on the U.S Dollar vs. German Mark financial currency data and the conclusions that may be drawn from this study.

2. The System

Since different pattern classifiers will exhibit different strengths and weaknesses, we propose a multiple neural-based classifier system for financial time-series prediction which contains an intelligent decision making scheme that fuses the predictions, such that each classifier's deficiencies are compensated for while preserving its strengths. A number of different strategies exist in combining classifier decisions. Two or more classifiers may be concatenated so that the output of one of them becomes the input to another, or they may be operated in parallel. We choose the later variant, where the group of classifiers to be combined can be viewed as a group of experts looking at the same problem from their individual points of view and stating their individual prediction about the future trend. The task performed by the decision module is to combine the predictions in a manner such that the overall uncertainty associated with the final decision is reduced. A block diagram of the proposed system is shown in Figure 1.

3. A Neural Network for Forecasting

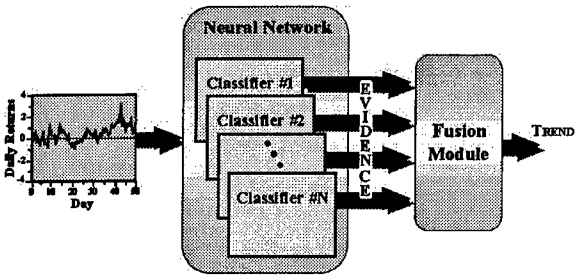


Figure 1. Fusion of multiple neural classifiers for improved financial market analysis.

The most successful Artificial Neural Network (ANN) to be applied to pattern recognition tasks is the standard fully connected multi-layered perceptron net. It learns a mapping between input and output pairs by adapting its weights through back-propagation learning algorithm. Figure 2 shows the one-layer architecture used in our financial time-series prediction experiments.

The model assumes there exists an underlying complex relationship between the current return and the prior returns over a twenty-day period, hence the input layer contains 20 neurons. Generally, a single output neuron having a nonlinear hyperbolic tangent activation function is used to produce values within the range of [-1,1], where its sign indicates the direction of change in the market [1], [2]. However, in our multiple classifier system we would like to be able to interpret the network outputs as Bayesian *a posteriori* probability estimates, which can then be easily combined using evidential reasoning methods. Richard and Lippmann [6] showed that Bayesian probabilities are estimated when the desired network outputs are 1 of M classes (one output is unity for the correct class, all others are zero), and the network is trained by minimizing the expected mean square error (MSE) or the cross-entropy cost function. Thus, we utilize two neurons in the output layer having a sigmoidal activation function [0-1]. One of the output neurons is used to indicate the prediction of an upward trend, while the other indicates a downward trend in the market as shown in Figure 2.

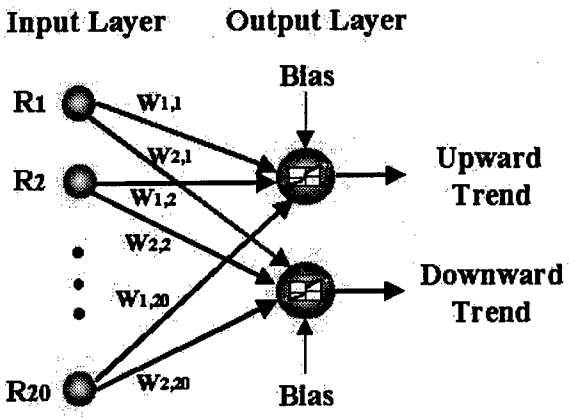


Figure 2. A single-layered perceptron used for financial forecasting.

The network parameters were optimized using back-propagation learning with the following variants. First, the weight update rule was modified to include a momentum term α, which along with the learning rate η were adapted during training from their initial values of 0.9, and 0.1, respectively. At each training epoch, the training patterns were presented sequentially to the network for weight updating. In addition, we incorporated the oddness symmetry "hint" into the learning process by presenting each training instance followed by its negation of both the input vector and its target value. In [2], it was shown that the symmetry "hint" improves the generalization ability of the network, by preventing overfitting, and by restricting the number of solutions the network may settle into. Finally, we experimented with using both the squared-error and the cross-entropy cost functions for optimizing the network weights.

3.1 Mean Square Error

The traditional mean square error function is the most popular cost function used in the majority of applications for optimizing the weights of a neural network. It has demonstrated good performance on real-world problems, and can be used for prediction, classification and regression. It assumes that the network's "error" will be normally distributed about the predicted values. In situations in which this is true, the Gaussian probability distribution is appropriate and the error term to be minimized is the mean square error. The error function and its derivative used for weight update are:

$$E_p = \sum_k \frac{(t_k - y_k)^2}{2\sigma^2}$$

$$\frac{\partial E_p}{\partial y_k} = \frac{(t_k - y_k)}{\sigma^2}$$

where, σ^2 is typically fixed to be 1 and t_k is the k^{th} neuron's target value and y_k its output. When the Bayesian *a posteriori* probabilities are estimated correctly, the classification error rate will be minimized, and the outputs sum to one such that they can be interpreted as probabilities.

3.2 Cross-Entropy Error

Another popular cost function measures the cross-entropy between actual outputs and desired outputs, which are treated as Bayesian probabilities. Motivation for its use lies in the assumption that the desired outputs are independent, binary, random variables, such that the network's "error" will be binomially distributed. Therefore, given binary target values of 0 and 1, we can write the learning objective in terms of the relative or cross-entropy of the target value to the actual output of the network. The minimizing error function becomes:

$$E_p = -\sum_k \left[t_k \log(y_k) + (1 - t_k) \log(1 - y_k) \right]$$

with its derivative easily expressed as:

$$\frac{\partial Ep}{\partial y_k} = -\left(\frac{t_k}{y_k} - \frac{1 - t_k}{1 - y_k} \right)$$

where, t_k is the k^{th} neuron's target value and y_k its output. This cost function can be interpreted as minimizing the Kullback-Liebler probability distance measure. It weights errors more heavily than the squared-error term, and thus the trained

network tends do a better job of predicting large changes in the market at the expense of misclassifying smaller variations. For our application this bias is desirable since a failure to detect large shifts in the market is far more costly than failing to detect smaller movements.

4. Combining Classifier Predictions

The fusion methodologies investigated for combining the predictions of the different neural classifiers ranged from simple techniques, requiring little computation such as majority voting and averaging, to the more computationally intensive evidential reasoning techniques: Bayesian, Dempster-Shafer's rule, and fuzzy integral fusion. In this section, we describe each method of combination employed in generating the final prediction decision.

4.1 Majority Vote

This scheme tallies the classification votes from all networks, then chooses the prediction yielding the maximum number or that which was indicated by the majority (e.g., at least 3 out of 5 classifiers).

4.2 Arithmetic Mean

In this combination scheme, we simply average the individual classifier outputs. The maximum of the averaged values is chosen as the correct prediction class.

$$\bar{y}_c = \frac{1}{n}\sum_i y_{c,i}$$

where, n is the number of classifiers.

4.3 Bayesian Evidential Reasoning

The Bayesian evidential reasoning technique is strongly founded upon the framework of probability theory, however the underlying assumptions it requires for the propagation of beliefs may or may not be true in practical situations. For example, Bayesian reasoning assumes that the pieces of evidence E_i to be aggregated are statistically independent. This assumption may not be true in cases where causal or contextual relationships exist, however for the purposes of fusing multiple neural forecasters we will assume that the evidence sources are "independent" with respect to the errors they make. Figure 3. shows the information fusion process under an evidential reasoning framework.

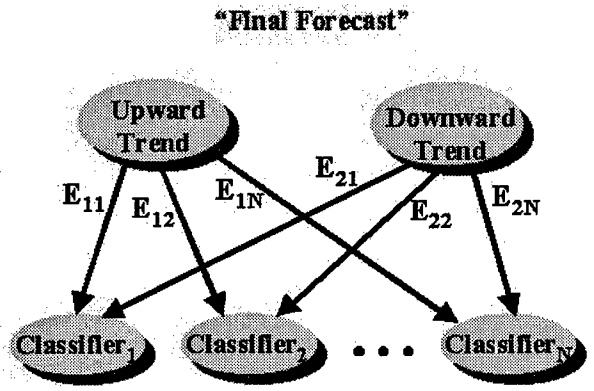


Figure 3. Information fusion through Bayesian evidential reasoning.

Bayesian theory uses an "Odds-Likelihood Ratio" formulation of Bayes' rule to aggregate the evidence from multiple sources. The *a priori* odds O(H) of a given class hypothesis H (e.g., upward trend, downward trend) is related to its *a priori* probability P(H) by the following relations:

$$O(H) = \frac{P(H)}{P(\sim H)} \quad \text{and} \quad P(H) = \frac{O(H)}{1+O(H)}$$

where ~H means "not H". The likelihood of the evidence E_i, given that the hypothesis H is true, is:

$$L(E_i \mid H) = \frac{P(E_i \mid H)}{P(E_i \mid \sim H)}$$

The class probabilities for each hypothesis may be estimated from training data, and the neural network outputs divided by these probabilities to produce scaled likelihoods, where the scaling factor is the reciprocal of the unconditional input probability.

The formula for updating the *a posteriori* odds of a hypothesis H, given the evidence E_i observed is:

$$O(H \mid E_1, E_2, ..., E_n) = O(H) \prod_{i=1}^{n} L(E_i \mid H) \cdot$$

and, the "belief" or *a posterior* probability for a hypothesis is simply:

$$P(H \mid E_1, E_2, ..., E_n) = \frac{O(H \mid E_1, E_2, ..., E_n)}{1 + O(H \mid E_1, E_2, ..., E_n)} \cdot$$

The final prediction is chosen to be that hypothesis H having the greatest probability given the accumulated evidence.

4.4 Dempster-Shafer's Evidential Reasoning

Dempster-Shafer's (DS) theory of evidence is another tool for representing and combining evidence, which is considered to be a generalization of Bayesian theory. It is more flexible than Bayesian when our knowledge is incomplete, by permitting the assignment of an ignorance term rather than forcing an over-commitment towards "belief" or "disbelief" in a hypothesis. Rather than representing the probability of a hypothesis H by a single value P(H), DS theory binds the probability to a subinterval [Bel(H),Pl(H)] of the interval [0,1], where Bel(H) - "belief" and Pl(H) - "plausibility" represent the lower and upper bounds on the probability, such that:

$$Bel(H) \leq P(H) \leq Pl(H).$$

When Bel(H)=Pl(H), Dempster-Shafer theory reduces to Bayesian.

According to D-S theory, the set of all possible outcomes (i.e., the sample space) in a random experiment is called the frame of discernment (FOD) denoted by Θ. For our problem, the frame of discernment would be Θ={upward trend, downward trend}. Associated with each of the neural network classifiers is a basic probability assignment (bpa), which expresses the degree to which the evidence confirms or supports a hypothesis. It is assigned according to the neural network output y_k, and is estimated from the statistics of the training set.

Given two bpa's $m_1(\bullet)$ and $m_2(\bullet)$ discerned in the same frame, their combined belief in a hypothesis H can be computed using Dempster's rule of combination:

$$m_{1\oplus 2}(H) = \frac{\sum_{B \cap C = H} m_1(B) m_2(C)}{1 - \sum_{B \cap C = \varnothing} m_1(B) m_2(C)}$$

This rule is applied recursively until the evidence from all n sources is aggregated. The output of the D-S fusion module is the following interval of belief:

$$\text{Belief(H)} = m_{1\oplus 2\oplus 3\oplus \ldots \oplus n}(H)$$

$$\text{Plausibility(H)} = 1 - \text{Belief(}\sim H\text{)}$$

$$\text{Belief Interval} = [\text{Bel(H)}, \text{Pl(H)}].$$

Then, the final prediction hypothesis having the largest amount of "support" with the smallest uncertainty (i.e., the difference Pl(H)-Bel(H)) is chosen.

4.5 Fuzzy Integral Fusion

The fuzzy evidential reasoning scheme views the outputs of multiple networks or experts as independent sources of "*objective*" or "*observed*" evidence, which is combined with an evaluation of the "*relevance*" or "*importance*" of that evidence with respect to each hypothesis. The combination of both types of information is accomplished using a fuzzy aggregation operator called the fuzzy integral. The fuzzy integral is a

nonlinear function defined with respect to a fuzzy measure, which is a generalization of a probability measure, that replaces the additivity property (i.e. P(A) + P(~A) = 1) with a weaker monotonicity condition. The "*relevance*" of an information source is captured by this fuzzy measure or density, which may be subjectively assigned by a human, or estimated from the training data. The fuzzy integral operator integrates the outputs of the neural network experts with respect to this aggregated relevance function to compute the possibility expectation between the pooled evidence and it's combined relevance. As shown in Figure 4, the fuzzy integration or possibility expectation may be interpreted as searching for the maximal agreement between the actual evidence, and its aggregated relevance.

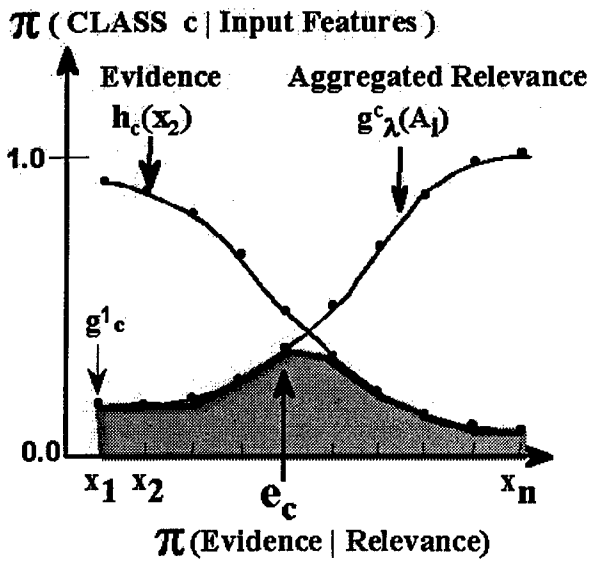


Figure 4. Fuzzy integration.

Algorithm:

For all c classes or hypotheses {

1) Sort classifier evidence:
$h_c(x_1) \geq h_c(x_2) \geq \ldots \geq h_c(x_n)$,
$A_i = \{x_1, x_2, \ldots x_i\}$
2) Find lambda parameter:
$$1+\lambda_c = \prod_{i=1}^{n}(1+\lambda_c g_c^i) \quad \lambda_c = (-1,+\infty)$$
3) Compute aggregated relevance:
$g_{\lambda_c}(A_1) = g_c^1$
$g_{\lambda_c}(A_i) = g_c^i + g_{\lambda_c}(A_{i-1}) + \lambda_c g_c^i g_{\lambda_c}(A_{i-1})$
4) Compute possibility expectation:
$$e_c = \oint_x h_c(x) \circ g_{\lambda_c}$$
$$= \underset{i=1}{\overset{n}{MAX}}\{MIN[h_c(x_i), g_{\lambda_c}(A_i)]\}$$

}
Compute final classification decision:
$$Class = \underset{c=1}{\overset{\#Classes}{MAX}}(e_c).$$

Separate aggregation networks are needed for fusing information regarding each hypothesis. The final prediction classification or hypothesis decision is taken to be the one returning the largest fuzzy integral value as shown in Figure 4.

5. Experimental Results

The data used to evaluate our system consisted of the closing prices of the U.S. Dollar versus the German Mark currency exchange rate over a four-year period. The prices (P_t) were normalized to compute the daily return (R_t) using the following formula:

$$R_t = \left(\frac{P_t - P_{t-1}}{P_{t-1}}\right) \times 100\%$$

A plot of the computed daily returns is shown in Figure 5. This normalized data was then divided into three sets with the first 500 samples used to train the neural networks, and the remaining samples divided into two test sets. We chose to combine five neural networks each trained in the same manner, although due to random weight initialization, each network started at a different point in the error surface. Table 1 presents the prediction hit rate results for the neural networks trained using the squared-error function and the "oddness" symmetry hint. The performance on the second test set is lower due to the fact that the training data is not representative of the current market status, but instead is "out-of-date". Table 2 presents the prediction hit rate results for the

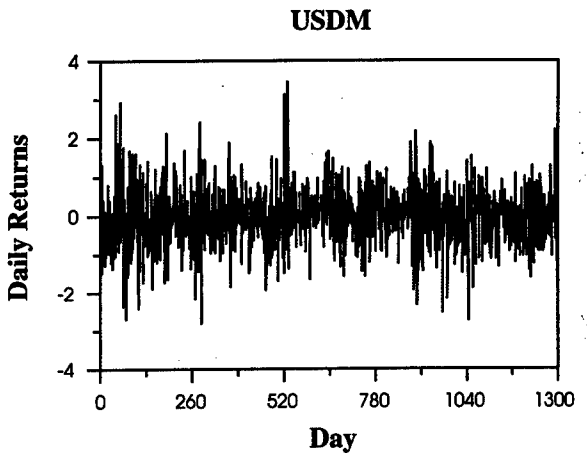


Figure 5. U.S. Dollar vs. German Mark exchange rates.

neural networks trained using the cross-entropy cost function and the "oddness" symmetry hint. The results are lower because this cost function tends to predict only large trends in the market at the expense of smaller variations.

Having trained our neural-based forecasters, the goal is to combine the outputs from the individual networks to obtain an overall prediction of the market trend. Five different fusion methodologies were implemented and tested. Table 3 & 4 present the prediction hit rates on the two different test sets obtained by combining the five neural networks trained using the MSE error and the "oddness" symmetry hint. By examining the results, we see that there is a clear benefit in using evidential reasoning methods for fusing the individual network predictions. We obtained nearly a 15% increase in performance over the predictions of the individual neural classifiers

6. Conclusions

We introduced a multiple "redundant" neural-based classifier system for financial market analysis. Each classifier was trained in an identical manner, however random weight initialization provided some network diversity. The objective functions used to optimize the network weights produced estimates of the Bayesian *a posterior* probabilities, which can easily be converted to scaled likelihoods, and then combined for higher level decision making. An intelligent information fusion scheme was used to combine the predictions of the individual classifiers, such that the accuracy and reliability of the final prediction was improved. We obtained nearly a 15% increase in performance over the predictions of the individual neural classifiers. In the future, we intend to investigate the generation of complementary forecasters through the use of time-frequency transformations.

References

[1] Tan, C., and Wittig, G. , "A Study of the Parameters of a Back-propagation Stock Price Prediction Model", in Proc. ANNES, 1993.

[2] Abu-Mostafa, Y.S., "Financial Market Applications of Learning From Hints", *Neural Networks in the Capital Markets*, Wiley, 1994.

[3] Jacobs, R., and Jordan, M., Nowlan, S., and Hinton, G., "Adaptive Mixtures of Local Experts", *Neural Computation*, 3:79-87, 1991.

[4] Xu, L., Krzyzak, A., and Suen, C., "Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition", IEEE Trans. Syst. Man Cybern., 22: 688-704, 1992.

[5] Benediktsson, J., and Swain, P., "Consensus Theoretic Classification Methods", IEEE Trans. Syst. Man Cybern., 22:418-435, 1992.

[6] Richard, M., and Lippmann, R., "Neural Network Classifiers Estimate Bayesian A Posteriori Probabilities", *Neural Computation*, 3:461-483, 1991.

[7] Chauvin, Y., and Rumelhart, D.E., *Back-propagation: Theory, Architectures, and Applications*, Erlbaum, 1995.

[8] Masters, T., Signal and Image Processing with Neural Networks: A C++ Sourcebook, pp.38-55,Wiley, 1994.

[9] Abidi, M., and Gonzalez, R., *Data Fusion in Robotics and Machine Intelligence,* Academic Press, 1992.

[10] Deboeck, G.J. (editor), *Trading on the Edge: Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets*, Wiley, 1994.

Table-1 Performance of networks trained using the MSE and "oddness" symmetry hint .

	% IN Samples Total Samples = 500 (1 ... 500 days)		% OUT Samples Total Samples = 253 (501 ... 753 days)		% OUT Samples Total Samples = 250 (754 ... 1003 days)	
Classifiers	**Correct**	**Error**	**Correct**	**Error**	**Correct**	**Error**
1	70.00	30.00	60.10	39.90	53.80	46.20
2	71.10	28.90	62.80	37.20	54.60	45.40
3	72.39	27.61	57.80	42.20	52.40	47.60
4	71.60	28.40	58.27	41.73	54.40	45.60
5	71.20	28.80	61.30	38.70	55.00	45.00

Table-2 Performance of networks trained using the Cross-Entropy error & "oddness" hint.

	% IN Samples Total Samples = 500 (1 ... 500 days)		% OUT Samples Total Samples = 253 (501 ... 753 days)		% OUT Samples Total Samples = 250 (754 ... 1003 days)	
Classifiers	**Correct**	**Error**	**Correct**	**Error**	**Correct**	**Error**
1	66.00	34.00	48.20	51.80	43.20	56.80
2	68.30	31.70	51.00	49.00	39.60	60.40
3	64.18	35.82	52.80	47.20	41.80	58.20
4	66.20	33.80	55.05	44.95	44.00	56.00
5	63.30	36.70	52.35	47.65	43.10	56.90

Table-3 Performance combining five networks trained using MSE error and "oddness" symmetry hint over the first test set consisting of (501 ... 753 days).

	Majority Vote	Arithmetic Mean	Bayesian Reasoning	Dempster-Shafer	Fuzzy Integral
Correct	56.80	54.50	65.80	60.90	64.00
Error	43.20	45.50	34.20	39.10	36.00

Table-4 Performance combining five networks trained using MSE error and "oddness" symmetry hint over the second test set consisting of (754 ... 1003 days).

	Majority Vote	Arithmetic Mean	Bayesian Reasoning	Dempster-Shafer	Fuzzy Integral
Correct	42.40	40.80	58.80	55.90	61.00
Error	57.60	59.20	41.20	44.10	39.00

Quantifying Operational Risk using Bayesian Networks and Monte Carlo Simulations

Prabhat Ojha

Vishrut Jain
School of Computing
National University of Singapore
Singapore - 119260

Christopher Marshall

Abstract

Operational Risk Management imposes a structured approach to dealing with potential losses in complex operational processes and resources in financial firms. Two unique properties of operational risk (as opposed to financial risk) make application of Bayesian Networks (BNs) attractive to this domain. Firstly, in the absence of mark-to-market assets, operational risk measurement requires integration of various data sources and expert judgements about risk. The ability of BNs to structure subjective beliefs and learn interactively from data is attractive in this regard. Secondly in absence of liquid markets where risks can be diversified away, operational risk needs to be internally actionable. The ability of BNs to structure conditional relationships between risk factors and draw probabilistic inferences and decision support is attractive in this regard. Monte Carlo (MC) simulation over a BN provides powerful capabilities for deriving more meaningful loss distributions rather than point probability estimates. A framework combining these two methodologies provide a way to both measure and manage operational risk in an integrated way. We have implemented a prototype system of the framework. Preliminary results demonstrate the practical promise of the framework.

1.Introduction

In recent years high profile losses in investment banks due to poor organizational design and trader fraud, like the Barings disaster [1, 2] or disasters due to inadequacies in information systems, like the Joseph Jett [3] case have focussed the attention of managers on operational risk. There is no universally accepted definition of operational risk, suggesting infancy of the field. Board of Governors of the Federal Reserve System Trading Activities Manual defines operational and systems risks as the "risk of human error or fraud, or that systems will fail to adequately record, monitor and account for transactions or positions." The Basle committee 1994 Risk Management Guidelines (Vol. 16) for OTC derivatives adopted a definition that has been used by a number of banks, which holds that Operational Risk is "Risk that deficiencies in information systems or internal controls will result in unexpected losses. This risk is associated with human error, systems failures and inadequate procedures and controls."

Practitioners agree that operational risk is not confined to back-office or "operations risk" but encompasses front-office operations or virtually any business process in the bank and includes elements of settlement risks, business interruptions risk and administrative or legal risks.

The absence of a liquid market for operational risks means that it needs to be measured using internal data. Since the ultimate aim of measurement is management of operational risks, the larger methodology should also be able to structure operational risks and provide capabilities for decision making for reduction of these risks. First generation operational risk measurement methodologies concentrate on measuring an aggregate value for operational risk for the purpose of capital allocation using cost, income or price volatility based models. These models are easy to implement as they are based on available accounting or market price data. However, such quick and highly aggregated risk numbers do not lead to actionable recommendations. The need to discover sources of operational risks in order to redesign processes or controls has lead to more advanced operational risk models. These models are not just focussed towards setting an aggregate capital value, but also towards discovering sources of operational risk, understanding of loss events and their relationships and discovering the effect of process and control redesign alternatives on these risks.

Most operational risks are a result of a complex sequence of related events, where the events themselves are uncertain. Usually we are dealing with low-probability/high impact events for which data is scarce by definition. However, operational managers can generally articulate their beliefs about the probabilities and impacts associated with these events. BNs and MC simulation when used together can allow us to structure these beliefs about event probability and conditional dependencies in a systematic framework. This paper describes one such framework used for

measuring risks associated with securities settlement process in the Singapore operations of a multinational investment bank.

In section 2 we briefly review literature on BNs and MC simulations as applied to BNs. In section 3 we choose a subset of operational risks associated with a securities settlements process in a bank and illustrate the use of BNs and MC simulations in quantifying these risks. In section 4 we present our conclusions and directions for future research

2.Tutorials

2.1 Bayesian Networks

Bayesian Networks (BN) are closely associated with subjectivist school, as opposed to frequentist school of reasoning (for discussion on some issues see [4]). This approach suggests those experts, or people identified as having deep knowledge in a specific domain are able to meaningfully articulate causal relationships and conditional dependence between variables they deal with. This approach allows us to define a systematic framework of probabilistic inference and decision analysis tasks when there is great deal of uncertainty, data is scarce and the most reliable source of knowledge is beliefs held by experts. The flexibility of BNs allows us to update opinions, as data becomes available.

Bayesian Networks (also called *belief networks*, *Bayesian belief networks*, *causal probabilistic networks*, or *causal networks*) [5-7] are directed acyclic graphs in which nodes represent random variables and arcs represent direct probabilistic dependencies among them. The structure of a BN is a graphical, qualitative illustration of the interactions among the set of variables that it models. The structure of the directed graph can mimic the causal structure of the modeled domain, although this is not necessary. When the structure is causal, it gives a useful, modular insight into the interactions among the variables and allows for prediction of effects of external manipulation. The numerical part of BN is a set of prior probabilities and conditional probabilities.

More formally, given V is a set of variables. Then a Bayesian belief network B over V is a pair (B_s,B_p). B_s is a directed acyclic graph with a node for each variable $v \in V$, called the *network structure*. B_p is a set of *assessment functions*, one for each variable v in V, defining a conditional probability of the variable (conditioned variable) given the variables that are its parents in B_s (conditioning variables). These functions quantify the strength of dependencies between the variables connected with an arc. Together, the assessment functions of a BN define a unique joint probability distribution over V that agrees with the interdependencies represented by the network structure. Once a BN is so constructed and initialized, it can be used to calculate the values of other variables given that a subset of V has been observed. Probabilities are updated using the Bayes' rule. Probabilistic inferences can be drawn from this network.

When a BN is used for probabilistic inference only it is frequently termed as "knowledge map" [8]. However, a popular use of BN is decision analysis [9]. A BN can be transformed into an influence diagram, which also incorporates utility and decision nodes [10]. Based on the utility functions and the probabilities encoded in the BN, decision alternatives can be studied. Another advantage of this particular formulation of Bayesian networks is that it allows for further conversion of an influence diagram into decision trees [11]. In this paper we shall rely on influence diagram and decision tree formulation of BN.

The original reference for influence diagrams is [12]. For some issues related to BN construction see [13, 14]. Algorithms for inference in BNs are discussed in [15][Huang, 1996 #12]. Some practical applications of BNs are described in [16]. Software that implements BN or influence diagrams are HEUGIN in [17], MSBN [18] and DATA [11].

2.2 Monte Carlo Simulation

Simulation is the process of building a mathematical or logical model of a system or a decision problem and experimenting with the model to gain insight into the system's behavior or to assist in solving the decision problem. A model is an abstraction of a real system. A BN can be seen as a *descriptive model* of V that that describes relationships between variables in B_s and provides information for evaluation in B_T. Monte Carlo (MC) simulation is sampling experiment whose purpose is to estimate the distribution of an outcome variable that depends on several probabilistic input variables. MC simulation can be seen as a way of managing the

uncertainty associated with input variables or testing the sensitivity of the model to its assumptions. The results of MC simulation are distribution of outcome variables obtained from thousands of combinations of values that the input variables could possibly take. General discussion on simulation applied to risk analysis can be found in [19, 20].

In their pure form, BNs operate on and calculate point probability estimates for the conditioned variables, analytically from the networks. Simulation can be used internally in a BN for updating and inference [21]. However, we are more interested in how simulation can strengthen modeling and decision making in a BN framework.

Since BNs are in essence models of relationships over a particular set V, where modeling assumptions and parameters are both noisy and uncertain, use of MC simulations allow us to do more robust decision analysis over the system. Simulation in this sense complements sensitivity analysis on influence diagram parameters [22] Further, point probability estimates are not very interesting for decision making in risk management as the expected value of the payoff nodes represent only long run expected averages. One at least wishes to know not just the mean of loss distribution (expected loss), but also a standard deviation (unexpected loss) and a high percentile like the 99% (catastrophic loss).

The way to do MC simulation is to specify uncertain parameters (like probabilities) as distributions on the conditioning nodes instead of simple point estimates. The MC algorithm then picks at values random from these distributions. The values of conditioned nodes and associated payoffs (expected values or expected utilities of decision alternatives) are then calculated in the normal way. Doing this hundred of times results in a distribution of outcomes that becomes a basis of more meaningful analysis.

Some of the good references looking at applications of simulation in decision analysis or artificial intelligence reasoning are [23-25]. A software implementing MC simulations on influence diagrams is DATA [11].

3. Exploratory Case Study

3.1 Context

Our case study was conduced in a mid-size, full service securities firm in Singapore. A securities firm is a financial intermediary between suppliers and demanders of liquidity. In this role it deals with a variety of investors, ranging from individuals to governments and corporations on one side and capital market instruments on the other side (for background on security firm's operations see [26, 27]).

The number of transactions in this firm is about 250/day with S$2.55 million per transaction value. This transactional throughput is about average for the industry. Its operating year of 250 working days a year is conventional for the industry. There were 10 full time employees working in the operations division of the firm during the time of the study.

The operations manager in this firm was concerned about some unusual losses mounting up in its securities settlement operations. He engaged one of the big-5 consulting firms in order to help with operational risk management and process improvement. Operational risks in the firm were divided into several main categories (like technology, human resources, transaction etc.), with each of the categories further subdivided into hundreds of sub-categories. For the purpose of this paper we will concentrate on a small subset of the operations dealing with securities settlement, which form a part of a much larger operational risk management project.

3.2 Methodology

The methodology for this research was as follows.

1) Development of a process flow diagram

2) Identification of major operational risks

3) Identification of causal structure of the events that lead to these risks.

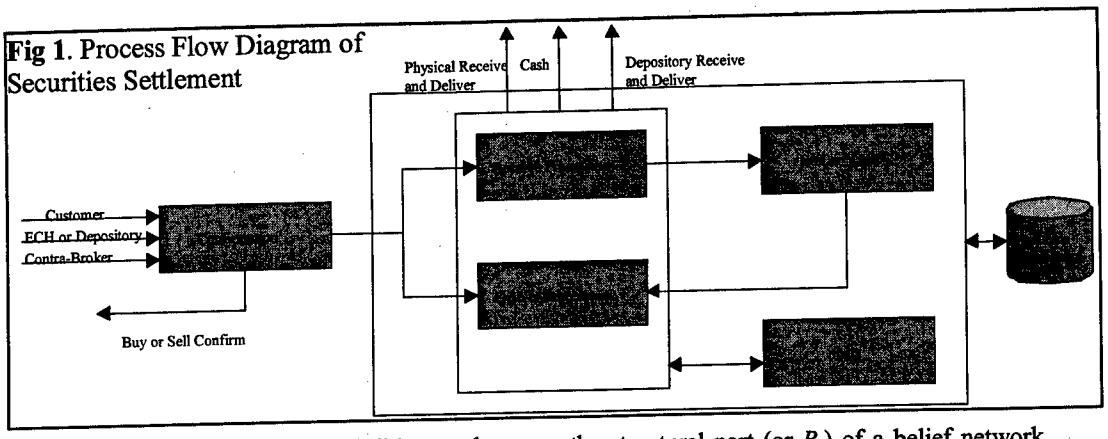


Fig 1. Process Flow Diagram of Securities Settlement

4) Determination of probabilities and loss distributions

5) Simulation over the belief network

6) Risk mapping and reporting

First a process flow diagram was constructed in discussion with the operational manager and the consultant in charge of this project. Several levels of these diagrams were constructed in order to clarify the flow of operations and to establish responsibilities for the different sub-processes and resources. The high level overview of the process is given in *Fig .1*. This diagram became the basis operational risk measurement strategy.

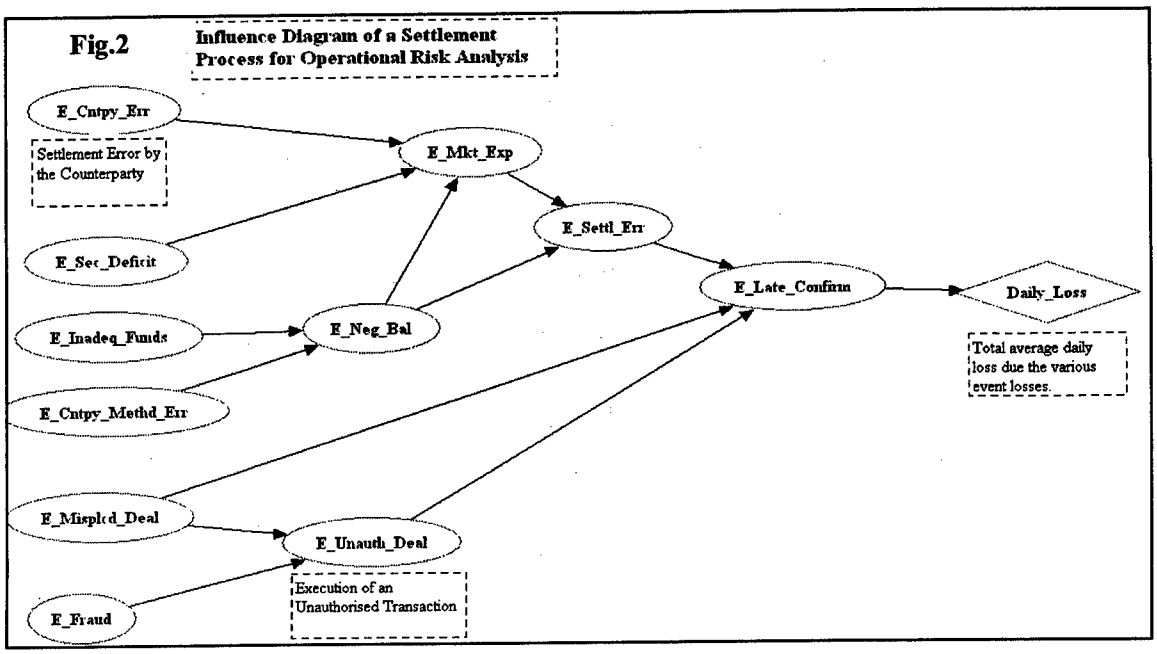
At the next stage we were concerned with finding out some of the operations risk that the firm was facing. We sent out questionnaires asking the employees about 5 of their "nightmare scenarios", or events which if they happened in the part of the process under their responsibility would cause a major loss (defined as being over a suitable daily limit set by the operations manager). This information was then triangulated with historical operating cost data and an important subset of these scenarios was chosen for further development.

The issue we aim to use as a basis of demonstration for this case study was late settlement of promised securities. A security firm enters into a legal obligation to settle a transaction within a set time period (usually within three days). Failure to do so exposes the firm to unbalanced position till the security is settled. This may lead to exposure to market risk , credit risk or liability risk.

After identification of appropriate problems the next stage was development of the structural part (or B_s) of a belief network. This was done in consultation with all the 10 people involved in the settlement problem. This required several structured group meetings. The elicitation was done in an iterative, top down method, where conditioning variables for the top level were identified and in their turn the causal influences that lead to them were identified. This stage also helped in revealing that several of the losses being realized in one part of the process were actually due to errors in other parts of the process. This formed the basis cost allocation at the reporting stage.

The next stage was determination of conditional probabilities and loss values associated with each event. Conditional probabilities were elicited using PROBES [28] a tool developed for this purpose. Probability was defined as the chance of a particular event occurring on a given day. Conditional probabilities were defined as the chance of an event occurring given its predecessor has occurred. For probability elicitation we asked the individuals under whose area of responsibility a particular event originated to work with PROBES. This was done because we believed that individuals who observe these events on a daily basis have a better idea about probabilities. Instead of eliciting point probability estimates, we elicited probability distributions.

For loss values a slightly different method was used. Loss was defined as the average daily loss associated with an event multiplied by the amount of time required for fixing the problem. Loss values due to specific event were determined by looking at historical cost data and by talking to the operations manager



because it was felt that the junior level people did not have a good overall idea of the effects of events on the firm as a whole. It was deemed reasonable to assume that losses were normally distributed and their specification involved assessment of their means and standard deviations.

This network structure was now constructed in DATA [11] and interactively refined. Only chance and value nodes were used since at this moment we were only interested in operational risk measurement and not in decision making.

After the network was initialized expected values of the aggregate operational risks were obtained. MC Simulation was then performed over this network by drawing the value of each chance node the probability distributions specified. In this manner, unexpected (equivalent to the Standard Deviation) and catastrophic (95% percentile) value. The description of this process and results follows in the next sub-section.

3.3 Results and Discussion

The causal network elicited for the settlement process was developed as an Influence Diagram in DATA (figure 2). Chance Nodes were used to represent loss events . For example, a Settlement Error by the Counterparty (*E_Cntpy_Err*) leads to a dollar loss for the firm and occurrence of the event could lead to further loss by causing a Market Exposure (*E_Mkt_Exp*) for the firm.

Each chance node can take two distinct outcomes (i.e. either an event occurs or it does not). The uncertainty associated with the probabilities of occurrence was accounted for by representing the probability of occurrence as a distribution e.g. a triangular or exponential distribution. The uncertainty of the estimates for the losses caused by the occurrence of an event was taken into account by representing the loss values for each chance node as a probability distribution e.g. normal distribution.

The arcs in the influence diagram between chance nodes represent the conditional probabilities. For example, given that the *E_Cntpy_Err* event has occurred the conditional probability that the event *E_Mkt_Exp* will occur is defined as a distribution. This helps propagate the uncertainty with respect to the probability of occurrence, and the dependencies between the events, through the network.

The *Daily_Loss* value node in the above influence diagram represents the aggregated dollar loss across all the events. For each simulation, each of the probability and loss distribution is sampled to determine the joint outcome of the events in the influence diagram. MC simulations were performed using DATA to determine the distribution of the *Daily_Loss*. The figure 3 in the next page is an output from an MC simulation for the above network.

From the distribution of the *Daily_Loss*, the average, standard deviation and 95% percentile loss values can be determined. The simulation runs for the illustrated causal network resulted in the daily loss values as presented in the figure.

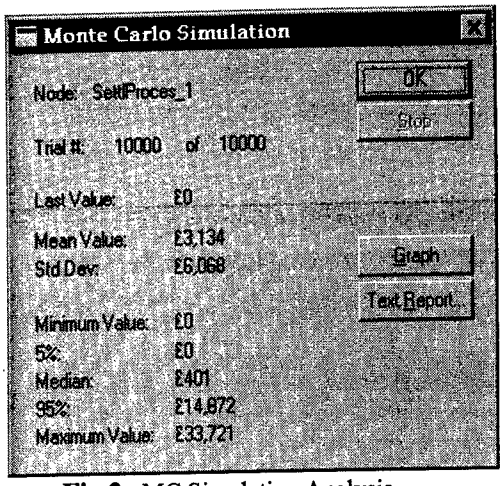
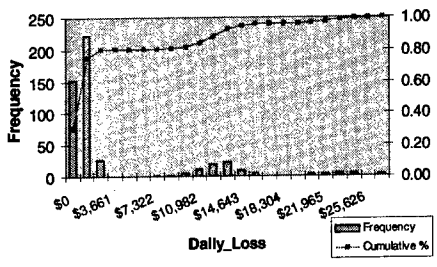


Fig 3 : MC Simulation Analysis

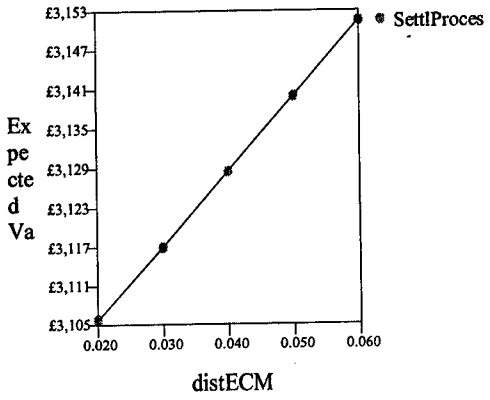
The expected daily loss was calculated as 3134. However, the unexpected loss comes up to 6,068, which is approximately twice the mean value. The 95% percentile value $14,872 represented Value-at-Operational-Risk due to late settlements at one day confidence interval. These three quantified values were then integrated with operational risks from other risk categories in order to arrive at an aggregate Value-at-Operational-Risk for the firm as a whole. The large spread of the loss distribution reflects the low-probability/high-impact nature of operational risks.

Fig. 4 Daily_Loss distribtuion



The other very useful feature of using an Influence Diagram in the Operation Risk Management process is the fact that they allow inferencing, sensitivity analysis with respect to the event's probability and loss distribution (figure 5). In the influence diagram under consideration, assigning extreme values to the loss distribution of the *E_Mkt_Exp* event

Fig. 5 Sensitivity distECME



resulted in a significantly large VaR value. This capability of an Influence Diagram is very important to Operation Managers who can then identify significant parameters of the causal structure and take suitable action to reduce the loss.

3. Conclusion

In this paper we have described an application of Bayesian networks and MC simulation in quantifying operational risk. We illustrated this application on a specific risk associated with a settlement process in a securities firm. The initial analysis look very interesting and this analysis will be expanded to other parts of the operational risk management process.

We are working towards further enhancements. We are considering specification of probabilities using any arbitrary distribution. The problems associated with this is there usually is very little data, and individuals with little statistical knowledge find it hard to specify their beliefs in terms of a distribution. We are also investigating the use of some opinion pooling methods in order to gain better triangulation on the beliefs of different experts. Some approaches based on quantile specification of probabilities look very promising.

As regards to loss distribution, we found that MC simulation over normal distribution does not readily reveal worst case scenarios, which are of great interest to managers. We are investigating the use of extreme value distributions in this case.

While the current implementation is promising for measurement of operational risks, we have not yet exploited the capabilities of Bayesian

networks for decision analysis. The complete system we are planning to implement will have a Bayesian network with direct data feed from firms operations systems. Addition of decision nodes coupled with simulation and optimization will allow looking at the change in risk profile contingent on various risk reduction actions by the manager and suggesting an optimal course of action.

References:

1. Ludwig, E., *Lesson from Leeson.* Risk, 1995. **8**(7).
2. Stoll, H.R., *Lost Barings: A Tale in Three Parts Concluding with a Lesson.* The Journal of Derivatives, 1995(Fall 1995).
3. Lynch, G.G., *Report of Inquiry into False Trading Profits at Kidder, Peabody & Co. Incorporated`*, . 1994, Davis Polk & Wardwell.
4. Henrion, M. *Uncertainty in artificial intelligence: Is probability epistemologically and heuristically adequate?* in *Expert Judgment and Expert Systems, Proceedings of the NATO Advanced Research Workshop on Expert Judgment and Expert Systems. 1986.* Porto, Portugal,Berlin, Germany: Springer Verlag.
5. Pearl, J., *Probabilistic reasoning in intelligent systems : networks of plausible inference.* 1988, San Mateo, Calif.: Morgan Kaufmann Publishers. xix, 552 p. : ill. ; 24 cm.
6. Jensen, F.V., *An introduction to Bayesian networks.* 1996, London: UCL Press. x, 178 p. : ill. + 1 computer disk (3 1/2 in.).
7. Ramoni, M., *Theory and practice of Bayesian belief networks.* 1999, :: Arnold,.
8. Howard, R.A., *Knowledge Maps.* Management Science, 1989. **35**: p. 903-922.
9. Shachter, R.D. and M.A. Peot . *Decision making using probabilistic inference methods.* in *Proceedings of the Eighth Annual Conference on Uncertainty in Artificial Intelligence (UAI-92).* 1992. San Francisco, CA: Morgan Kaufmann Publishers.
10. Cooper, G.F. *A method for using belief networks as influence diagrams.* in *Proceedings of the Workshop on Uncertainty in Artificial Intelligence.* 1988. Minneapolis, Minnesota.
11. DATA, *Decision Analysis by Tree Age (DATA)*, . 1997, TreeAge Software, Inc: Williamstown, MA.
12. Howard, R.A. and J.E. Matheson, *Influence Diagrams*, in *The Principals and Applications of Decision Analysis, Vol III*, R.A. Howard and J.E. Matheson, Editors. 1981, Strategic Decisions Group: Palo Alto. p. 719-762.
13. Henrion, M., ed. *Some practical issues in constructing belief networks.* Uncertainty in Artificial Intelligence 3, ed. L.N. Kanal, T.S. Levitt, and J.F. Lemmer. 1989, Elsevier Science Publishers B.V. (North Holland). 161-173.
14. Pearl, J., *Fusion, propagation, and structuring in belief networks.* Artificial Intelligence, 1986. **29**(3): p. 241-288.
15. Henrion, M., *An introduction to algorithms for inference in belief nets*, in *Uncertainty in Artificial Intelligence 5*, M. Henrion, *et al.*, Editors. 1990, Elsevier Science Publishers B.V. (North Holland). p. 129-138.
16. Heckerman, D., A. Mamdani, and M.P. Wellman, *Real-world applications of Bayesian networks.* Communications of the ACM, 1995. **38**(3): p. 24-26.
17. Jensen, F.V., *An introduction to Bayesian networks.* 1996, New York, NY: **SpringerVerlag NewYork Inc.**
18. MSBN, *Microsoft Bayes Network*, . 1996, Microsoft Corporation: Richmond.
19. Vose, D., *Quantitative risk analysis : a guide to Monte Carlo simulation modelling.* 1996, New York: Wiley. p. cm.
20. Evans, J.R. and D.L. Olson, *Introduction to simulation and risk analysis.* 1998, Upper Saddle River, N.J. :: Prentice Hall,. xvii, 279 p. : ill.
21. Shachter, R.D. and M.A. Peot, *Simulation approaches to general probabilistic inference on belief networks*, in *Uncertainty in Artificial*

Intelligence 5, M. Henrion, *et al.*, Editors. 1990, Elsevier Science Publishers B.V. (North Holland). p. 221-231.

22. Laskey, K.B. *Sensitivity Analysis for probability assessments in bayesian networks*. in *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*. 1993. San Mateo, CA: Morgan Kaufmann Publishers.
23. Slade, G., *Generating Explanations for goal-based decision making*. Decision Sciences, 1992. **23**(6): p. 1440-1461.
24. Raghunathan, S. and P. Tadikamalla, *Use of Stochastic Simulation in Knowledge - Based Systems*. Decision Sciences, 1992. **19**(3): p. 1333-1356.
25. Widman, L.E. and K.A. Loparo, *Artifical Intelligence, Simulation and Modeling*. Interfaces, 1990. **20**(2): p. 48-66.
26. Reddy, M., *Securities Operations*. Second ed. 1995, New York, NY: New York Institute of Finance.
27. Hoffman, D. and M. Johnson, *Operating Procedures*, in *Risk Magazine*. 1996.
28. Lau, A.H. and T.Y. Leong, *PROBES: A Framework for Probability Elicitation from Experts*, . 1999, Medical Computing Lab, School of Computing, National University of Singapore: Singapore.

A Study on Stock Data Mining by Map Recognition

Nianyi Chen*, Wenhua Wang and Dongping Daniel Zhu*****

*Laboratory of Data Mining, Shanghai, P. R. China

**Salomon Smith Barney, 250 West Street 8th Floor, New York, NY 10013, USA

***Zaptron Systems, Inc., 415 Clyde Avenue #108, Mt. View, CA 94043, Email: dan@zaptron.com

Abstract - Stock data analysis for price forecasting and trend prediction has been a challenging problem that attracts researchers from different fields. Some use statistical methods, while others use neural network based approaches. This paper reports on a preliminary study on stock market data analysis using a hyperspace data mining approach that is built upon a projective geometrical method. Discussions include data separation, feature selection, data pattern identification, and model building. Application of this method to stock performance classification and market speculation prediction is described. Preliminary results with real-world financial data seem to provide useful insights on how to discriminate the performance of different companies and to identify the market speculation manipulated by large investors.

Key Words: stock data, data mining, projection, time series, pattern recognition, classification.

1. Introduction

There are a number of objectives in financial data analysis. They include investment evaluation, building of mathematical models to predict market prices [15], quick identification of clusters from survey data, demand forecast based on customer satisfaction (or consumption utilization) [17], consumer behavior analysis in terms of satisfaction, preference, subject appreciation [15], correlation and association analysis to relate goals (targets) to factors, and case studies by computer model simulations.

Traditional analytical methods include those from classical statistics including linear and nonlinear regression analysis, factor analysis, correlation and association, time series analysis, and those from artificial intelligence including artificial neural networks (ANN), fuzzy expert systems, genetic algorithms, and so on. In pattern recognition applications, PLS (partial least square) method is often used to find quantitative target-factor relationships. However non-linearity exists among targets and factors that calls for new methodology. To develop a new methodology for financial data mining solutions, a number of important issues need to be addressed, such as feature selection, data separation, and model building.

How to select and use financial factors to describe the underlying operation of a financial system is an important but complicated problem. People often rely on theories in micro-economics, macro-economics, and econometrics to select features that best describe financial systems. In most cases, features used in financial analysis are extracted by human intelligence. However, financial system structures are very complicated, and they are often described by a large number of seemingly unknown factors. The difficulty is how to choose the right set, or a reduced set, of the factors that correlate the financial activity with the structure of a financial system. It seems that the empirical rules developed by human intelligence can also be discovered by computer software that implement powerful methods designed for feature selection and feature reduction purpose.

The data separability criteria, implemented in the MasterMiner software and reported herein, are rather useful in selecting key factors that influence the financial performance of a business. People often use linear and nonlinear regression in data separation that gives poor results. MasterMiner is useful in simplifying the selection of nonlinear terms in regression. It has been compared favorably against other popular software. The former uses far less terms in mathematical models, and produces lower prediction residue error squared sum (PRESS) than the latter.

This paper describes a hyperspace data mining method [3] [7] that has a number of advantages over those based on pure ANN or pure regression (such as PCA, principal component analysis) methods often used in financial data mining. The reported method has proven to be very effective in dealing with non-linear, non-uniformly sampling, non-Gaussian and multi-variant cases in non-financial applications, such as petrochemical, materials design and process controls. As such, it is reasonable to extend its use to financial data mining. Examples are included to show the efficacy of this new method for stock performance classification and speculation prediction.

2. Review on Classification Techniques

Classification techniques are grouped into two categories: decision theoretic (or statistical) and syntactic [14]. In the statistical approach, features are extracted from input patterns, and classification carried out by partitioning the feature space in probabilistic terms. In this method, however the important structural relationship among features is lost often times. Bayesian reasoning, as a major statistical method, resolves uncertainty by resorting to the *principle of indifference* – probabilities are distributed *uniformly* among all events known to have relevance, leading to inconsistencies in decision. Another obstacle to Bayesian method is that a classifier is required to have complete and accurate knowledge (in a database) of both *a priori* and *conditional* probability distributions, or the classifier is forced to guest anyway no matter how impoverished the information is. Therefore the validity of this procedure in practice is questionable. In contrast, syntactic methods extract structural information from data, and a class is characterized by several sub-patterns and relationships among them. Both techniques, however, are not adaptive to the uncertainty associated with real-world data, and have limitations in applications.

AI-based methods include fuzzy logic, neural networks and genetic algorithms, and they have received wide attentions in recent years, especially as the Internet grows. Fuzzy logic is precise reasoning about imprecise concepts for human reasoning, and fuzzy sets handle uncertainty effectively in that they have no well-defined boundaries, and the transition from full membership to non membership is gradual [1][13]. One could map the human knowledge into fuzzy rules, and make classification by fuzzy reasoning [18]. One could also use fuzzy clustering methods, such as fuzzy c-means, to classify patterns into different categories. However, a *pure* fuzzy technique has limitations: (1) lack of adaptive learning ability: it cannot learn classification knowledge from data, and (2) incompleteness and fuzziness in representing experts' knowledge: even experts cannot clearly describe approximate reasoning under uncertainty and they often make wrong decisions.

Crisp neural nets (supervised or unsupervised), on the other hand, can mimic the biological information processing mechanism in a very limited sense. They have been used as alternatives to traditional classifiers [2][6][12]. They have a number of advantages, including (1) high computation rates because of massive parallelism, (2) adaptivity in learning decision rules, and (3) a greater degree of robustness or tolerance against uncertainty. However, crisp neural classifiers have intrinsic *shortcomings*: (1) they represent knowledge by distributed crisp weights that, unfortunately, have no explicit physical meanings. By adjusting weights, such a net can only extract knowledge at low-level (represented by numerical weights) rather than at high-level; (2) they cannot directly process symbolic data (linguistic values, e.g., "very high," and "about 55 grams,") because its weights can only store crisp numerical values (e.g., "-55.88."). In summary, a *pure* neural approach is not suitable to handle fuzzy and uncertain knowledge arising in the complex real-world data.

To overcome the limitations of pure neural and pure fuzzy approaches, the neurofuzzy methods are studied that offset the demerits of one paradigm by the merits of another. In a narrow sense, a fuzzy neural net is a fuzzy-operation-oriented neural net, implemented by fuzzifying inputs, output and weights, and using fuzzy set operations. However, the narrow-sense fuzzy neural network can not extract fuzzy rules from data because no explicit physical meanings are attached to the crisp or even fuzzy weights, whereas physical meanings are directly related to fuzzy rules used in classification. In a general sense, a fuzzy neural network is a fuzzy-reasoning-oriented neural network with adaptive learning and fuzzy reasoning. More importantly, such a method is capable of extracting fuzzy rules from given data. However, many fuzzy neural networks are Crisp-Input-Crisp-Output (CICO) models, not applicable to the cases that are described by a Crisp-Input-Fuzzy-Output (CIFO), Fuzzy-Input-Crisp-Output (FICO), or Fuzzy-Input-Fuzzy-Output (FIFO) model. As an improvement, a reasoning-oriented fuzzy neural network, called Crisp-Fuzzy Neural Network (CFNN) is proposed in [19]. CFNN includes FIFO, FICO, CIFO, and CICO model. Furthermore, Genetic-CFNN (GCFNN) is used to heuristically initialize fuzzy weights of a CFNN to avoid bad local minima.

3. Brief Background of Hyperspace Data Mining

The hyperspace data mining method [3][7] is a novel approach to nonlinear optimization for pattern recognition problems, and it has proved to be a powerful tool for design and decision optimization in many non-financial applications, including fault diagnosis and metallurgy and optimization [4][5][11]. New applications for data mining are being explored in different fields, including stock analysis, environment emission controls, computer products service data analysis, data network controls, and tobacco production optimization. This paper describes a preliminary study on stock data analysis using this method.

The basics of this method is the MREC (Map RECognition by hidden projection) methodology. MREC is an effective approach to statistical pattern recognition, and it seems to outperform the classical PCA, Fisher and PLS (Partial Least Square) methods in many applications. It is equally applicable to nonlinear problems that arise in many applications [3][4][5]. The MREC methodology consists of three steps: (1) data separation by a *hidden geometric transform*, (2) feature selection by *data geometric pattern* ("one-sided" or "inclusive" type), and (3) building models that reduce a complex nonlinear problem to a set of simple linear models in sub-spaces. All these functions are built into software MasterMiner.

MREC - Statistical pattern recognition methods are based on computerized recognition of m-D graphs (or their 2-D projections) of sample distribution in a m-D space. Independent variables (features) influencing the model are used to span an m-D space. If one can describe samples of different classes as points with different colors (or labels) in the space, a mathematical model can be obtained by data mining that describes the relationship (or regularity) between targets (goals) and features. In MREC, the hyper polyhedron model is used where samples are classified into class "1" (red) or class '2" (blue). Without loss of generality, it is assumed that a hyperspace or its subspace contains one and only one optimal zone that can be enclosed by a concave or convex hyper polyhedron. This polyhedron is used to describe the boundary of the optimal zone in which all sample points are of type "1" or red. This assumption is always true, since a hyperspace can be divided into a number of subspaces that can always be enclosed by a hyper polyhedron.

Unlike the regression methods (linear, nonlinear, logistic regression, etc.) or the neural nets that provide *quantitative* solutions, MRE can provides *semi-quantitative* and *qualitative,* as well as *quantitative* solutions. This is advantageous because real-world data exhibit strong noise, and quantitative models would be *too precise to* represent them. The PCA-based regression builds linear models without data separation, shown in Figure-1, whereas MREC regression first tries to separates data, and then builds more realistic models from a *reduced set* of data, shown in Figure-2.

Data Separability - The data separability test of MREC is designed to explore the possibility of separating data into different populations or clusters in the hyperspace. Building a model for a non-linear problem is possible only if the data set is *separable*. At each iteration, MREC chooses the "best" projection map with *maximum separation* from a series of *hidden projections*, and discards those samples outside the optimal zone (see the red box in Figure-2). After each projection, samples of class "1" (red) are automatically enclosed by a "tunnel" (the intersection of two tunnels are shown to form an "auto-square" in Figure-3), and *a reduced data set* is formed that contains only samples within the intersection. Then a second MREC is performed on this reduced set to obtain the next "best" projection to further separate data into different classes. After a series of such projections, a complete (close to 100%) separation could be realized, and the resulting data set is used to build a *very accurate* model. The physical meaning of MREC is explained by Figure-3, where each "auto-square" is formed by two "tunnels" in the original m-d space, and several such tunnels would form a *hyper-polyhedron* in the m-d space. This hyper-polyhedron, enclosing all or most "1" (red) but no "2" (blue) samples, defines an *optimal* zone in the m-d space. MREC has been shown to be much more powerful than various regression methods.

Back Mapping - After the MREC transform of data from the original measurement (or feature) space into a number of orthogonal sub-spaces, one needs to back map the transformed data into the original feature space to derive mathematical models for practical use. Two methods, called linear and non-linear inverse mapping (LIM, NLIM) or PCBs (principal component mapping) [7][11] have been developed whereby a point in a low-dimensional principal component subspace is continuously back-projected to a high-dimensional space until the original feature space. Table-2 and Figure-5 give one example of LIM where a set of linear equations (inequalities) are obtained from (red) class "1" samples inside the auto-box by MasterMiner. These equations represent the model that is sought.

Feature Reduction - The rate of data separation, **R**, is defined as $\mathbf{R} = (1\text{-}N2/N1)$, where N_1 and N_2 are respectively the number of class "1" and "2" samples inside the polyhedron. If **R** is larger than 70%, the separability is "acceptable," otherwise it is "unsatisfactory." **R** is used as a criterion in feature reduction – a feature can be removed if **R** remains the same after being removed from the model. **R** has been used to reduce feature number by 1/3 to 1/2.

Concave Polyhedron -- Since MREC only forms a *convex hyper-polyhedron*, it may not separate data that form a *concave* rather than a *convex* polyhedron in the space. In these cases, the BOX method, shown in Figure-4, offers a powerful solution whereby samples of class "2" are cut off from the polyhedron so that all samples inside have type "2."

4. Stock Performance Classification by Price Ratio

In Shanghai Stock Exchange (SEE), in April of each year, all listed companies would publish an Annual Economical Report with detailed financial data. These data will exert an influence on the price of the stocks over a long period of time (for example, one year). It is possible to use these data to span a hyperspace, and map the data of each stock onto this hyperspace. We then classify the samples (representative points in the hyperspace) of every stock as either "good" (class "1") or "bad" (class "2") points, according to the company performance in one year. MasterMiner is used to find the distribution regularity of the two types of data points, and build a mathematical model that represents the data in the hyperspace. The model so found it used to classify companies into "good" or "bad" class as investment guidance. This method has been successfully used on the stock data available from the SSE. The classification and modeling results are rather encouraging.

In one analysis, the stock data of 56 out of 828 companies on SSE are selected for performance classification. In this study, the target γ is the ratio of the stock price in January to that in December of the year. Samples from 29 "good" companies with γ above 1.0 are of class "1," and data from 28 "bad" companies with γ less than 1.0 are of class "2." Table-1 lists the futures that are used in computation. Figure-3 shows the result of data separation on stock data for 1997. It is seen that the data separation by MasterMiner is close to 96%, a fairly good separation rate in practice.

Table-1 List of Features Used

Feature	Definition or meaning
CO	Comprehensive index (debt, liguidability. ..)
S	Market return rate = stock price/earning
IN%	Annual increase rate of earning = E(k)/E(k-1)
J%	Return rate on net asset
M%	(Profit per share) / (net asset/total shares)
Jb	(Stock price) / (net asset/total shares)

5. Prediction of Stock Speculation

Every day, the SSE publishes a Stock Index to show the general trend of each stock. For a particular stock on a particular day, when an obvious and sudden deviation from the general (normal) trend is observed, and no physical evidence is available to justify this deviation, it is reasonable to contribute this sudden change to the result of stock *speculation*, an operation secretly controlled by a large investor in an effort to manipulate the stock market for quick profit.

Stock speculation has patterns that seem to be detectable by data mining methods. When speculation happens, in general the price of a speculated stock will go through several ups and downs. ("waves") before reaching a "top price" at a peak, called top price peak, that is followed by a sharp fall-down. Therefore, it is very important to detect such a speculation pattern and predict the top price in real-time, to determine if there is a stock speculation that is going on.

MasterMiner is used to process SSE stock data to identify the speculated stocks and predict the top price before it falls down. More than 20 price-time curves of various stocks have been used as the training set. Time series analysis algorithms are developed, and used on 22 stocks including Tsinghua Tongfang, Xiaxin Electronic, and etc. It has been observed that the *K-curves* often exhibit several small peaks before the *top price peak*, followed by a sharp fall-down. However, recognizing the top price peak before a sharp fall-down is a challenging task that calls for innovation. In this study on speculation prediction, our target is the price at the end of the day, and features are the daily price and the total number of stock exchanges during the last N days (N is 7 or longer) or their functions. Samples representing the "top price" pattern, i.e., a peak followed by a sharp fall-down, are defined as class "1," and the rest class '2." After spanning a hyperspace with sample data, discrimination of the top price peak from other peaks is realized by the hyper polyhedron modeling method. The MREC methodology is able to identify certain behaviors or regularities of the "top price." The reliability of the regularities so found has been tested by a leaving-one method for cross validation. When data of last 10 days are used, a prediction accuracy of 70% is achieved. Improvement on prediction performance is under study to make this method more useful in practice.

6. Conclusions

The reported method of hyperspace data mining has been successfully used in industrial process optimization and controls. In this study, we try to use it for stock market analysis. Although the preliminary results are rather exciting, new methodology should be developed to supplement current method, making it more practical for financial data analysis. Data from stock markets often exhibit much stronger noise than those from a smoothly operated factory. The former also exhibits the nature of a time series that is influenced by many unknown factors. The development of a novel feature selection method is an imminent task.

References

[1] J. Bezdek, "Fuzzy Models for Pattern Recognition: Methods that Search for Patterns," IEEE Press, 1992.

[2] G. Carpenter and S. Grossberg, "Neural Nets for Vision and Image Processing," The MIT Press, 1992.
[3] N. Chen, Z. Li, Y. Chen and D. Zhu, "N-factor pattern recognition for process optimization and materials intelligent design," Proc. of the 1st Int'l Conf. Multisensor-Multisource Information Fusion, Las Vegas, NV, July 6-9, 1998.
[4] N. Chen and et al., Chemometrics and Intelligent Laboratory Systems, vol. 45, p.329, 1999
[5] N. Chen and et al., Journal of Physics and Chemistry of Solids, vol. 58, p.731, 1997
[6] N. Chen and W. Zeng, "Artificial neural network method applied to enthalpy of fusion of transition metals, J. of Alloys abd Compounds, Vol. 235, 1997.
[7] N. Chen, W. Wang and D. Zhu, "Data mining for industrial diagnosis," AAAI Spring Symposium on Equipment Diagnosis and Maintenance, Stanford University, Palo Alto, CA, March 22-24, 1999.
[8] N. Chen, "Design and production optimization using data mining," (in preparation)
[9] N. Chen and D. Zhu, "Intelligent materials modeling by hyperspace data mining," 2nd Int'l Conf. Intelli Processing and Manuf. of Materials (IPMM'99), Hawaii, July, 1999.
[10] N. Chen, W. Wang and D. Zhu, "Data mining application in biomedical pattern recognition," the 2nd Int'l Conf. on Information Fusion. Sunnyvale, CA, July, 1999.
[11] H. Liu, Y. Chen and N. Chen, "PCB method applied to material design - computer aided synthesis of a super-conductor," J. of Chemometrics, Vol. 8, pp. 429-443, 1994.
[12] I. Guyun, "Advances in Pattern Recognition Systems Using Neural Net Tech," World Science, 1994
[13] A. Kandel, "Fuzzy Techniques in Pattern Recognition," John Wiley, 1982.
[14] A. Kulkarni, "Artificial Neural Networks for Image Understanding," Van Nostrand Reinhold, 1994.
[15] Y, Wang, D. Zhu, "Further results of fuzzy utility functions with applications – dynamic case," INFORMS Fall Meeting, American Operations Research Society, Oct. 1998.
[16] D. Zhu and N. Chen, "A data mining approach to pattern recognition for detection and diagnosis," INFORMS Fall Meeting, American Operations Research Society, Oct. 1998.
[17] Y. Wang, D. Zhu, "A study on fuzzy utility functions with in financial data mining – static case," Proc. Int'l Conference on Multisource-Multisensor Information Fusion (Fusion98), Las Vegas, July, 1998.
[18] H. Zhang, D. Zhu, "A framework for building expert systems," IEEE 18, Monterey, CA, Nov., 1997
Y. Zhang and A. Kandel, "Compensatory Genetic Fuzzy Neural Networks and Their Applications," *Series in Machine Perception Artificial Intelligence*, Volume 30, World Scientific, 1998.
[19] Y. Zhang and A. Kandel, Compensatory Genetic Fuzzy Neural Nets and Applications, *Series in Machine Perception Artificial Intelligence*, Vol. 30, World Scientific, 1998.

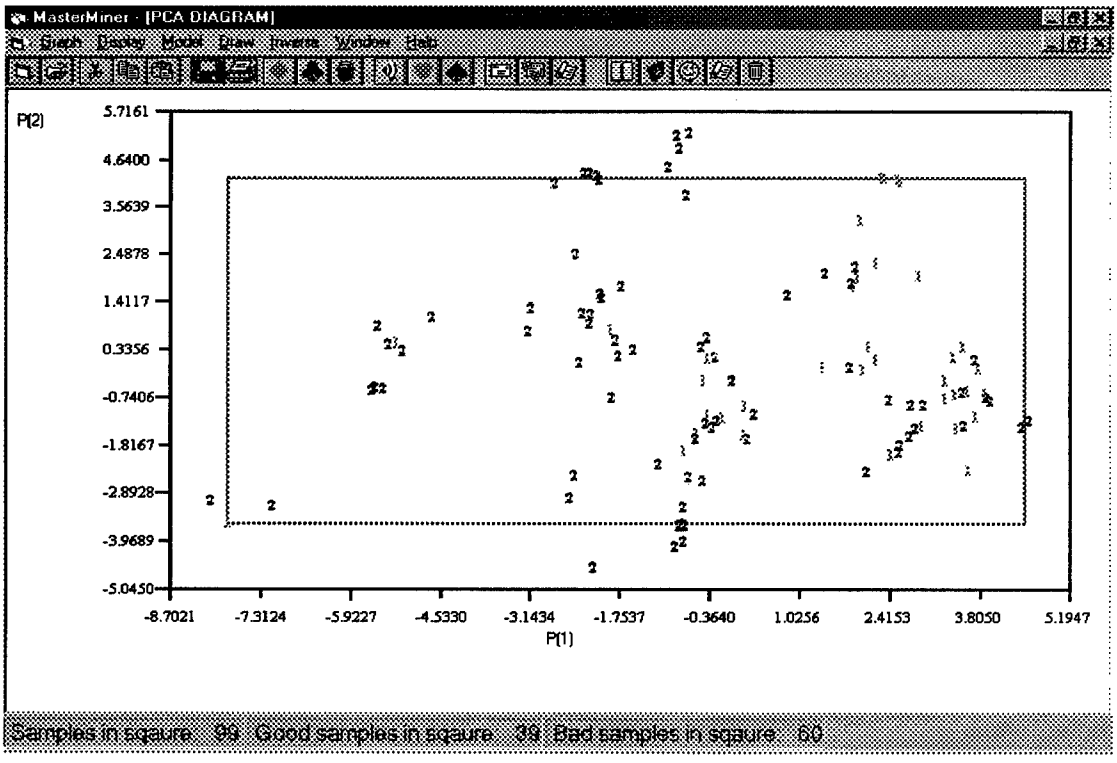


Figure-1 Principle component analysis – no data separation and "1" and "2" are inside the box.

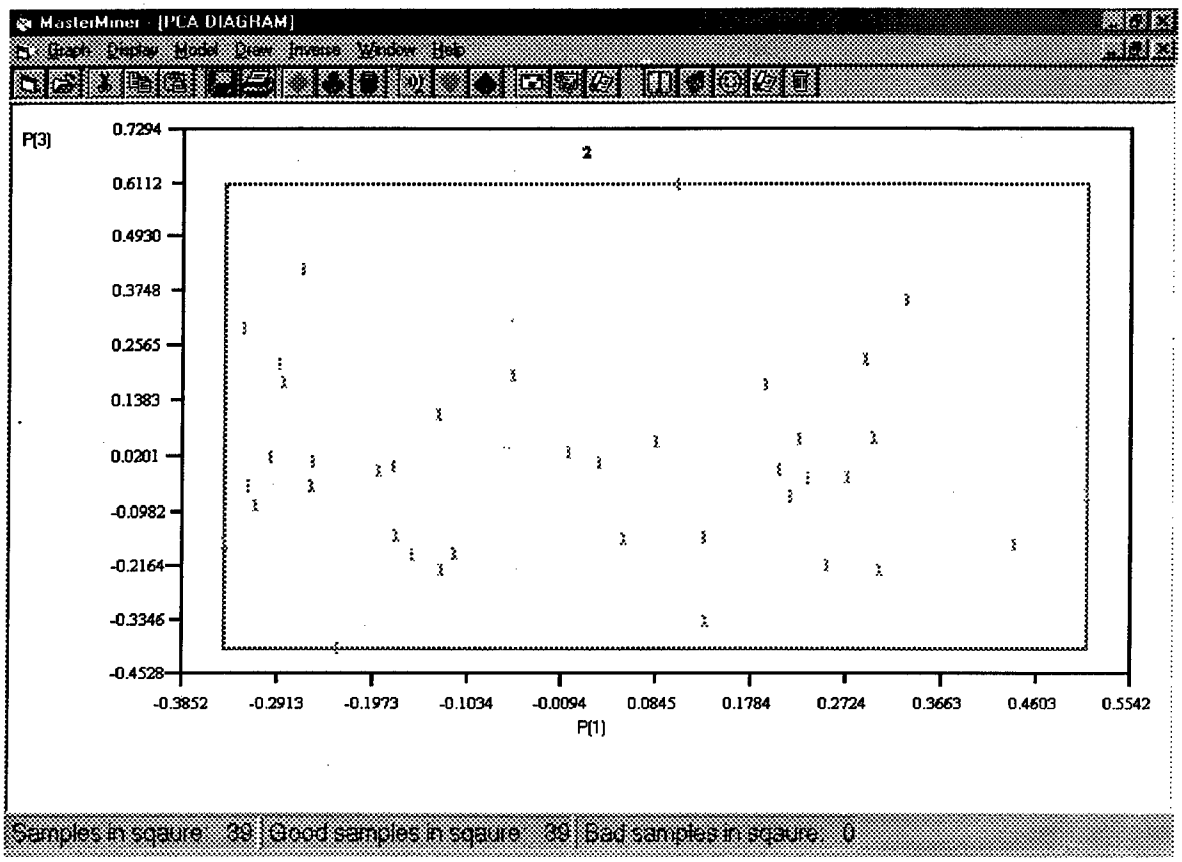


Figure-2 MREC – complete data separation, all "1" points (red) are inside the box.

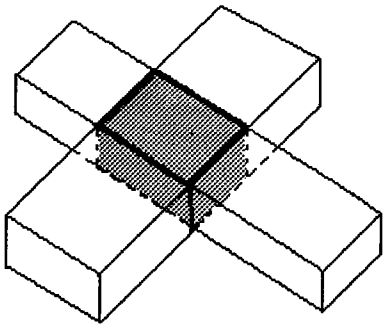


Figure-3. A polyhedron by intersection of two "tunnels."

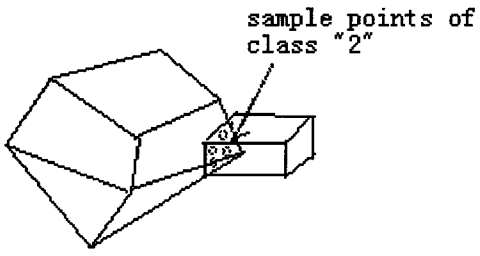


Figure-4 Build a convex polyhedron from two others.

Table-2 Model obtained my MasterMiner™

All inequalities obtained in original space:

+7.842<=-0.865[a1]+2.089[a2]+0.254[a3]+3.115[a4]<=+8.658

-0.645<=+1.089[a1]+2.236[a2]-0.077[a3]-1.317[a4]<=+0.028

-7.243<=+1.560[a1]+0.031[a2]+0.369[a3]-3.570[a4]<=-6.521

-4.197<=+0.096[a1]-2.008[a2]-1.203[a3]-0.802[a4]<=-3.344

-8.447<=+1.501[a1]-1.071[a2]+0.015[a3]-3.572[a4]<=-7.653

-1.661<=+0.076[a1]+0.537[a2]-0.580[a3]-0.747[a4]<=-0.994

where [a1], a[2], a[3] and a[4] are original features. The Auto-Box on the right side covers all red points, showing 100% data separation.

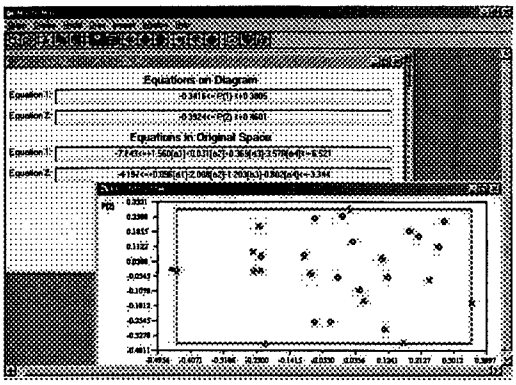


Figure-5 The auto-box with only "1" points and mathematical equations (above) computed.

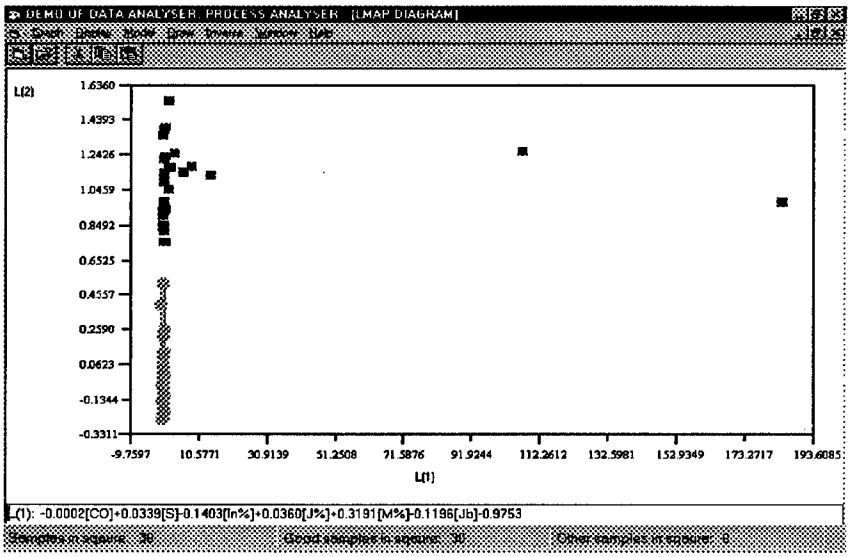


Figure-6 Performance classification, the lower cluster (red) for "good," and upper cluster (blue) for "bad." The two points on the far right represent two companies that are experiencing internal financial transition.

An Universal Method for Non-Linear Systems Equations and Non-Linear Programming Problems

Kamyshnikov A.Vladimir
Department of Economy, Tomsk State Architectural University,
634003, Tomsk-3, POB 3046 Russia E-mail: kva_son@yahoo.com

1. INTRODUCTION

Optimizational approach to solving queries of complex systems synthesis is in itself a big reserve for elevating quality of planning, management and projecting. The choice of optimizational aim areas of changing parameters is the task of particular economic and technical brunches of science. What concerns the optimization mechanism is the subject of mathematical programming.

The success of linear programming in increasing effectiveness of economic modeling and optimization of planning is well known. It is less obvious in technique, management and projecting as for added accuracy of above mentioned it is important to take into consideration nonlinear effects. Creation of "simplex method" and appearance of powerful PCs made linear programming an important tool for solving different problems but at the same time showed its weakness.[1)] Most of the queries cannot be adequately solved by linear programming model because they include nonlinear goal functions or constraints. All the above enumerated attracted special attention of mathematicians to the progress in nonlinear programming research.

It is important to mention that many nonlinear optimizational queries that exist in economy and technique are described as including convex or concave functional and convex areas of possible solutions parameters (for more information see Casten theorem). In literature on the subject, attention is mainly attracted to convex programming. The reason is the possibility to produce universal method for solving basic form equations as well as others with deviations from the basic form.[2)] One cannot guarantee the same for other nonlinear equations with wider range of parameters. Let me leave aside the prior and talk more about the later, as it is the subject of my research.

2. DESCRIPTION

Before I turn to the description of my research, let me first speculate a little on main definitions of nonlinear equations. These are tasks where two qualifications are relaxed: dividing and adding, which means that goal function and constraints might be nonlinear and variables can take values from some multitude, including discrete multitude. To the above mentioned problems of mathematical programming, refer the following:

- Tasks of integer programming with linear and nonlinear goal functions and constraint functions;
- Combined models of integer programming problems with linear and nonlinear goal functions and constraint functions;
- Problems of discrete programming, in which the value of variables is chosen out of given value multitude of rational numbers, not necessarily integer, also with linear and nonlinear goal function and constraint function;
- Problems of nonlinear programming with convex and concave functions;
- Problems of nonlinear programming with multiple-optima functions on non-convex or/and convex or/and non-linked zone of variables;
- Solving of equations and systems of equations, with functions that I identified above.

A large number of books are devoted to the problem of solving above-mentioned queries. It is enough to tell that G. Vagner's "Basic operational research", published in 1973, gives references to more than 162 publications, "Nonlinear programming", by Bazara and Shetti, published in 1982 sources 600 books. Literature on the subject is not of course bound to what was brought up here. The reason for such an attention lies in the provisions of practice. The absence of effective and quick method for finding global optimum in such problems leads to artificial simplification of mathematical model in practice. As a consequence system characteristics in economy, finance, technique, chemistry, physics, architecture, etc deteriorate.

To my knowledge, there is such an opinion among specialists that to create a universal method for

[1)] Mokhtar S. Bazaraa, C.M. Shetty, "Nonlinear programming, Theory and Algorithms", John Wiley and Sons, New York, 1979, p.8
[2)] Judin D.V., "Mathematical Programming", Moscow Press, Moscow, 1982, p.39

solving nonlinear problems, like Simplex -method that is used for linear queries, is impossible and the only way out are specialized methods for each particular type of programming problem. I hope now I can prove this information not up-to-date. Application of specialized algorithms requires exact correspondence of real model and problems that can be solved with the help of this particular algorithm. In case of any deviations you have to change it usually by simplifying mathematical model to the requirements of algorithm. It is important to mention also that initially the reality was already simplified in the model (in order to make it equal to the needs of an algorithm). As a result you answer the question that was not asked by the practice. In order to meet the requirements of imperfect mathematical method we replace the model of real natural conditions by the picture that either you or algorithm found suitable. Such fractional results are inapplicable for problems of integer programming because this data is used for forming planned decisions in complicated situations. Such problems with nonlinear functions are important for non-proportional fluctuation of expenses, determination of productivity value, quality assurance, in technical and economic addendum, in the sphere of nonlinear physical laws and others.

To begin with the description of my research, I want to insist on the fact that I *created a unique and general method* for solving non-linear queries of any level of complication. Queries that this method is solving are used in calculation of optimization problems and solution of nonlinear equations, for such areas as Modeling and Simulation, Operations Management, Machine Vision, Robotics & Automation. Using my method I successfully solved test queries given by Wolf, H.H.Rosenbrock and J.D. Powell, as well as different sets of non-linear equations. Using it, I am capable of solving quite complicated non-linear optimization problems of the classic form:

$$\min F(X)$$
subject to:
$$Wi(X) <= 0,\ i = 1, \ldots, m_1,$$
$$Gi(X) = 0,\ i = 1, \ldots, m_2,$$

Where X is an n-vector, and the functions F, Wi and Gi should not have gaps and can be linear, nonlinear, multi-extreme, convex and non-convex. My universal method of finding optimal solution for the above-enumerated programming queries consist of two well-known and simple algorithms: method of *proportionally deformed polyhedron* and method of *gradient descending*. It is well known that these two methods usually produce only local, separate solutions. But it is not true for my technique. In the algorithm these simple methods let me always acquire global solutions. Obviously there is no magic in it. The results obtained by me (proved theoretically) are the effects of extending the theory of convex functions. My algorithm was checked during two years of investigations and illustrated by solving of several hundred, better to say thousand of, tests and real nonlinear equations (not all of them were included in this work as examples) and also applied theoretically. My algorithm meets all the requirements that I imposed before creating it. I was very thorough while testing it. My testing included the following steps:

- I solved all test equations that I could find in the literature on the subject.
- All real problems from the literature were solved too.
- Successfully tested it on several hundred self-created equations with number of variables starting from 1 up to 500
- Solved equations with implanted linear and nonlinear constraints, including discrete constraints on variables
- Solved several sets of equations for clients via e-mail

Results obtained were always satisfactory which means that algorithm was able to find global optimum for multi-extreme function, as I was only calculating that type of function. Program that embodied my method and that can also be used, as subprogram is in FORTRAN-77, MS DOS. Resolution time depends on the type of a problem or number, and complexity of functions. For Pentium 200, I did not find an equation for which solution time was longer than an hour. In my practice programming, a query with 100 variables was solved within 15 minutes. Algorithm stops its search when either global optimum is found, or a particular number of iterations is achieved.

3. COMPARISON

It is worth mentioning that from the beginning of realization of existence of both local and global optimum or optimums, we find information on algorithms for their location. It is hard for me to make any connections or comparison of my method with others. In accordance with the information of Northwestern University and the Argonne National Laboratory of in USA, such a method of mine does not exist. Scientists consider it unrealistic to expect to find one general nonlinear optimization code that is going to work for every kind of nonlinear model. Instead one should try to select a code that fits the problem one is solving.[3)]

[3)] For more information see http://www.mcs.anl.gov/otc/Guide/faq/nonlinear-programming-faq.html

I will try here to bring about possible comparison points. It is feasible to relate solution time of my algorithm and others. I was comparing my algorithm to stochastic search code by Georgian professor Chichenadze. Test task # 14 was solved within 14 seconds while it took Chichenadze 10 hours.[4] Modified method of nonlinear and stochastic optimization systems based on ideas of Professors L. Ingber and J. Powel allows to make solution more precise and finds global optimum with set before probability but it does not guarantee theoretically global solution. For this algorithm there will emerge a counterexample that cannot be calculated by it. This cannot be said about my algorithm which finds global optimum every time. From the calculation point of view it does not have common error - the rounding default of the results accuracy. Error is equal to the error of optimum calculation in optimal point, which is determined by the chosen software and type of the computer. However I should add that any universal algorithm (and mine is not an exception) will not be as good for solving some particular problems as algorithm specifically created for them. To my knowledge there is no other algorithm with such broad universality as mine.

4. EXAMPLES

Computing examples (numbered N_1 to N_14) using the proposed algorithm are presented below to show the efficacy of the method.

N_1. Equation with multi-extreme goal function.
Minimize:

EE1=exp{SIN(10*XX)*SIN(123+XX^0.73+3*XX^2.13+XX^1.55)}
where :
EE2=ABS{[X1^0.9+X2^0.8]/[1.1+COS(X3^2.3+X4^1.3)]} -3<0
EE3=XX-ABS{(X1^0.9+X2^0.8)/[1.1+COS(X3^2.3+X4^1.3)]}=0
Solution:
EE1= 0.3682, X1= 1.9484, X2= 1.5357, X3= 2.6601, X4= 1.445, EE2=-0.3259, EE3=0.00

N_2. Equation with multi-extreme goal function.

Minimize:
$EE_1=3(SIN(4(3+X_1)))^2+(SIN(4+X_1))^{12}$
IF $X_1<7.9$ OR $X_1>8.1$ then $EE_1=2+SIN(EE_1)$
where(<=0):
$EE_2=-X_1$.
Solution:
$EE_1=0.00061823$, $X_1=7.995$.
Control:
$EE_1=0.001509568$, $X_1=8.00$

N_3.. Equation with multi-extreme goal function. Minimize:

$EE1=SIN((X_1+SIN(X_1)+4)^2/(10+EXP(SIN(33X_1))))$
IF $X_1<16$ or X_1 gt 16.2 then $EE_1=2+EE_1$
where(<=0):
$EE_2=-X_1$.
Solution:
$EE_1=-0.99999999$, $X_1=16.097$
Control:
$EE_1=-0.975$, $X_1=16.1$

N_ 4. Equation with multi-extreme goal function (function of Rozenbrok).
Minimize:$EE_1=100(X_2-X_1^2)^2+(1-X_1)^2$
where(<=0):
$EE_2=-X_1$.
Solution:
$EE_1=0.00$, $X_1=0.99999$

N_ 5. Equation with multi-extreme goal function (function of Powell).
Minimize:$EE_1=(X_1+10X_2)^2+5(X_3-X_4)^2+(X_2-2X_3)^4+10(X_1-X_4)^4$
Where:
$X_1., X_2, X_3, X_4.>=0$
Solution:
$EE_1=0.00$, $X_1=0,00$, $X_2=0.00$, $X_3=0.00$, $X_4=0.00$

N_ 6. Equation with multi-extreme goal function.

Minimize
$EE_1=Sum(I=1,\ldots,100)(X_1I)^2+10SIN(3.1416*3/2/IX_I)$
$+10SIN(5*3.1416*3/2/IX_I)$
$+10SIN(9*3.1416*3/2/IX_I)+10SIN(13*3.1416*3/2/IX_I)+10SIN(17*3.1416*3/2/IX_I)$
$+10SIN(21*3.1416*3/2/IX_I)+10SIN(25*3.1416*3/2/IX_I)+10SIN(27*3.1416*3/2/IX_I)$
$+10SIN(33*3.1416*3/2/IX_I)+10SIN(41*3.1416*3/2/IX_I)+10SIN(81*3.1416*3/2/IX_I)$
where (<=0):
$EE_j=X_I-200$, j=1,...,100
Solution:
$EE_1=-8999.999$,
$X_1=0.999, X_2=1.999, X_3=2.999, X_4=3.999, \ldots,$
$X_{99}=98.999$, $X_{100}=99.999$

N_7. Equation with multi-extreme goal function and non-convex and non-linked goal function.

Minimize:
$EE_1=(SIN(X_1)+SIN(7X_1))$
where(<=):
$EE_2=(COS(3.22X_1)+1)$

[4] Chichenadze V.K., "Solution of non convex nonlinear optimization problems", Moscow "Science", Moscow, 1983

$EE_3=(X_1-10)$.
Solution:
$EE_1=-0.5928$, $X_1=4.879$, $EE_2=0.00000321$, $EE_3=-5.12096$

N_ 8. Equation with multi-extreme goal function and non-convex and non-linked goal function

Minimize:
$EE_1=0.001((X_1-10)(X_1-1)(X_1-2)(X_1-3)*(X_1-4)(X_1-5)(X_1-6)(X_1-7)(X_1-8)(X_1-9)+COS(17X_1)(X_2-5)(X_2-6)$
where (<=0)
$EE_2=(SIN(X_1+X_2))^2$
Solution:
$EE_1=-42.934069$, $X_1=1.29246267$, $X_2=11.2795776$, $EE_2=0.0000321$

N_ 9. Equation with multi-extreme goal function.

Minimize:
$EE_1=(X_1-22)^2+(SIN(13X_1)-3))^4$
where (<=0):
$EE_2=-X_1$
Solution:
$EE_1=0.00022543$, $X_1=21.98646728$
Control:
$EE_1=0.000354$, $X_1=21.99$

N_ 10. Equation with multi-extreme goal function.

Minimize:
$EE_1=22SIN(0.001X_1)(X_1-5)^2(1/(X_1+4))(SIN(13*X_1)+COS(20*X_1)$
where (<=0):
$EE_2=-X_1$
Solution:
$EE_1=-35.2194$, $X_1=8.6238524$
Control:
$EE_1=-35.143360$, $X_1=8.62$

N_ 11. Equation with multi-extreme goal function

Minimize:
$EE_1=3(SIN(4(3+X_1)))^2+(SIN(4+X_1))^{12}+(X_2-22)^2+(SIN(13X_2-3))^4$
$+SIN((X_3+SIN(X_3)+4)^2/(10+EXP(SIN(33X_3))))$
$+22SIN(0.01X_4)(X_4-5)^2(1/(X_4+4))*$
$+(SIN(13X_4)+COS(20X_4))+100(X_6-X_5^2)^2+(1-X_5)^2+(X_7+10X_8)^2+5(X_9-X_{10})^2$
$+(X_8-2X_9)^4+10(X_7-X_{10})^4$
where :
$X_1,X_2,\ldots,X_1 >= 0$
Solution:
$EE_1 = -36.07$, $X_1= 1.3200E-001$
$X_2= 22.20$, $X_3= 16.035$, $X_4= 8.6238$, $X_5= 1.030$, $X_6= 1.0622$
$X_7= 4.6579E-009$, $X_8= 4.65E-009$, $X_9= 5.3838E-003$, $X_{10}= 5.3836E-003$

N_ 12. Equation with multi-extreme goal function.

$EE1=3(SIN(4(3+X_1)))^2+(SIN(4+X_1))^{12}$
$EE2=(X_2-22)^2+(SIN(13X_2-3))^4$
$EE3=SIN((X_3+SIN(X_3)+4)^2/(10+EXP(SIN(33*X_3))))$
$EE4=SIN(0.01X_4)*(X_4-5)^2*22(1/(X_4+4))(SIN(13X_4)+COS(20X_4))$
IF $X_4 > 10$ then EE4=1
IF $X_4 < 9$ then EE4=EE4*10
$EE5=100*X_6-X_5^2)^2+(1-X_5)^2$
$EE6=(X_7+10X_8)^2+5(X_9-X_{10})^2+(X_8-2X_9)^4+10(X_7-X_{10})^4$
$EE7=4/3(X_{11}^2-X_{11}*X_{12}+X_{12})^2)$
IF EE7 < 0 then EE7=0
Minimize:
$EE(1)=EE1+EE2+EE3+EE4+EE5+EE6+EE7^{*0.75}-X_{13}$
where (<=0):
$EE_2=X_{13}-2$
$EE_{16}=-X_{11}-X_{12}-X_{13}$
Solution:
EE(1)=-35.088, $X_1=1.32129E-001$, $X_2=22.6815$, $X_3=4.359E-002$, $X_4=8.623$
$X_5=1.041$, $X_6=1.085$, $X_7=3.4943E-002$, $X_8=0.00$, $X_9=8.978625E-002$
$X_{10}=8.74408E-002$, $X_{11}=3.36281E-003$, $X_{12}=1.25656E-003$, $X_{13}=1.43497$
$EE_2=-5.6502E-001$, $EE_{16}=-1.43959$

N_ 13. Equation with multi-extreme goal function and non-convex and non-linked goal function.

W=35, R=7, T1=2.
XX=X1^0.5+(X2/X1)^0.5+(64/X2)^0.5
Minimize:
EE1=ABS(XX-3-(W-2*SIN(0.5*SIN(XX^0.3)))/R*T1*0.5+(2+(SIN(22*XX^3))^2)))
where (<=0):
EE2=(SIN(3.1416*X1))**2-0.1
EE3=(SIN(3.1416*X2))**2-0.1
EE4=(X1+1E-3)-X2
Solution:
EE1= 0.00, X1=1.999, X2=5.0906, EE2= -0.099, EE3= -0.021, EE4= -3.0904

N_14. Begin = 17:48:29, end = 17:48:42.

Queries with multi-extreme goal function.
XX=ABS((X1^0.9+X2^0.8)/(1.1+COS(X3^2.3+X4^1.3)))+Sum(j=1,...,100)(0.1*J*Xj^(1/j))
Minimize:
EE1=EXP(SIN(10*XX)*SIN(123+XX^0.73-3*XX^2.13+XX^1.55))
where (<=0):
EE2=ABS((X1^0.9+X2^0.8)/(1.1+COS(X3^2.3+X4^1.3)))-3
Solution:
EE(1) = 0.3679, X1,...,X100 = 1.

The Entity-Relation-Problem (ERP) Model for General MIS

Yanzhang Wang
Institute of Systems Engineering
Business School, Dalian University of Technology
Dalian, 116024 P. R. China, Email: yzwang@dlut.edu.cn

ABSTRACT – An adequate model for system structure is very important in the design and development of information systems, especially for large-scale business and financial management systems. With rapid changes and advancement in computer hardware and software techniques used for management information systems (MIS), demands on information are also rapidly expanding and system requirements are constantly changing. This implies that systems must be generalized and can be reconfigured. In this paper a general-purpose structural model for information systems is proposed, based on paradigm of the entity-relation-problem (ERP) knowledge representation. This model classifies all information of a system as entities, relations and problems. Under ERP, any MIS can be decomposed into three categories of management tasks: entity management, relation management, and problem management. They can be independently managed and reconstructed according requirements. Using the REP model, a number of real-world general-purpose MIS have been developed that demonstrate the efficacy of the proposed method.

Key Words: MIS, structural model, knowledge representation, entity, relation, problem, data bases and data warehouses.

1. INTRODUCTION

As electronic and computer techniques and other information techniques are rapidly changing and advancing, the support environments of information systems are renewed unceasingly. However, demands on information for social or economical purposes are quickly expanding and changing as competition intensifies. In the world, especially in China, many MIS had been or are being developed, but only a few of them are successful. Some need modifications right after building, making MIS expensive and less efficient in practice.

In our analysis of practical MISs, it has been found that a good structural model of must be generalized and easily reconstructed to satisfy the changes in techniques and demands. In order to obtain this good structure, a system must be designed on the basis of general characteristics of engineering, technologies, as well as social and economic bases, rather than on any special use. It must also be modular for easy reconstruction to fit the changes in techniques or demands. Many different system structures have been studied [1-11]. But most of them can only be adapted to a narrow field, or tied to specific database technologies, short of capturing the macro-structure of a system and the demand changes.

A knowledge representation system, the Entity-Relation-Problem (ERP) approach, had been proposed in [12]. This system classifies knowledge into two types: knowledge on objective system (system knowledge) and knowledge on manager's subjective behaviors (problem knowledge). In the ERP system, the Entities-Relations (ER) describes the objective knowledge, and the Problem (P) describes the problem knowledge. This paper gives a general structural model of MIS based on the ERP system. First, the basic mathematical descriptions of ERP system are reviewed. Then the model for entities, relations and problems are introduced. Finally, a few general-purpose MISs are described that are developed using the EPR model to show its adaptability and flexibility.

2. BASICS OF THE ERP SYSTEM

The ERP model is a general knowledge system to represent entity-relation-problem [12]. It is independent of any special application field. The major property of ERP is modularity, in which the representation of entities, relations and problems can be divided or composed flexibly.

(1) Description of Entity

Let e represent the concept of an entity, and A_e denote an attribute set of the entity, namely

$$A_e = \{ a_1, a_2, ..., a_q \}$$

Here a_i ($\forall i \in \{1, 2, ..., q\}$) is the ith attribute of the entity **e** of a system **S**, it may be a string of characters describing the concept of the attribute, or a flag value of the attribute; q is the number of the attributes related to the entity.

Let E stand for the power set on the attribute sets of all entities considered for an objective system, and E is called it the entity set of the system,

$$E = \{e_i\},\ i = 1, 2, ..., p$$

(2) Description of Relations

Denote E the entity set of a system S, then a relation r on E is defined as

$$r = \{O, I, B, C\}$$

where C is the attribute set of the relation r; O a related entity set, $O \subseteq E$ and $O \neq \phi$; I is a relating entity set, $I \subseteq E$ and $I \neq \phi$; B is a relation matrix, namely

$$B = [b_{ij}]$$

and

$$b_{ij} = \begin{cases} 1 & \text{if } i\text{th \& } j\text{th attributes are related} \\ 0 & \text{otherwise} \end{cases}$$

Furthermore, a relation set R of system S can be constructed as

$$R = \{r \mid r = \{O, I, B, C\}, I \text{ and } O \subseteq E\}$$

Thus an objective system S can be described by entity and relation set E and R as follows

$$S = \{E, R\}$$

(3) Description of Problems

For a given objective system S

$$S = \{E, R\},$$

a problem p on S can be defined as

$$p = \{O, I, X, R_p, C\}$$

where O is a goal entity set, $O \subseteq E$ and $O \neq \phi$; I is the condition, input, or control entity set, $I \subseteq E$ and $I \neq \phi$; $X \subseteq E$ a related entity set; R_p a relation set on the problem, $R_p \subseteq R$; C an attribute set of the problem.

3. THE MACRO STRUCTURE AND ENVIRONMENT OF MIS

Using the ERP model, the macro structure and environment of an MIS is depicted in Figure 1. In the environment of an MIS, decision making or management is the basic goal or service. The *entity* and *relation* information is the reflection and description of the intrinsic attributes and mechanisms of the MIS, and the *problem* information is the reflection of subjective activities in decision making and management. This means that entity and relation depend mainly on objective activities, and it does not change by the subjective activities of managers, whereas the problem information is dependent on the subjective activities and it changes very often.

Relations are the reflections or representation of inherent relations between attributes or units in a MIS. As we have more in-depth understanding of the environment of an MIS, requirements on management quickly increase. Current models in data structures or management routines do not satisfy this change in requirements. They need to separate relations from information to make specialized management on relation information.

In large-scale MIS, macro structure should consist of modules for entity, relation and problem management. As is well known, RDB (Relation Database) technique is based on ER knowledge, but it does not separate relation information from ER, and is not suited to problem management.

OODB (Object-Oriented Data Base) supports problem management well, but it can not conveniently share and reconstruct entity and relation information between objects because it encloses data and relations in objects. Data warehouse techniques are general-purpose, so it can suit wide application requirements. This paper does not discuss OODB, rather it studies the problem of how to build ERP models using these techniques at macro levels. In fact, OODB is the result of the increasing demands on relation and problem information, and data warehouses are developed to generalized information service. They also show that the separation of relations and problems from the systems itself is rather necessary.

4. MANAGEMENT SYSTEM STRUCTURE OF ENTITY INFORMATION

Entity information is the descriptions or indications of attributes, units or elements in objective systems, and it is denoted by a set ***E*** defined in the ERP model. Since ***E*** is a power set of all entities, a hierarchical structure is suggested, as shown in Figure 2. Here the subordinative relations between entities are applied to management, and these hierarchical subordinative relations are consistent with the understanding of objective systems. To this structure, RDB and data warehouse techniques can be directly applied.

5. MANAGEMENT SYSTEM STRUCTURE OF RELATION INFORMATION

Relation information is the descriptions of inherent relations between attributes, units or elements in an objective system. According to the definition of relation knowledge in the ERP, relation information or knowledge can be represented by a set ***R***.

For a relation $r \in R$, we have

$$r = \{ O, I, B, C \}.$$

Based on this subset, a data structure shown as Figure 3 can be designed to manage the relation information. Here ***O*** and ***I*** can be the formal parameters or code sets on set ***E***, ***B*** a relation matrix in sparse form, and ***C*** a characteristic set of ***r***. For complex cases, ***B*** may be fuzzy and may include some operators, such as identification and simulation operator, to handle relations of a system.

Relation information management uses general operations, such as addition, deletion, modification, and inquiry, as well as separation and integration. Although all the RDB or OODB techniques can be used to implement the management of relations in ERP, RDB offers higher efficiency in inquiry, and OODB offers more convenience data processing.

6. MANAGEMENT SYSTEM STRUCTURE OF PROBLEM INFORMATION

Problem information refers to the requirements on information. In the procedure of remodeling objective world, managers or decision makers solve different problems by using various information about the problems, thereby forming the demands on information. For efficient services of information, the supply of information should be organized by problems.

Based on ERP knowledge representation, for a given objective system $S = \{ E, R \}$, a problem ***p*** on ***S*** can be described as

$$p = \{ O, I, X, R_p, C \}.$$

This set ***p*** can be used to build the management structure of the problems. Here sets ***O***, ***I*** and ***C*** are closely related to the problems, and the information about them is obtained from interacting with decision makers. The information on ***X*** and R_p may be automatically generated from ***S*** information.

Figure 4 shows the structure of a problem management system, where the problem forming is included, and interface with users is strengthened. It is easy to see that OODB and OOP techniques can be easily implemented.

7. APPLICATIONS

In accordance with the idea of the structural model proposed above, a generalized MIS has been developed in which system *management* is combined with system *generation*. This model satisfies the changing requirements on information management in government and enterprises, especially in macro or large-scale cases. Up to now, more than one 100 departments in government or enterprises have adopted this model in macro management, and some of them have been in use for more than 5 years with considerable social and economical benefits [10][11][12].

This model has a powerful subsystem for entity and relation management, where the management pattern is consistent with the hierarchical knowledge of objective systems. In particular, a *special Executive Command System (ECS),* together with a problem management module, has been developed with multimedia and touch screen. In ECS, problems can be easily formulated by executives with minimum training. Moreover general-purpose tools supporting data processing can be introduced to ECS, allowing processing of tables, graph, statistics and so on.

8. CONCLUSIONS

The structural ERP knowledge representation system has the following major advantages:

(1) The model is generalized and can be used to configure general-purpose MIS and to satisfy varying system requirements in practice.

(2) It has a modular structure, in which entity, relation and problem information can be divided or composed flexibly. This leads to flexible configuration of general MIS.

(3) This model supports data analysis and in-depth decision making, because it can be easily connected with quantitative or qualitative analysis based on relation and problem information.

(4) A visual and reusable software system can be developed for system management and system generation.

9. ACKNOWLEDGEMENT

The author acknowledges the support by the National Natural Science Foundation of China (69674042) and China's Fund for Outstanding Researchers for the 21st Century.

10. REFERENCES

[1] Mark Astley and Gul A. Agha, " A visualization model for concurrent systems", Information Sciences, Vol. 93, Nos. 1/2, 1996.

[2] G.H.W.M. Bronts, S.J. Brouwer, C. L. j. Martens and H.A. Proper, "A unifying object role modelling theory", Information Systems, Vol. 20, No. 3, pp.213-235, 1995.

[3] D. Avison and G. Fitzgerald, "Information systems development: methodologies, techniques and tools", Blackwell Scientific Publications, Oxford, UK, 1988.

[4] P. Chen, "The entity-relationship model: toward a unified view of data", ACM Trans. on Database Systems, 1(1); 9-36, 1976.

[5] R. Elmasri, J. Weeldreyer & A. Hevner, "The category concept: An extension to the entity-relationship model", Data & Knowledge Engineering, 1:75-116, 1985.

[6] Otto Rauh and Eberhard Stickel, "Standard transformations for the normalization of ER schemata", Information Systems, Vol. 21, No.2, pp. 187-208, 1996.

[7] Robert T. Chi and Efraim Turban, "Distributed intelligent executive information systems", Decision Support Systems 14, 117-130, 1995.

[8] Robert B. Jackson, David W. Embley and Scott N. Woodifield, "Developing formal object-oriented requirements specifications: a model, tool and technique", Information Systems Vol.20, No. 4, pp273-289, 1995.

[9] Juan Miguel Medina, Olba Pons and Maria Amparo Vila, "GEFRED: A generalized model of fuzzy relational Databases", Information Sciences 76, 87-109, 1994.

[10] Zhongtuo Wang, "Applications of computers in business ---- a new system paradigm", Publsihing house of Dalian University of Technology,1994.

[11] Yanzhang Wang, Bai Yang and Zhongtuo Wang, "Some considerations about decision making and IIIDSS", Proceedings of International Conference on Information & Systems, International Academic Publishers, pp.674-677, 1991.

[12] Yanzhang Wang, Bai Yang, and Zhongtuo Wang, "ERP System of knowledge representation for DSS", Proceedings of International Conference on Information & Systems, International Academic Publishers, pp.852-856, 1991.

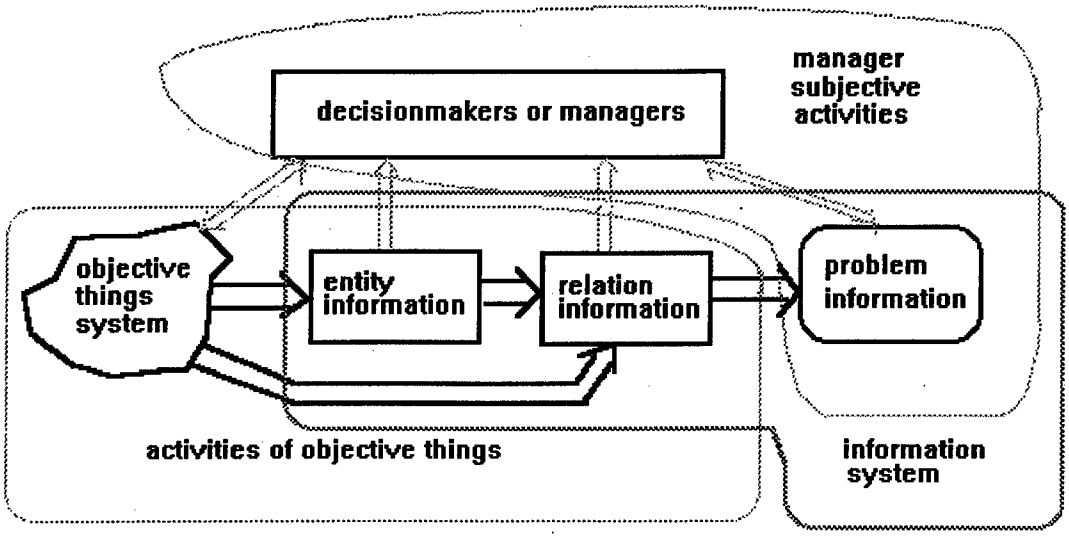


Fig. 1 Macro structure and environment of information systems

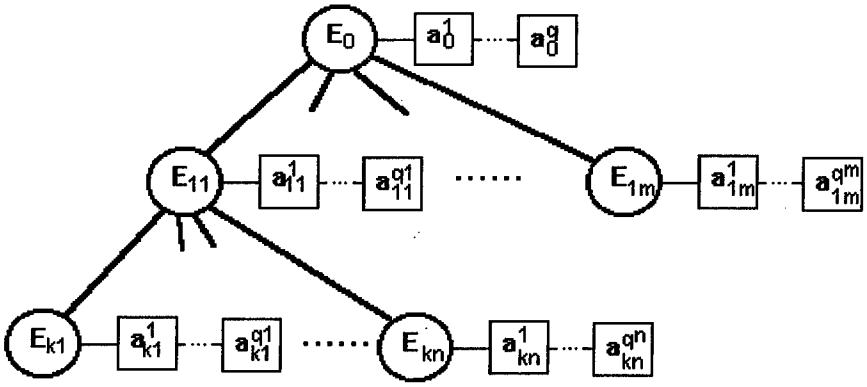


Fig. 2 the hierarchic structure of entity information

record	set O	set I	matrix B	attribute set C
record 1	O_1	I_1	B_1	C_1
record 2	O_2	I_2	B_2	C_2
...	...	...	...	...
record n	O_n	I_n	B_n	C_n

Fig. 3 the structure of relation information

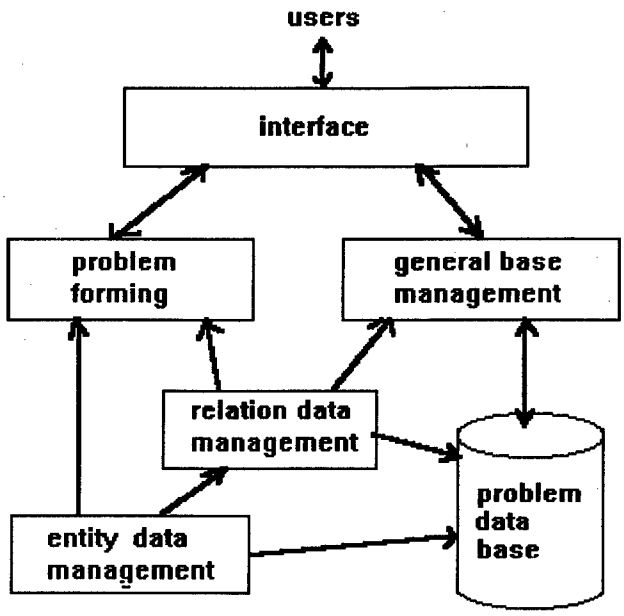


Fig. 4 the structure of problem information management

Session RB2
Multisensor Target Tracking and Recognition Of Small-to-Medium Sized Targets
Chair: Oliver Drummond
Consulting Engineers, Culver City, CA, USA

ARTAS: An IMM-based Multisensor Tracker

R.A. Hogendoorn, C. Rekkas* and W.H.L. Neven**

Summary
ARTAS (an acronym for ATM Surveillance Tracker and Server) is the new operational Surveillance Data Processing and Distribution (SDPD) system at Amsterdam Airport and is being evaluated at different sites in France, The Netherlands, Portugal and the UK. ARTAS was developed by Eurocontrol, in co-operation with a consortium of industrial partners, in order to be used as a basis for the development of SDPD systems in Europe. The ARTAS system consists of a tracker, responsible for maintaining up-to-date target state vectors, a server, which handles client subscriptions (e.g. from the ATC display system) and delivers the target state vectors to these clients and a Man-Machine Interface/Supervision module, for system control and air-situation display. An ARTAS system co-operates with adjacent ARTAS systems by exchanging target state vector information.
The main features of the ARTAS Tracker are

- tracking with up to thirty radars (PR, SSR or CMB)
- on-line estimation of the radar systematic errors
- on-line estimation of radar false plot maps
- on-line estimation of the radar accuracy and coverage
- high-accuracy position and velocity-vector estimation
- responsiveness to target manoeuvres
- insensitivity to clutter
- target type identification

All these features are realised through the use of state-of-the-art estimation and identification algorithms, such as the IMM (Interacting Multiple Model) algorithm and Dempster-Shafer reasoning, and an object-oriented architectural design.

Track Data Server

ARTAS is designed as a track data server. Track data users can subscribe to a certain service and receive the track data in ASTERIX format via a local-area or wide-area network. Users can be ATC centres, flightplan data processing systems (FDPS), air-traffic flow management units and so on (figure 1). Each user can have a dedicated service, taking into account requirements with respect to data contents and update frequency. An ARTAS unit also receives its input data from the radars via the local -area or wide-area network. Furthermore, an ARTAS unit can communicate via the network with other, adjacent, ARTAS units in order to provide a continuous air-picture to its users. Track data from adjacent units is used to accelerate the initiation of tracks at the border of the unit's own domain of interest (DOI) and to smooth the transition of a track from one unit's DOI to another unit's DOI. Finally, when there is sufficient coverage of the own unit's DOI by adjacent ARTAS units, the adjacent ARTAS units can take over the surveillance in case of an own unit failure. Thus, enhancing the overall reliability of the surveillance.

A prime requirement for handling multisensor data is the ability to cope with sensor alignment errors, i.e. systematic radar errors like position bias, range- and azimuth bias, but also time-stamping bias and transponder-delay error. The latter is an example of a, so-called, micro-error: a systematic error that depends on the object being tracked. The former errors are macro-errors; they only depend on the sensor involved. Unfortunately, both macro- and micro-errors may change in time, due to e.g. changing atmospheric conditions and radar maintenance. Therefore, the ARTAS Tracker contains modules that dynamically estimate and correct both the macro- and micro-errors.

* National Aerospace Laboratory NLR
PO Box 153
8300 AD EMMELOORD
The Netherlands
e-mail: hogend@nlr.nl, neven@nlr.nl
* Eurocontrol Brussels
96, rue de la Fusée
B-1130 BRUSSELS
Belgium
e-mail: christos.rekkas@eurocontrol.be

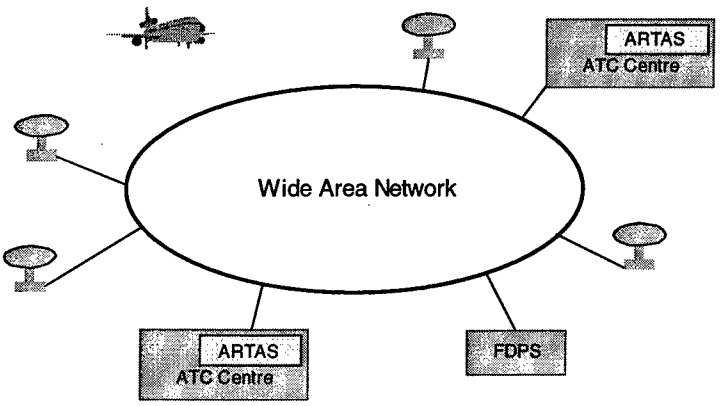


Figure 1.The ARTAS Environment

Another requirement for handling multisensor data is a proper treatment of coordinate transformations. This becomes a more obvious problem when the size of the system area becomes large. ARTAS uses WGS84 as a reference system. Measurement processing and track update processing are done in local Cartesian systems, such that the error, induced by coordinate transformations, always stays below a required level of a few meters.

The internal structure of an ARTAS unit is shown in figure 2. The Router Bridge is the interface to the external network. It pre-processes the incoming radar data, i.e. it performs format checks and sectorisation of the plot data and keeps track of the operational status of the radars. The Server is responsible for the handling of ARTAS user requests and the distribution of the track data, according to the different user services. Furthermore, the Server is responsible for Track and Service continuity across the borders of the DOIs of adjacent ARTAS units, i.e. track data users are not aware the fact that targets cross this border. The simplest service that is provided is a regular broadcast of all track data. MMI/Supervision is the man-machine interface and supervision unit. It provides a basic display of the unit tracks and control functions for the ARTAS unit. The Tracker, finally, is responsible for keeping an up-to-date air picture. An ARTAS unit consists of two identical chains of Router Bridge/Tracker/Server/MMI/Supervision subunits. The Trackers in both chains operate in a multiple-computation redundancy mode; that is, there is a master and a slave Tracker that both perform the same processing, except that the slave Tracker does not provide any output.

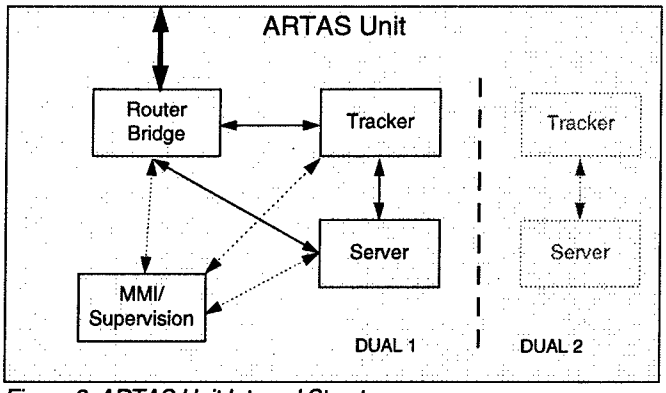


Figure 2. ARTAS Unit Internal Structure

Instead, the slave Tracker performs some additional processing to keep master and slave in synchronisation.
All the ARTAS subunits run on off-the-shelf hardware and are programmed in ADA, except for the MMI, which is programmed in C++.

The ARTAS Tracker

Basically, the task of the tracker is to provide estimates of the aircraft state vector for each target in the domain of interest of the ARTAS unit. It makes use of, maximally, 30 sensors. In the operational system, the sensor types are primary radar (PR) and secondary surveillance radar (SSR). A prototype ARTAS2 tracker, which is an extension of the operational ARTAS tracker, additionally handles aircraft-derived data, received either through Mode-S or through automatic dependent surveillance (ADS).

Track continuation uses the reports of all available sensors to estimate the state of a target. Each track extrapolation/update cycle is based on the reports of a single sensor, though. Subsequent cycles, however, may be of entirely different sensors. Prior to the track update, all the relevant reports are corrected for micro-errors (systematic errors that vary from target to target) and slant-range effects. Track continuation is discussed in more detail below.

The integration of aircraft-derived position, speedvector and roll-angle information at the tracking filter level results in a clear performance improvement. This was demonstrated to Eurocontrol and European national administrations in February 1999, using the ARTAS2 prototype tracker. Figures 3 and 4 show the decrease of the course error after a turn, when aircraft-derived data is used (simulated Mode-S radar data; averaged for 25 tracks)

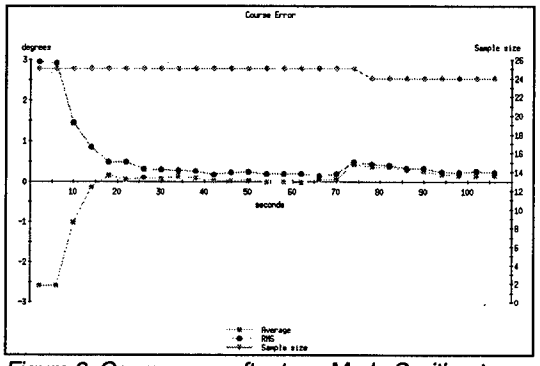


Figure 3. Course error after turn; Mode-S without aircraft-derived data

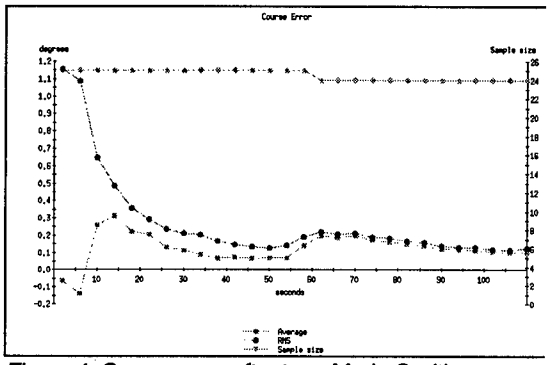


Figure 4. Course error after turn; Mode-S with aircraft-derived data

As explained earlier, all sensors and all tracked objects have their own local Cartesian system that may change in time when objects move. The effect of this is clearly visible in figures 3 and 4; the increase in course error after about 60 seconds is due to relocation of the track-local Cartesian system.

Track continuation uses the reports of all available sensors to estimate the state of a target. Each track extrapolation/update cycle is based on the reports of a single sensor, though. Subsequent cycles, however, may be of entirely different sensors. Prior to the track update, all the relevant reports are corrected for micro-errors (systematic errors that vary from target to target) and slant-range effects. Track continuation is discussed in more detail below.

Track initiation is done based on the reports of single sensors only. It is based on multiple-hypothesis tracking (MHT) and is done retrospectively [3]. Considering the fact that a new target generally enters the coverage of the Tracker with only mono-radar visibility, the gain of a shorter track initiation delay did not warrant the additional complexity of a multiradar initiation in a civil ATC environment. This trade-off is not valid in a military environment, though. It is foreseen to extend the track initiation to a multisensor initiation in the scope of an on-going evaluation.

The ARTAS Tracker maintains aircraft and non-aircraft tracks since, in many cases, the best way of dealing with anomalies, like reflections and sidelobes, is to track them and to identify them as being non-aircraft. To that end, the ARTAS Tracker contains a track type identification module, which identifies tracks using Dempster-Shafer reasoning [4]. The criteria, used in the track type identification, are based on radar environment characteristics, target behaviour and a set of models for specific anomalies, like reflections and sidelobes. An advantage of Dempster-Shafer reasoning

is the ease with which additional criteria, like target signature information, can be incorporated into the classification process.

Track Continuation

For the ARTAS Tracker, a Bayesian approach to track continuation was adopted. This approach did prove to yield a high-performance tracker, as NLR experience with the JUMPDIF prototype tracker has shown [1].

Basically, there are four major problems that occur during track continuation

1. Non-linear aircraft dynamics during a turn
2. The association of measurements with existing tracks
3. The occurrence of outlier measurements (non-Gaussian measurement noise)
4. Sudden starts and stops of manoeuvres

For each of these problems, adequate solutions were already developed for the JUMPDIF prototype [2]; the result, an Interacting Multiple-Model Probabilistic Data-Association (IMMPDA) algorithm with Extended Kalman Filters (EKF) [1]. This four-mode IMMPDA EKF was used in extensive performance tests. The results of these performance tests were used as a basis for the ARTAS Tracker performance requirement specification. A number of improvements, with respect to the JUMPDIF tracker, were made in the ARTAS Tracker, though.

For target resolution situations, new joint probabilistic data-association (JPDA) algorithms were developed [3] that avoid the track coalescence property of conventional JPDA, while performing considerably better than the probabilistic data-association (PDA) algorithm in situations with targets closely together.

In the ARTAS2 tracker, the IMMPDA algorithm was extended to incorporate aircraft-derived data. This extension is called ADD-IMMPDA. Furthermore, the IMM track extrapolation was adapted to handle the situation where very accurate position reports are received with a low sampling rate. This may be the case when aircraft position reports are obtained by means of differential GPS.

The ARTAS Tracker is required to track targets down to zero groundspeed. For these targets, a simplified two-model (manoeuvring flight, straight flight) IMMPDA filter is developed.

Initially [2], a two-model (climb/descent, level flight) IMMPDA filter for SSR mode-C measurements was developed. In the ARTAS Tracker this filter was improved by a three-model (climb, descent, level flight) IMMPDA filter in order to be more responsive to changes in the rate of climb/descent. Furthermore, two algorithms to estimate the target altitude in absence of SSR mode-C information were implemented. One algorithm, Triangulation, is discussed in more detail below. Although not as accurate as mode-C based height, the performance of the triangulation algorithm often is surprisingly good. Another algorithm, Height-from-Coverage, uses the assessed coverage of all radars that detect or do not detect the target, to calculate a height interval for the target. This is used as a fallback in cases where neither mode-C nor triangulated height is available.

For centralised multisensor track continuation, a key problem is the accurate estimation and correction of systematic errors. The solution developed for the ARTAS Tracker is a dynamic estimation and correction of the macro- and micro-systematic errors of all involved measurements, before they are used within the track extrapolation/track update cycle. This essentially reduces the multisensor problem to a single-sensor problem. The time sequence of track extrapolation/track update cycles, obviously, contains track extrapolation/track update cycles for all the available sensors. The difference between cycles for different sensors is the use of a different measurement matrix for the Extended Kalman filters.

Figure 5 shows a track, departing from Schiphol airport that uses biased measurements from three different radars. Figures 6 and 7 show the ARTAS Tracker estimates of the groundspeed and SSR mode-C height of this track, respectively. Without an effective elimination of systematic errors, groundspeed and height would contain a substantial number of irregularities.

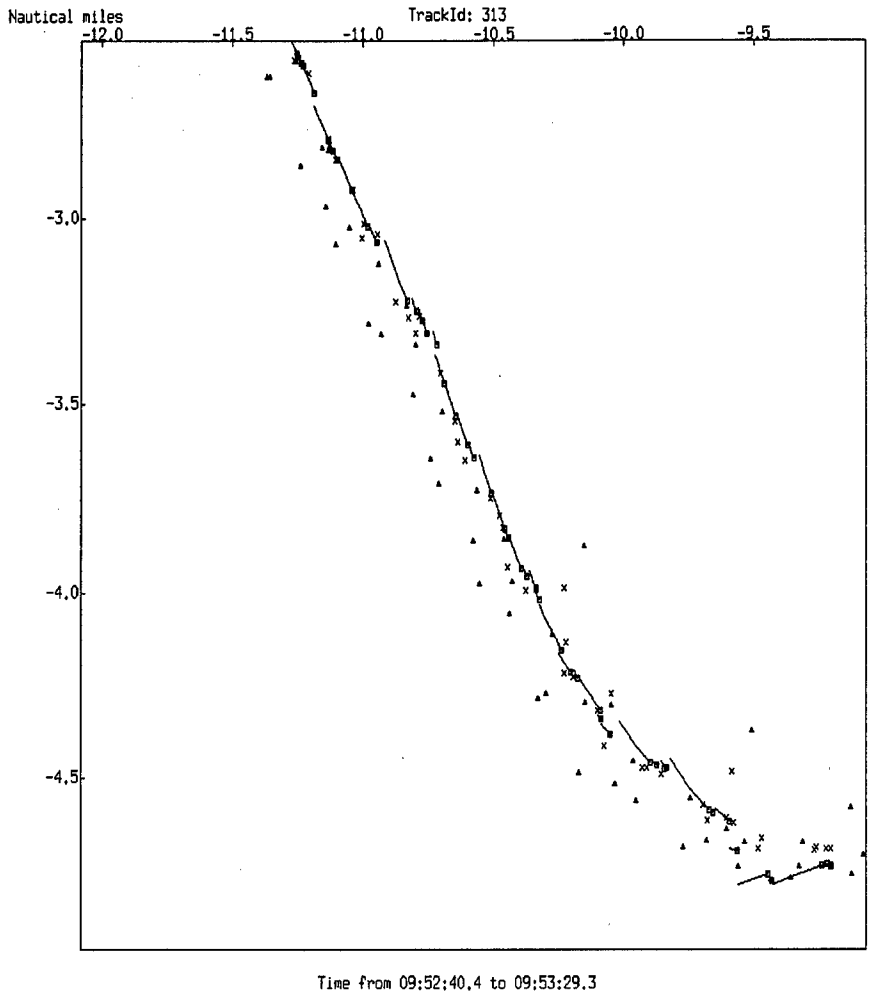


Figure 5. Departure from Schiphol airport, using biased measurements from three different radars (triangles indicate raw plots, crosses nearest-neighbour plot positions (corrected for the estimated radar biases) and squares the updated track position. The vectors indicate the predicted flightpath up to the next measurement instant).

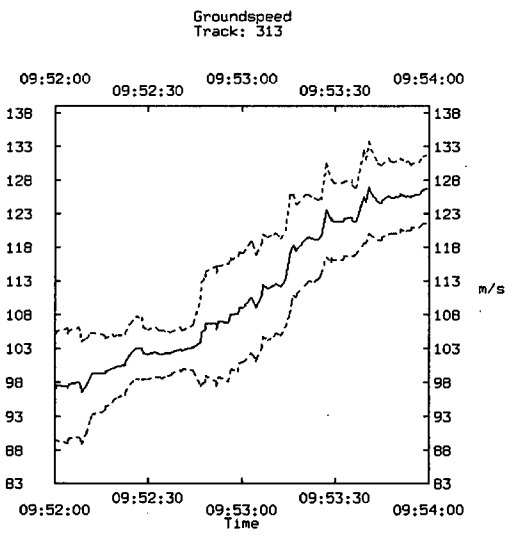


Figure 6. Track groundspeed estimate

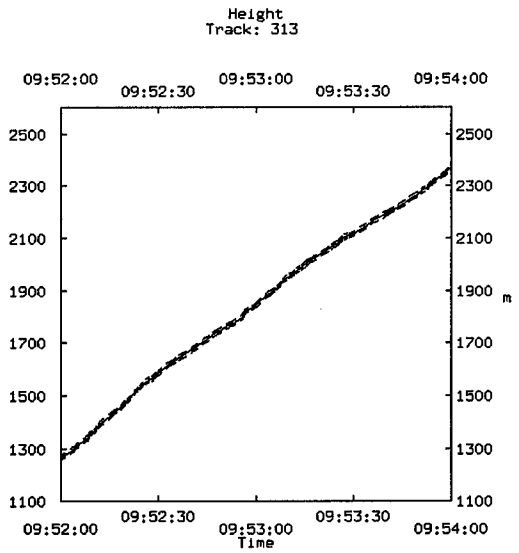


Figure 7. Track mode-C height estimate

Macro-Error Estimation

The ARTAS Tracker estimates the following (macro-) systematic errors:

- range bias
- azimuth bias
- range gain (a range error proportional to the range)
- antenna squint (non-verticality of the plane of the radar beam)
- verticality error (antenna rotation axis not perpendicular)
- time-stamping bias

The problem with dynamic estimation of the (macro-) systematic errors is that, in principle, the filter equations are coupled with the track continuation equations of the individual tracks. It is, of course, very well possible to make a selection of a small number of well-behaved tracks and to solve the resulting set of equations. In ARTAS, a different approach is taken [6], which decouples the equations for (macro-) systematic error estimation from the track continuation equations. Effectively, it comes down to a weighted integration of the innovations of all tracks and filtering these weighted results with a Kalman filter. Due to this decoupling, the filtering equations become independent of the individual track maintenance equations. This algorithm is implemented in the ARTAS Tracker. It uses a selection of non-manoeuvring tracks when it is necessary to save CPU-load without jeopardising the speed of convergence of the macro-error estimation process. Figures 8 and 9 show results of the (macro-) systematic-error estimation process on a 2-radar PR scenario.

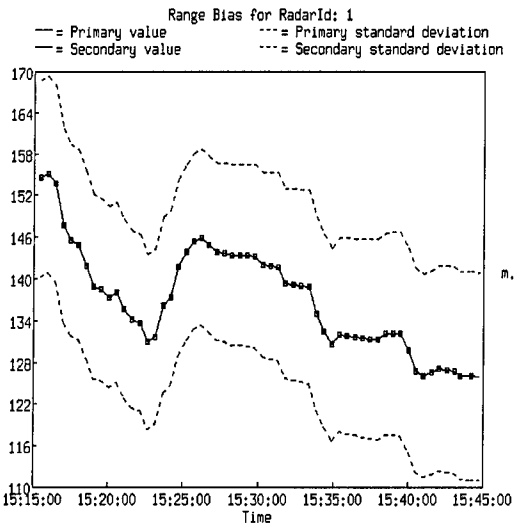


Figure 8. TAR estimated range bias

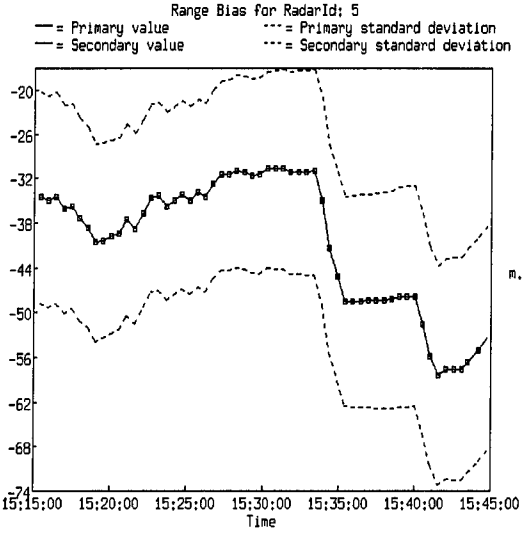


Figure 9. LAR estimated range bias

Triangulation-based Micro-Error Estimation

After estimation of the systematic radar errors that are radar-dependent only (macro errors), the track-related errors (micro errors) can be estimated. Within the ARTAS Tracker, these micro errors consist of the transponder delay error (i.e. the difference between the actual delay and the nominal value of 3 microsecond as specified by ICAO) and the geometric height, estimated from range-azimuth position measurements in a multiradar environment.

A general solution to this problem is to extend the state vector of an object with these components and to extend the corresponding extended Kalman filter equations accordingly. Since this is a very costly solution (in terms of CPU), we have looked for a robust method that is not coupled with the track continuation equations. In situations where an SSR radar has a co-located primary radar, a robust method to estimate the transponder delay error is to average the difference in range measurements of the two radars. In other situations, the transponder delay error and geometric height estimations are coupled.

Consider the situation that two non-co-located radars observe an object at the same moment in time. To perform triangulation, we use the difference between the projections of the plots to a common 2-dimensional Cartesian coordinate system (the track-local coordinate system) as the

innovation term in a Kalman-like filter update step for the estimation of the transponder delay error and the geometric height.

Since a simultaneous measurement of one object by two non-co-located radars is quite unusual, we perform a triangulation on the basis of a triplet of projected plot positions (under the condition that the track groundspeed and course are constant). The first and third projected positions are interpolated to the time of the middle plot.

The performance of this algorithm depends, among others, on the geometric configuration of the radars involved: the middle plot should be from a different radar than the other two plots, with a line-of-sight opposite to that of the other radars, and as close to the object as possible.

In figure 10, we see a part of a track from a live data collection. The recording was made for 3 secondary and 2 primary radars, but the Tracker was run with only the primary plot data. The track is flying at FL 290 (8840 m); the plots are not corrected for systematic radar errors. The estimate of the geometric height and the 1-sigma margin are given in figure 11; the initial estimate is 6000 m.

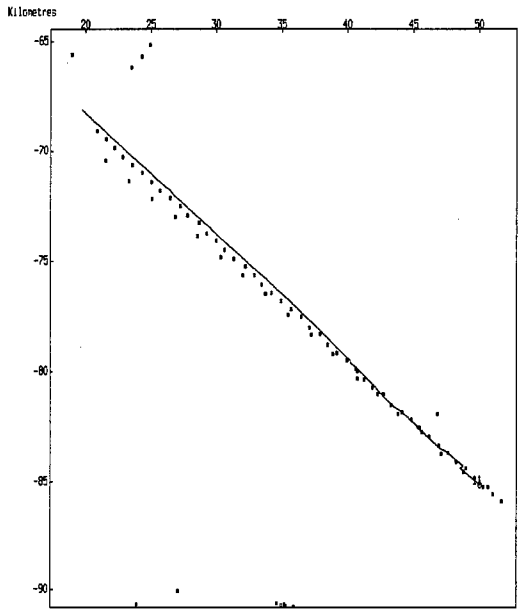


Figure 10. Track observed by 2 PR radars

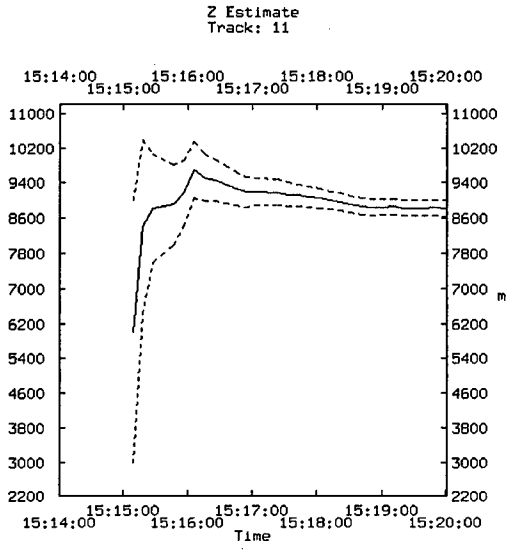


Figure 11. Triangulated height as function of time

Multisensor Environment Assessment

The use of aircraft-derived data increases the complexity of the multisensor situation enormously. In addition to the rather small set of radars with sensor characteristics that are, generally, well known, the sensors on-board each and every aircraft have to be taken into account. This creates two problems:

- the estimation of the on-board sensor characteristics;
- the estimation of additional micro-errors.

The ARTAS tracker already contains modules to estimate the radar sensor characteristics. These are part of the, so-called, Multiradar Environment Assessment (MREA). In ARTAS2, these modules will be extended to become a Multisensor Environment Assessment (MSEA).
New methods will have to be developed for the estimation of the large amount of additional micro-errors, such as time-stamping bias, drift in position and differences in atmospheric conditions (pressure altitude).

Conclusions

Adequate systematic error estimation is a pre-requisite for accurate multisensor tracking. In the ARTAS Tracker, several powerful methods are employed for the on-line estimation of both macro- and micro-systematic errors. These methods provide accurate estimates of the systematic errors as shown by a number of examples. By having accurate systematic error estimates, the

multisensor problem is essentially reduced to a time-sequential single-sensor problem, which is, obviously, much easier to solve.

The incorporation of aircraft-derived data in tracking increases the complexity of the systematic error estimation dramatically. New estimation methods will have to be developed to deal with this problem.

References

[1] H.A.P. Blom and Y. Bar-Shalom, The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficients, IEEE Trans. on Automatic Control, Vol. 33, 1988, pp. 780-783.

[2] H.A.P. Blom, R.A. Hogendoorn and B.A. van Doorn, Design of a Multisensor Tracking system for Advanced Air-Traffic Control. In: Y. Bar-Shalom (ed.), Multi-Target-Multisensor Tracking: Applications and Advances, Vol. II, Artech House, 1992.

[3] E.A. Bloem and H.A.P. Blom, Joint Probabilistic Data Association Methods avoiding track coalescence, Proc. 34th IEEE Conf. on Decision and Control, December 1995, pp. 2752-2757.

[4] R.A. Hogendoorn, H.A.P. Blom, Bayesian Track Initiation by time-reversion of Trajectory Models, Proc. IEEE Int. Conf. on Control and Appl., April 1989, Jerusalem, paper WA-1-5.

[5] F. Voorbraak, A Computationally Efficient Approximation of Dempster-Shafer Theory, Department of Philosophy, University of Utrecht, 1987.

[6] B.A. van Doorn and H.A.P. Blom, Systematic Error Estimation in Multisensor Fusion Systems, SPIE Conf. on Signal and Data Processing of Small Targets, April 1993, Orlando, pp. 450-461.

An Adaptive Bayesian Approach to Fusion of Imaging and Kinematic Data

B. Rozovskii
Center for Applied Mathematical Sciences
University of Southern California
Los Angeles, CA 900089-1113, U.S.A.

A. Tartakovsky
Center for Applied Mathematical Sciences
University of Southern California
Los Angeles, CA 900089-1113, U.S.A.

G. Yaralov
Center for Applied Mathematical Sciences
University of Southern California
Los Angeles, CA 900089-1113, U.S.A

Abstract - *The primary goal of this paper is to develop an approach to fusion of two streams of data – imaging and kinematic – for optimization of target identification. Specifically, we focus on the fusion of range-profile data (obtained from a high range resolution sensor) and kinematic information for observation-to-track assignment and target recognition. A more reliable target identification is possible due to the strong correlation between kinematic characteristics and range profile. Indeed, the range profile signal depends on the range to the target and its aspect angle, while the latter is related to the target velocity (via Euler's equation), thus yielding a strong correlation of both types of data with the aspect angle. Effective estimation of the aspect angle is therefore the key to successful target identification. The dynamics of the aspect angle is modeled by a Markov process with a switching parameter. The latter parameter describes transitions from one target maneuver to another. In this model the state process and the observation are nonlinear. This rules out application of standard methods of estimation based on Kalman filter and necessitates the use of a nonlinear filtering algorithm. The crucial part of the fusion and identification algorithm is the fully coupled optimal nonlinear filter for the aspect angle. This filter allows us to compute recursively joint unnormalized posterior distributions of the target class and aspect angle. Then specially designed adaptive sequential multihypothesis classification procedures, which exploit the optimal nonlinear estimates of the aspect angle for all classes, are used to identify targets of interest.*

Key Words: high range resolution sensor, range-profile data, kinematic data, nonlinear filtering, fusion of imaging and kinematic data, sequential identification.

1 Introduction

We propose a new method of target identification based on fusion of imaging and kinematic measurements. Our approach is fairly general, however in this paper for the sake of concreteness we concentrate on fusion of high range resolution radar (HRRR) imaging data (in the form of range profiles) and standard kinematic data (e.g. range, velocity, etc.). The final goal is to improve performance of target recognition. This problem was addressed by several authors [7, 10]. In particular, Libby and Maybeck [10] proposed a version of the dynamic programming method (the Viterbi-Larson-Peschon algorithm [9, 19]) to estimate the most probable "path" of aspect angles given both kinematic data and HRRR-profiles. This estimate is needed to compute an approximation to *a posteriori* probability of target class. The Libby-Maybeck algorithm is designed to utilize fixed size samples (i.e. it is "one-stage" or "batch" algorithm).

In contrast, we propose a sequential algorithm for joint target tracking and recognition by fusing of kinematic data and HRRR-profiles on the basis of optimal nonlinear filtering. The nonlinear filtering provides an accurate and robust recursive algorithm for estimation of aspect angles. These estimates serve as input data for an optimal multihypothesis sequential test for target identification. We remark that the dynamic programming method is time consuming and its computational complexity grows fast when the number of observations increases. Also, it is shown that the developed sequential identification algorithms two to four times faster than the best fixed sample size test.

2 Problem Formulation and Basic Mathematical Model

We consider a scenario where the target recognition algorithm must relate the target to one of M predetermined classes (hypotheses): $H_1, H_2, \ldots, H_M$ (e.g. H_1 = SU27, H_2 = MIG31, H_3 = F18, H_4 = A6, H_5 = A10, H_6 = UFO).

At each successive time $t_k, k = 1, 2, \ldots$, the input information for the identification algorithm consists of range profile measurements (target signature) $\boldsymbol{r}(t_k) = (r_1(t_k), \ldots, r_m(t_k))$, where $r_i(t_k)$ is the wave envelop in the ith range cell at time t_k, and the vector of kinematic parameters $\boldsymbol{z}(t_k)$ (target's velocity, position, etc.). Thus, the observed process $\boldsymbol{y}(t_k)=(\boldsymbol{r}(t_k), \boldsymbol{z}(t_k))$ consists of two different in nature components – range profile and kinematic measurements. Our objective is to incorporate these data in the target identification algorithm in an optimal way.

The data observed up to time t_k will be denoted $\boldsymbol{Y}_k = (\boldsymbol{y}(t_1), \ldots, \boldsymbol{y}(t_k))$, i.e. $\boldsymbol{Y}_k = (\boldsymbol{R}_k, \boldsymbol{Z}_k)$ where $\boldsymbol{R}_k = (\boldsymbol{r}(t_1), \ldots, \boldsymbol{r}(t_k))$, $\boldsymbol{Z}_k = (\boldsymbol{z}(t_1), \ldots, \boldsymbol{z}(t_k))$. Write also $\boldsymbol{\Phi}_k = (\boldsymbol{\phi}(t_1), \ldots, \boldsymbol{\phi}(t_k))$.

If the target belongs to class H_j, the range profile signal (target's signature) $\boldsymbol{r}(t_k) = \boldsymbol{r}_j(t_k)$ at time t_k is given by

$$\boldsymbol{r}_j(t_k) = \boldsymbol{S}_j(\boldsymbol{\phi}(t_k)) + \boldsymbol{N}_j(t_k), \qquad (1)$$

where

$$\boldsymbol{N}_j(t_k) = (N_{j,\lambda_1}(t_k), \ldots, N_{j,\lambda_m}(t_k))$$

is the noise component;

$$\boldsymbol{S}_j(\boldsymbol{\phi}(t_k)) = (S_j(\lambda_1, \boldsymbol{\phi}(t_k)), \ldots S_j(\lambda_m, \boldsymbol{\phi}(t_k)))$$

is the range profile signal of the target; $\boldsymbol{\phi}(t_k)$ is the target pose (aspect angle) at moment t_k; λ_i is the time lag to the ith range resolution element; m is the total number of range resolution elements.

There is a number of simulation tools that can be used to synthesize target signatures with different levels of fidelity: XPATCH, URISD, conditionally Gaussian model, etc. [7].

In the Bayesian framework, the decision making algorithms (sequential or non-sequential) are based on the posterior probabilities $P(H_j|\boldsymbol{Y}_k)$, $j = 1, \ldots, M$. If the prior distribution of classes is unknown, the adaptive version of the generalized likelihood ratio approach can be applied. In this case, the generalized likelihoods (averaged over the trajectories $\boldsymbol{\Phi}_k$) are replaced by their adaptive versions $P(\boldsymbol{Y}_k|\widehat{\boldsymbol{\Phi}}_k^j, H_j)$ where $\widehat{\boldsymbol{\Phi}}_k^j$ is an estimate of the aspect angle path (see Section 4 for more details). In either case (Bayesian and non-Bayesian) implementation of the identification procedure requires estimation of the sequence of aspect angles. In [10] the posterior probabilities of classes $P(H_j|\boldsymbol{Y}_k)$, which can be obtained by averaging of the joint posterior distribution $P(H_j, \boldsymbol{\Phi}_k|\boldsymbol{Y}_k)$ over $\boldsymbol{\Phi}_k$, are approximated by using (conditional) estimates $\widehat{\boldsymbol{\Phi}}_k^j = \arg\max_{\boldsymbol{\Phi}_k} P(\boldsymbol{\Phi}_k|\boldsymbol{Y}_k, H_j)$. The Larson-Peschon-Viterbi (dynamic programming) algorithm was used in [10] to compute these estimates.

A well known drawback of the dynamic programming approach to estimation in hidden Markov models (HMM) is that it does not have a sequential structure. For example, the optimal trajectory $\widehat{\boldsymbol{\Phi}}_k^j$ might differ substantially from the first k entries of the $\widehat{\boldsymbol{\Phi}}_{k+1}^j$. Another disadvantage of this approach is high computational complexity.

To overcome these drawbacks we propose to use the optimal nonlinear filtering (ONF) algorithm for HMM. In this approach, instead of computing $P(H_j, \boldsymbol{\Phi}_k|\boldsymbol{Y}_k)$ one computes the joint filtering density $P(H_j, \boldsymbol{\phi}(t_k)|\boldsymbol{Y}_k)$. In contrast to $P(H_j, \boldsymbol{\Phi}_k|\boldsymbol{Y}_k)$ the probability $P(H_j, \boldsymbol{\phi}(t_k)|\boldsymbol{Y}_k)$ allows for efficient recursive computation and the resulting identification algorithm can be implemented sequentially.

The relationship between kinematic and range profile data is very strong. In (1) the range profile signal depends on the range to the target and its aspect angle $\boldsymbol{\phi}$. On the other hand, the latter is related to the target velocity vector by the Euler's equation in the inertial coordinate system:

$$\dot{\boldsymbol{v}}(t) = \boldsymbol{Q}(t)\boldsymbol{v}(t) + \boldsymbol{f}(t), \qquad (2)$$

where $\boldsymbol{v}(t)$ is target velocity, $\boldsymbol{f}(t)$ is target acceleration and

$$\boldsymbol{Q}(t) = \begin{bmatrix} 0 & -q_3(t) & q_2(t) \\ q_3(t) & 0 & -q_1(t) \\ -q_2(t) & q_1(t) & 0 \end{bmatrix},$$

where $\boldsymbol{q}(t) = (q_1(t), q_2(t), q_3(t))^T$ is target angular velocity. It is related to the aspect angle $\boldsymbol{\phi}(t) = (\phi_1(t), \phi_2(t), \phi_3(t))^T$ as follows [5, 6]

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} \dot{\phi}_1 \cos\phi_2 \cos\phi_3 - \dot{\phi}_2 \sin\phi_3 \\ \dot{\phi}_1 \cos\phi_2 \sin\phi_3 + \dot{\phi}_2 \cos\phi_3 \\ -\dot{\phi}_1 \sin\phi_2 + \dot{\phi}_3 \end{bmatrix},$$

where the dependence of ϕ_k and q_k on t is omitted. In other words $q_1(t), q_2(t)$, and $q_3(t)$ are rates of change of the roll, pitch and yaw angles, respectively. Both types of data, kinematic and non-kinematic, are strongly correlated with the aspect

angle ϕ. In fact, as one can see from (1), (2), target velocity and range profile are coupled by the aspect angle ϕ and the angular velocity q. Effective estimation of ϕ is the key to successful target identification.

The acceleration $f(t)$ and the angular velocity of the target $q(t)$ are only partially observable and their dynamics is difficult to predict (at least in the case of maneuvering non-cooperative target). To take the above uncertainty into account, we model $\phi(t)$, $f(t)$ as a pair of stochastic processes. In particular, for $f(t)$ we assume the white noise acceleration model [2]. The latter means that $f(t_k)$, $k = 1, 2, \ldots$ is a sequence of independent Gaussian random variables with zero mean and unknown intensity σ^2.

The dynamics of the aspect angle $\phi(t)$ is modeled by an interactive Markov diffusion process. Specifically for a target from class H_j its aspect angle $\phi_j(t) = (\phi_{j1}(t), \phi_{j2}(t), \phi_{j3}(t))$ is given by the stochastic differential equation

$$\begin{aligned} \dot{\phi}_j(t) &= F_j(\theta_j(t), \phi_j(t), t) \\ &+ G_j(\theta_j(t), \phi_j(t), t)\dot{\beta}_j(t), \end{aligned} \tag{3}$$

where $\dot{\beta}_j(t)$ is a white noise and $\theta_j(t)$ is a switching parameter. This parameter models switches between target modes (maneuvers). We will assume that $\theta_j(t)$ is a Markov type jump process. Its intensity matrix Λ_j characterizes *a priori* probabilities of transitions between maneuvers. The functions F_j, G_j and Λ_j bear *a priori* knowledge about target kinematics. Of course these functions are class specific. We will omit the index j related to the target class in $\phi_j(t)$, $\theta_j(t)$, etc. when it does not lead to ambiguity.

The above model of the aspect angle is a refinement and generalization of the standard interactive multiple model (IMM) approach to modeling dynamics of non-cooperative maneuvering targets.

To complete the description of our model, it remains to consider the structure of measurements in more detail. It is natural to assume that the variance of the acceleration $f(t_k)$ dominates the variance of errors in velocity measurements. So we can assume safely that the velocity measurements are noise free. For the sake of simplicity we will ignore the range measurements. In this case the kinematic measurements are $z_k = v(t_k)$. Write $w_k = f(t_k)(t_{k+1} - t_k)$. Explicit discretization of (2) gives

$$z_{k+1} = C_k z_k + w_k, \tag{4}$$

where $C_k = I + (t_{k+1} - t_k)Q_k$ and

$$Q_k = \begin{bmatrix} 0 & -q_{3,k} & q_{2,k} \\ q_{3,k} & 0 & -q_{1,k} \\ -q_{2,k} & q_{1,k} & 0 \end{bmatrix},$$

$q_{i,k} \approx q_i(t_k)$ with $\dot{\phi}_i(t_k)$ approximated by

$$\dot{\phi}_i(t_k) \approx \frac{\phi_i(t_{k+1}) - \phi_i(t_k)}{t_{k+1} - t_k}, \qquad i = 1, 2, 3.$$

The random variables w_k, $k = 1, 2, \ldots$ are supposed to be independent Gaussian vectors with zero mean and covariance matrix $(t_{k+1} - t_k)^2\sigma^2 I$, where I is the identity matrix. The intensity parameter σ^2 is unknown but can be estimated easily assuming that the overlook time $t_{k+1} - t_k$ is sufficiently small.

It must be noted that the range profile signal $S_j(\phi)$ depends also on a number of (in general) unknown parameters including amplitude, group time lag, and number of range resolution elements, etc. There is also a number of sensor dependent error sources that contribute to the noise distribution. Here we assume that the noise N is Gaussian with zero mean and covariance Σ. However, more realistic models of noise can also be incorporated in our model without much difficulty. Unknown parameters in (1) can be estimated via combination of signature simulation and signature collection.

In conclusion of this section, we remark that the mathematical model of target dynamics and observation (1)-(4) belongs to the general type of hidden Markov models. According to terminology of HMM approach the process $X(t) = (\phi(t), \theta(t))$ is the state process and $y_k = (r_k, z_k)$ is the observation process.

Note also that in our model both the state process and the observation are nonlinear. This rules out application of standard methods of estimation based on Kalman filter. In the next section we describe novel nonlinear filtering techniques based on spectral separating scheme that allows us to compute the joint posterior distribution $P(H_j, \phi(t_k) \in A | Y_k)$ in an efficient manner.

3 Data Fusion Based on Nonlinear Filtering

Our approach to fusion of kinematic and range profiling data is based on the Bayesian approach. We start with M hypotheses $\{H_1, H_2, \ldots, H_M\}$ regarding the type of the target. We will compute sequentially posterior distributions of the hypothesis H_j, $P(H_j | Y_k)$, given the measurements

$\boldsymbol{Y}_k$, $k = 1, 2, \ldots$ and identify the target by using modern sequential multiple hypothesis testing techniques [3, 16, 17, 18].

The crucial part of our fusion and identification algorithm is the fully coupled optimal nonlinear filter for the aspect angle. This filter must compute recursively the joint posterior distributions $P_j^k(A_n) = P(H_j, \phi(t_k) \in A_n | \boldsymbol{Y}_k)$ where A_n is the nth bin of aspect angle (it is assumed here that the viewing sphere is partitioned into N angular bins). In what follows we write $\phi_k = \phi(t_k)$ and $\theta_k = \theta(t_k)$ for brevity.

Given a new set of measurements $\boldsymbol{y}_{k+1}$ at time t_{k+1}, the filtering distribution $P_j^k(A_n)$ is updated according to the Bayes rule

$$P_j^{k+1}(A_n) = \tag{5}$$

$$\frac{P(\boldsymbol{y}_{k+1} | H_j, \phi_{k+1} \in A_n, \boldsymbol{Y}_k) P_j^k(A_n)}{\sum_{j=1}^M \sum_{n=1}^N P_j(\boldsymbol{y}_{k+1} | H_j, \phi_{k+1} \in A_n, \boldsymbol{Y}_k) P_j^k(A_n)}.$$

By integrating out A_n's (respectively H_j's) one can obtain from (5) the posterior distributions $P(H_j | \boldsymbol{Y}_{k+1})$, $j = 1, \ldots, M$ (respectively $P(\phi_{k+1} \in A_n | \boldsymbol{Y}_{k+1})$, $n = 1, \ldots, N$).

Fusion of kinematic and non-kinematic measurements is facilitated by the *correction term* $P(\boldsymbol{y}_{k+1} | H_j, \phi_{k+1} \in A_n, \boldsymbol{Y}_k)$.

Formula (5) provides a general form of the nonlinear filter. In this form it cannot be implemented efficiently since we have yet no means to compute the correction term. However, filter (5) can be refined by using two important properties of the HMM (1)-(4):

(i) The kinematic measurements $\boldsymbol{z}_k$ and range-profile data $\boldsymbol{r}_k$ are *conditionally* independent given ϕ_k, ϕ_{k-1} and $\boldsymbol{z}_{k-1}$. (Note that without the conditioning $\boldsymbol{z}_k$ and $\boldsymbol{r}_k$ are strongly correlated.)

(ii) $\mathbf{X}_k = (\phi_k, \theta_k)$, $k = 0, 1, 2, \ldots$, is a homogeneous Markov chain.

Write

$$P_j^{k+1,i}(A_n) = P(H_j, \phi_k \in A_n, \theta_k = i | \boldsymbol{Y}_k).$$

Obviously, $P_j^{k+1}(A_n) = \sum_i P_j^{k+1,i}(A_n)$.

Using (i) and (ii) we obtain

$$P_j^{k+1,l}(A_n) = \tag{6}$$

$$\frac{\rho_{j,k+1}^n \sum_{m,i} p_j(m,i,n,l) \zeta_{j,k+1}^{n,m} P_j^{k,i}(A_m)}{\sum_{j=1}^M \sum_{n,l} \rho_{j,k+1}^n \sum_{m,i} p_j(m,i,n,l) \zeta_{j,k+1}^{n,m} P_j^{k,i}(A_m)}$$

where

$$p_j(m,i,n,l) =$$

$$P(\phi_{k+1} \in A_n, \theta_{k+1} = l | H_j, \phi_k \in A_m, \theta_k = i),$$

$$\rho_{j,k+1}^n = P(\boldsymbol{r}_{k+1} | H_j, \phi_{k+1} \in A_n),$$

$$\zeta_{j,k+1}^{n,m} = P(\boldsymbol{z}_{k+1} | H_j, \phi_{k+1} \in A_n, \phi_k \in A_m, \boldsymbol{z}_k).$$

Formula (6) demonstrates an important fact: in a setting with fully coupled kinematic and non-kinematic measurements conditioning on ϕ_{k+1}, ϕ_k and $\boldsymbol{z}_k$ decouples the correction term into the product of the kinematic and the non-kinematic conditional correction terms $\zeta_{j,k+1}^{n,m}$ and $\rho_{j,k+1}^n$ respectively.

In contrast to (5), the filter given by (6) is practically implementable. Indeed, with the use of the models (1) and (4) it is readily checked that

$$\rho_{j,k+1}^n \approx \frac{1}{(2\pi)^{r/2} |\boldsymbol{\Sigma}|} \times$$

$$\exp\left\{ -\frac{1}{2} (\boldsymbol{r}_{k+1} - \boldsymbol{S}_j(a_n))^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{r}_{k+1} - \boldsymbol{S}_j(a_n)) \right\},$$

and

$$\zeta_{j,k+1}^{n,m} \approx \frac{1}{|(2\pi)^{1/2} \Delta_{k+1} \sigma|^3} \times$$

$$\exp\left\{ -\frac{(\boldsymbol{z}_{k+1} - \boldsymbol{C}_{j,k} \boldsymbol{z}_k)^T (\boldsymbol{z}_{k+1} - \boldsymbol{C}_{j,k} \boldsymbol{z}_k)}{2\sigma^2 \Delta_{k+1}^2} \right\},$$

where $\Delta_{k+1} = t_{k+1} - t_k$, a_n is the center of mass of the bin A_n and $\boldsymbol{C}_{j,k} = \boldsymbol{C}_k(\phi_k = a_n, \phi_{k+1} = a_m)$ ($\boldsymbol{C}_{j,k}$ depends on the number of class j, since ϕ_k is class specific).

Optimal nonlinear filter (6) can be greatly simplified by switching to unnormalized filtering distributions [4]. Specifically, one can show that

$$P_j^{k+1,i}(A_n) = \widetilde{P}_j^{k+1,i}(A_n) / \sum_{n=1}^N \widetilde{P}_j^{k+1,i}(A_n),$$

where the unnormalized filtering distribution $\widetilde{P}_j^{k+1,i}(A_n)$ is given by

$$\widetilde{P}_j^{k+1,l}(A_n) = \tag{7}$$

$$\widetilde{\rho}_{j,k+1}^n \sum_{m,i} p_j(m,i,n,l) \widetilde{\zeta}_{j,k+1}^{n,m} \widetilde{P}_j^{k,i}(A_m)$$

with

$$\widetilde{\rho}_{j,k+1}^n = \tag{8}$$

$$\exp\left\{ \boldsymbol{S}_j(a_n)^T \boldsymbol{\Sigma}^{-1} \boldsymbol{r}_{k+1} - \frac{1}{2} \boldsymbol{S}_j(a_n)^T \boldsymbol{\Sigma}^{-1} \boldsymbol{S}_j(a_n) \right\},$$

$$\widetilde{\zeta}_{j,k+1}^{n,m} = \tag{9}$$

$$\exp\left\{\frac{1}{\sigma^2\Delta_{k+1}^2}(z_{k+1}^T C_{j,k} z_k - \frac{1}{2}|C_{j,k}z_k|^2)\right\}.$$

Note that since the unnormalized distribution $\widetilde{P}_j^{k+1}(A_n) = \sum_l \widetilde{P}_j^{k+1,l}(A_n)$ achieves maximum at the same point as the normalized one, $P_j^{k+1}(A_n)$, the former one is usually sufficient for the purpose of target identification. In addition, (7) is much simpler than (6). The main advantage of the unnormalized filter (often referred to as Zakai filter) is its linearity. This property allows us to implement powerful numerical schemes for data fusion and target identification.

Computational complexity is the most serious roadblock on the way of practical implementation of the optimal nonlinear filters (6) and (7). The most computationally expensive part of algorithms (6) and (7) is evaluation of the sum $u_j^l(n) = \sum_{m,i} p_j(m,i,n,l) f_j(m,i,n,l)$, where $f_j(m,i,n,l)$ is equal to $\widetilde{\zeta}_{j,k+1}^{n,m} \widetilde{P}_j^{k,i}(A_m)$ in the case of (7) and $\zeta_{j,k+1}^{n,m} P_j^{k,i}(A_m)$ in the case of (6). Standard numerical algorithms for solving this problem have computational complexity $O(N(\ln N)^2)$, where N, the number of aspect angle bins, can be of the order of $10^4 - 10^5$. If the algorithm is applied straightforwardly, this translates into $10^8 - 10^{10}$ operations per step. Thus a real time implementation of the above algorithm is not obvious. Several novel numerical techniques which address this problem were introduced recently. In particular, the Spectral Separating Scheme (S^3) [11, 12, 13] and the Stochastic Domain Pursuit (SDP) method [14, 15] appears the most promising. Both algorithms reduce the online computational complexity to the level $O(N)$. Due to the lack of space, we do not discuss the details of adaptation of S^3 and SDP methods to this particular setting and leave it to the interested reader.

4 Identification Algorithms

We develop two types of sequential identification algorithms – completely Bayesian and non-Bayesian.

4.1 Bayesian Algorithm

The first identification algorithm is based on comparison of *a posteriori* probabilities with each other and with a threshold level that is defined based on a given misclassification rate. Note that the decision statistics (posterior probabilities) exploit both kinematic and HRRR-profile data. In other words, it is completely coupled algorithm.

Let H_j, $j = 1, \ldots, M$ be the set of M hypotheses regarding the type of the target and $\mathbf{\Pi}(0) = (\pi_1(0), \ldots, \pi_M(0))$ be the vector of prior probabilities assigned to these hypotheses. This distribution may represent human factors (e.g. the operator's judgment expressed in the form of subjective probabilities), or the statistical estimates for the particular tactical situation, or a combination of both. Strictly speaking, in order to obtain an *a posteriori* distribution of hypotheses, $\mathbf{\Pi}(t_k) = (\pi_1(t_k), \ldots, \pi_M(t_k))$, $\pi_j(t_k) = P(H_j|\boldsymbol{Y}_k)$, we should average the joint distribution $P_j^k(A_m) = P(H_j, \boldsymbol{\phi}(t_k) \in A_m|\boldsymbol{Y}_k)$ over m:

$$\pi_j(t_k) = \sum_{m=1}^{N} P_j^k(A_m), \tag{10}$$

where N is the number of aspect angle bins.

The recognition (identification) algorithm at the kth step is as follows:

- if $\max_j \pi_j(t_k) < C_{\alpha,M}$, go to the step $k+1$, where $C_{\alpha,M}$ is a threshold level which depends on the predefined probability of misclassification α (typically chosen between 0.01 and 0.1) and the number of target classes M;
- if $\max_j \pi_j(t_k) \geq C_{\alpha,M}$ and $\pi_\kappa(t_k) = \max_j \pi_j(t_k)$, the observation process is stopped and the target is identified as belonging to the class H_κ.

This sequential classification algorithm has the following important properties [3, 16, 17, 18]. If the threshold is chosen as

$$C_{\alpha,M} = 1 - \alpha/M, \tag{11}$$

then the algorithm belongs to the class of identification procedures for which $\alpha_j = \Pr(\text{accepting } H_j|H_j \text{ is wrong}) \leq \alpha$ (i.e. the probability of misidentification α_j does not exceed the given level $\alpha \in (0,1)$). Moreover, in this case the algorithm minimizes asymptotically (when α is small enough) the expected sample size for all hypothesis.

More specifically, let

$$L_{ji}(t_k) = \ln\left[\frac{P(\boldsymbol{Y}_k|H_j)}{P(\boldsymbol{Y}_k|H_i}\right]$$

be the log-likelihood ratio of target classes H_j and H_i, where $P(\boldsymbol{Y}_k|H_l) = \sum_{n_1,\ldots,n_k} P(\boldsymbol{Y}_k, \boldsymbol{\phi}_1 \in A_{n_1}, \ldots, \boldsymbol{\phi}_k \in A_{n_k}|H_l)$. Next, let $\boldsymbol{E}_l$ denote the expectation when the observations correspond to the class H_l (under distribution $P(\boldsymbol{Y}_k|H_l)$) and let

$$Q(j,i) = \lim_{k\to\infty} \frac{1}{k} \boldsymbol{E}_j L_{ji}(t_k) \tag{12}$$

be the parameter which characterizes the distance between jth and ith classes. Next, let τ_B denote the sample size (stopping time) of the Bayes identification algorithm. Obviously, τ_B is the first time t_k such that the statistic $\max_j \pi_j(t_k)$ exceeds the threshold $C_{\alpha,M} = 1 - \alpha/M$.

By analogy with [18] it can be shown that if $k^{-1}L_{ji}(t_k)$ converges strongly completely to $Q(j,i)$ as $k \to \infty$ (the definition of strong complete convergence see in [18]), then the proposed sequential identification procedure minimizes any positive moment of the observation time for small α and

$$E_j \tau_B^n \sim \left| \frac{\ln \alpha}{\min_{i \neq j} Q(j,i)} \right|^n \quad \text{as } \alpha \to 0 \qquad (13)$$

for any $n > 0$ (for $n = 1$ this is the expected sample size). The right hand side of (13) may serve as a reasonable approximation to the nth moment of the observation time for small probability of misclassification. (See [3] and Section 5 for the results of simulation).

4.2 Non-Bayesian Algorithm

Another identification algorithm is based on the conditionally optimal estimators of the aspect angle, $\widehat{\phi}_j(t_k) = E[\phi(t_k)|H_j, \boldsymbol{Y}_k]$, for each hypothesis H_j. These estimates represent the output of nonlinear filters for the aspect angle. The algorithm does not require any knowledge of an *a priori* distribution of hypotheses and at the same time has the same asymptotic performance as the previous method [16, 17, 18].

Let

$$L_{ji}(t_k, \Phi_k^j, \Phi_k^i) = \ln[P(\boldsymbol{Y}_k|H_j, \Phi_k^j)/P(\boldsymbol{Y}_k|H_i, \Phi_k^i)]$$

be the "conditional" log-likelihood ratio of the hypotheses H_j and H_i (we stress the difference as compared to the statistic $L_{ji}(t_k)$ defined in Section 4.1) and define the adaptive version of the log-likelihood ratio by the recursion

$$\widehat{L}_{ji}(k+1) = \widehat{L}_{ji}(k) + \ln \left[\frac{P\{y_{k+1}|H_j, \widehat{\phi}_j(k), \boldsymbol{Y}_k\}}{P\{y_{k+1}|H_i, \widehat{\phi}_i(k), \boldsymbol{Y}_k\}} \right]$$

where $\widehat{\phi}_l(k) = \widehat{\phi}_l(t_k)$. The identification algorithm at the kth step is as follows:

- if $\max_j \min_{i \neq j} \widehat{L}_{ji}(k) < B_{\alpha,M}$, go to the step $k+1$, where $B_{\alpha,M}$ is a threshold which depends on the given probability of misclassification α and the number of target classes M;
- if $\max_j \min_{i \neq j} \widehat{L}_{ji}(k) \geq B_{\alpha,M}$ and $\min_{i \neq \kappa} \widehat{L}_{\kappa i}(k) = \max_j \min_{i \neq j} \widehat{L}_{ji}(k)$, the observation process is stopped and a target is identified as belonging to the class H_κ.

In other words, the classifier is based on simultaneous application of a number of one-sided sequential probability ratio tests, acting in parallel, each of which intends to test the hypotheses H_j against all other alternatives. The algorithm stops observation at time

$$\tau_{\text{NB}} = \min\left\{k : \max_j \min_{i \neq j} L_{ji}(k) \geq B_{\alpha,M}\right\} \qquad (14)$$

and decides in favor of the class H_κ if

$$\min_{i \neq \kappa} L_{\kappa i}(\tau_{\text{NB}}) = \max_j \min_{i \neq j} L_{ji}(\tau_{\text{NB}}).$$

It can be shown that the probability of misidentification, $\alpha_j = \Pr(\text{accept } H_j | H_j \text{ is wrong})$, does not exceed the predefined level α when

$$B_{\alpha,M} = \ln[(M-1)/\alpha]. \qquad (15)$$

Also, the asymptotic formula (13) is valid for τ_{NB} whenever

$$\lim_{k \to \infty} \frac{1}{k} E_j \widehat{L}_{ji}(k) = Q(j,i) \quad \text{for all } i, j,\ i \neq j,$$

where $Q(j,i)$ is defined in (12). Thus, both proposed identification algorithms are optimal for small α.

The block diagram of the algorithm is shown in Figure 1. At the kth step the algorithm performs three tasks:

(i) computing optimal (nonlinear) filtering estimates $\widehat{\phi}_l(t_k) = E_l[\phi(t_k)|H_l, \boldsymbol{Y}_k]$ $(l = 1, \ldots, M)$ using nonlinear filtering algorithm described in Section 3;

(ii) computing the matrix of adaptive log-likelihood ratios $\|\widehat{L}_{ji}(t_k)\|$ $(i, j = 1, \ldots, M,\ i \neq j)$ which exploit the estimates $\widehat{\phi}_l$ up to time t_{k-1};

(iii) thresholding.

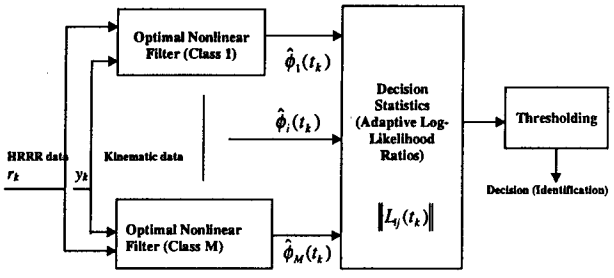


Figure 1: **Block-diagram of the data fusion and identification algorithm**

5 Performance and Conclusions

The target recognition performance of the proposed sequential identification algorithms is shown in Table 1, where we illustrate the system performance in the case where the log-likelihood ratios $\widehat{L}_{ji}(t_k)$ can be well approximated by the Gaussian processes with independent increments with means $\boldsymbol{E}_j\widehat{L}_{ji}(t_k) = Q(j,i) > 0$. The values of $Q(j,i)$ define the distances between classes H_j and H_i (see (12)). The performance of algorithms is evaluated in terms of the expected sample size required for the identification when the probabilities of misidentification $\alpha_j = \Pr(\text{accept } H_j | H_j \text{ is wrong})$ are fixed at the level $\alpha = 0.01$. In simulations the prior distribution of classes was assumed to be uniform, $\pi_j(0) = 1/M$, $j = 1,\dots,M$, and the number of classes $M = 3$. In the table, $\widehat{\alpha}_j$ and $\widehat{\boldsymbol{E}}_j\tau$ are the estimates of the error probabilities and expected sample sizes of tests obtained by the Monte Carlo technique, α_j is the given constraints, and $\boldsymbol{E}_j\tau$ is the expected sample size computed by the asymptotic formula (13).

It turns out that the thresholds (11) and (15) guarantee only the inequalities $\alpha_j \le \alpha$ for all $j = 1,\dots,M$. In general this choice does not guarantee the equalities $\alpha_j = \alpha$, which should be satisfied at least approximately to compare different algorithms correctly. To obtain accurate approximations for error probabilities we evaluated average overshoots of log-likelihood ratios over the boundaries and applied the nonlinear renewal theory techniques [3]. As a result, to guarantee the equalities $\alpha_j = \alpha$ for all j the thresholds can be different for different hypotheses (due to different overshoots). Particularly, it is seen from Table 1 that the thresholds C_3 and B_3 for H_3 differ from the thresholds $C_1 = C_2$ and $B_1 = B_2$. In other words we applied a slightly more general sequential algorithms compared to algorithms described in Section 4.1 and Section 4.2. For instance, the stopping time of the non-Bayes algorithm is

$$\tau_{\text{NB}} = \min(\tau_1, \tau_2, \dots, \tau_M),$$

$$\tau_i = \min\{k : \min_{n \ne i} \widehat{L}_{in}(t_k) \ge B_i\}, \quad i = 1,\dots,M$$

(compare with (14)). The decision is made in favor of the class H_κ if $\tau_{\text{NB}} = \tau_\kappa$.

The results presented in the table allow us to make the following conclusions.

1. The theoretical (asymptotic) estimates (13) give a reasonable approximation to the expected sample size even for moderate probabilities of errors.

2. Proposed sequential identification algorithms have almost the same performance – the difference between expected number of observations required to achieve the probability of misidentification $\alpha = 0.01$ is negligible.

3. Since the best fixed sample size identification algorithm takes 34 observation, the sequential algorithms are in average two to four times faster. Thus potentially the proposed sequential algorithms are better as compared to the non-sequential dynamic programming approach developed in [10].

The asymptotic formula (13) suggests a way of comparison of different data fusion methods in terms of highest recognition performance: the greater distances $Q(j,i)$ between classes, the better the data fusion algorithm is. The distances $Q(j,i)$ have a simple information-theoretic interpretation. Indeed, the value of $\boldsymbol{E}_j L_{ji}(t_k)$ is nothing but the Kullback-Leibler information distance between probability distributions $P(\boldsymbol{Y}_k|H_j)$ and $P(\boldsymbol{Y}_k|H_i)$. Hence $Q(j,i)$ is the effective (average) information distance between classes H_j and H_i per one observation. Fusion of data allows us to increase the effective distance between classes. The potential increase of $Q(j,i)$ defines the efficiency of the data fusion algorithm. This important issue will be considered elsewhere.

Table 1: **Performance of Sequential Identification Algorithms for Three Classes.** The number of trials used in the simulations is 10^5. The distances between classes are $Q(2,1) = Q(1,2) = 0.18$, $Q(3,2) = Q(2,3) = 0.5$; $Q(3,1) = Q(1,3) = 1.28$. The best fixed sample size test that meets the constraint on the probability of misidentification $\alpha = 0.01$ takes 34 observations.

Results for the Bayesian Algorithm					
	Error Prob. & Thres.			Exp. Sample Size	
	α_j	$\ln C_j$	$\widehat{\alpha}_j$	$\widehat{\boldsymbol{E}}_j\tau_{\text{B}}$	$\boldsymbol{E}_j\tau_{\text{B}}$
H_1	0.01	3.16	0.0097	18.83	17.54
H_2	0.01	3.16	0.0091	21.46	17.54
H_3	0.01	2.93	0.0098	7.24	5.85
Results for the non-Bayesian Algorithm					
	α_j	B_j	$\widehat{\alpha}_j$	$\widehat{\boldsymbol{E}}_j\tau_{\text{NB}}$	$\boldsymbol{E}_j\tau_{\text{NB}}$
H_1	0.01	3.16	0.0097	18.75	17.54
H_2	0.01	3.16	0.0106	20.85	17.54
H_3	0.01	2.93	0.0100	7.17	5.85

References

[1] Y. Bar-Shalom and T.E. Fortmann. *Tracking and Data Association.* Academic Press Inc., New York, 1988.

[2] Y. Bar-Shalom and Xiao-Rong Li. *Estimation and Tracking: Principles, Techniques, and Software.* Artech House, Inc., Boston, 1993.

[3] V.P. Dragalin, A.G. Tartakovsky, and V.V. Veeravalli. Multihypothesis sequential probability ratio tests, part II: accurate asymptotic expansions for the expected sample size. *IEEE Transactions on Information Theory* (submitted September 1998, expected publication 2000).

[4] R.J. Elliott, L. Aggoun, and J.B. Moore. *Hidden Markov Models: Estimation and Control.* Springer-Verlag, New York, 1995.

[5] B. Etkin. *Dynamics of Atmospheric Flight.* John Wiley & Sons, New York, 1972.

[6] H. Goldstein. *Classical Mechanics* (2nd Ed). Addison-Wesley Publ. Co., Reading, MA, 1980.

[7] S.P. Jacobs and J.A. O'Sullivan. High resolution radar models for joint tracking and recognition. *1997 IEEE National Radar Conference.* 99-104, 1997.

[8] A.H. Jazwinski. *Stochastic Processes and Filtering Theory.* Academic Press, New York, 1970.

[9] R.E. Larson and J. Peschon. A dynamic programming approach to trajectory estimation. *IEEE Transactions on Autom. Control,* 11(3):537-540, July 1966.

[10] E.W. Libby and P.S. Maybeck. Sequence comparison techniques for multisensor data fusion and target recognition. *IEEE Transactions on Aerospace and Electronic Systems,* 32(1):52-65, January 1996.

[11] S. Lototsky, R. Mikulevicius, and B.L. Rozovskii. Nonlinear filtering revisited: a spectral approach. *SIAM J. Control and Optimization,* 35(2): 435-461, March 1997.

[12] S. Lototsky and B.L. Rozovskii. Recursive nonlinear filter for a continuous-discrete time model: separation of parameters and observations. *IEEE Transactions on Autom. Control,* 43(8): 1154-1158, 1998.

[13] S. Lototsky, C. Rao, and B.L. Rozovskii. Fast nonlinear filter for continuous-discrete time multiple models. *Proceeding of the 35th IEEE Conference on Decision and Control,* 4:4060-4064, Kobe, Japan, Omnipress, Madison, 1997.

[14] C. Rao. *Nonlinear Filtering and Evolution Equations: Fast Algorithms with Applications to Target Tracking.* PhD thesis, Department of Mathematics, University of Southern California, Los Angeles, August 1998.

[15] C. Rao, B. Rozovskii, and A. Tartakovsky. Domain pursuit method for tracking ballistic targets. Technical Report # CAMS-98.9.2, Center for Applied Mathematical Sciences, University of Southern California, September 1998. (Available at `http://www.usc.edu/dept/LAS/CAMS/`).

[16] A.G. Tartakovsky. *Sequential Methods in the Theory of Information Systems.* Radio & Communications, Moscow, 1991.

[17] A.G. Tartakovsky. Multihypothesis invariant sequential tests with applications to the quickest detection of signals in multiple-resolution element systems. *Proceedings of the 1998 Conference on Information Sciences and Systems,* Vol. 2, 906-911, March 1998, Princeton University, Princeton, NJ.

[18] A.G. Tartakovsky. Asymptotic optimality of certain multialternative sequential tests: non-i.i.d. case. *Statistical Inference for Stochastic Processes,* 1999.

[19] A.J. Viterbi. Error bounds for convolution codes and an asymptotically optimal decoding algorithm. *IEEE Transactions on Information Theory,* 13: 260-269, April 1967.

Partial Branch and Bound Algorithm for Improved Data Association in Multi-frame Processing

Aubrey B. Poore
Department of Mathematics
Colorado State University
Fort Collins, CO, U.S.A.

Xin Yan
Department of Mathematics
Colorado State University
Fort Collins, CO, U.S.A.

Abstract *The problem of data association remains central in multitarget, multisensor, and multiplatform tracking. Lagrangian relaxation methods have been shown to yield near optimal answers in real-time. The necessarity of improvement in the quality of these solutions warrants a continuing interest in these methods. A partial branch-and-bound technique along with adequate branching and ordering rules are developed. Lagrangian relaxation is used as a branching method and as a method to calculate the lower bound for subproblems. The result shows that the branch-and-bound framework greatly improves the solution quality of the Lagrangian relaxation algorithm and yields better multiple solutions in less time than relaxation alone.*

Keywords: Lagrangian Relaxation Algorithm, Branch-and-Bound, Multidimensional Assignment Problem, Multitarget Tracking

1 Introduction

The multiframe data association problem for multitarget and multisensor tracking is formulated as a multidimensional assignment problem. This formulation is a superset of almost all MHT approaches to multiframe processing. The construction of real-time solutions to this fundamental problem has been achieved by the use of Lagrangian relaxation techniques, but the quest for improvements approaching optimality without going to full branch and bound techniques will remain a fundamental problem for some time. In this work, we examine a couple of techniques for improving the solution quality and demonstrate their effectiveness on some difficult tracking problems.

The multiframe data association problem is formulated as a multidimensional assignment problem[1, 2] as

$$
\begin{aligned}
\text{Min} \quad & \sum_{i_1=0}^{M_1} \cdots \sum_{i_N=0}^{M_N} c_{i_1 \cdots i_N} z_{i_1 \cdots i_N} \\
\text{St:} \quad & \sum_{i_2=0}^{M_2} \cdots \sum_{i_N=0}^{M_N} z_{i_1 \cdots i_N} = 1, \\
& \quad \text{for } i_1 = 1, \ldots, M_1, \\
& \sum_{i_1=0}^{M_1} \cdots \sum_{i_{k-1}=0}^{M_{k-1}} \sum_{i_{k+1}=0}^{M_{k+1}} \cdots \sum_{i_N=0}^{M_N} z_{i_1 \cdots i_N} \\
& \quad = 1, \qquad (1) \\
& \quad \text{for } i_k = 1, \ldots, M_k \\
& \quad \text{and } k = 2, \ldots, N-1, \\
& \sum_{i_1=0}^{M_1} \cdots \sum_{i_{N-1}=0}^{M_{N-1}} z_{i_1 \cdots i_N} = 1, \\
& \quad \text{for } i_N = 1, \ldots, M_N, \\
& z_{i_1 \cdots i_N} \in \{0, 1\} \quad \text{for all } i_1, \ldots, i_N.
\end{aligned}
$$

Here, $c_{0 \cdots 0}$ is arbitrarily defined to be zero and is included for notational convenience. The zero index is used to representing missing data, false alarms, initiating tracks and terminating tracks. We assume that the binary variables $z_{i_1 \cdots i_N}$ with precisely one nonzero index are free to be assigned and that the corresponding cost coefficients are well-defined. Actually these cost coefficients with exactly one nonzero index can be translated to zero by cost shifting

without changing the optimal assignment.

The only known methods for solving this problem optimally are enumerative in nature, with branch-and-bound being the most efficient. However, such algorithms are too slow for real-time applications. Because of the noise included in the cost coefficients, it is sufficient to find suboptimal solutions that are within the noise level of the true solution.

There are a variety of Lagrangian relaxation based methods, and large classes of these are explained in publications [3, 4, 5]. Because a moving window is usually used for track maintenance and track initiation, we prefer an algorithm that first relaxes an N-dimensional assignment problem down to a 2-dimensional assignment problem, optimizes over the multipliers, and then restores feasibility using an (N-1) dimensional assignment problem. This procedure is repeated on successive recovery problems until a 2-dimensional assignment problem is reached, which can be solved optimally in polynomial time. This algorithm is summarized in [1].

One of the key user inputs in this particular Lagrangian relaxation algorithm is the number of nonsmooth optimization steps that are taken. Usually the quality of the solution improves with the number of iterations. This improvement is not monotone, but has slight variations up and down, as the number of iterations increases. With 20 or so iterations, the solution quality is generally well within the noise level of the underlying problem; however, there is an ever increasing demand for better solution quality or examination of the relation between a "good" solution and one that is "better".

There are several approaches to improving solution quality. One approach is to increase the number of iterations in the multiplier adjustments (nonsmooth optimization steps). Another is to first generate a good solution and then use a local search technique to examine the solutions in a neighborhood of the existing one. Although we have had only marginal improvements with this approach [6], we believe that this avenue still needs to be explored. Another currently popular approach is to use the K-best solutions of the two dimensional assignment [7] problem to examine different potential solutions. Disappointingly, our testing shows that this approach has also failed to lead to any appreciable improvements to the relaxation solutions.

In this work, we develop two algorithms for improving the solution quality. The first is a heuristic to decide which solutions of the two dimensional assignment problem (that arises in the nonsmooth optimization iterations) lead to improved solutions of the data association problem. The second algorithm is an enumerative technique framed in the partial branch and bound paradigm and that generally yields uniformly improved multiple solutions.

An overview of the Lagrangian relaxation algorithm is in Chapter 2. In Chapter 3 the index alignment selection heuristic and the partial branch-and-bound algorithm are presented. Numerical results are presented in Chapter 3.

2 Lagrangian Relaxation Algorithm and Inner Problem

2.1 Overview of Lagrangian Relaxation Algorithm

For an N-dimensional assignment problem the Lagrangian relaxation algorithm consists of N-1 stages. In the first stage, the last (N-2) constraint sets are relaxed by applying Lagrangian multipliers $(u^3, \ldots, u^n)$. The dual problem is:

$$\Phi_n(u^3, \ldots, u^n) \equiv$$

$$\text{Min} \quad \text{Ł}_n(z^n; u^3, \ldots, u^n) \equiv$$

$$\text{Min} \quad \sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} \cdots \sum_{i_n=0}^{M_n} c^n_{i_1 i_2 \cdots i_n} z^n_{i_1 i_2 \cdots i_n}$$

$$+ \sum_{k=3}^{n} \sum_{i_k=0}^{M_k} u^k_{i_k} [\sum_{i_1=0}^{M_1} \cdots \sum_{i_{k-1}=0}^{M_{k-1}}$$

$$\sum_{i_{k+1}=0}^{M_{k+1}} \cdots \sum_{i_n=0}^{M_n} z^n_{i_1 \cdots i_n} - 1] \equiv$$

$$\text{Min} \quad \sum_{i_1=0}^{M_1} \cdots \sum_{i_n}^{M_n} \left[c^n_{i_1 \cdots i_n} + \sum_{k=3}^{n} u^k_{i_k} \right] z^n_{i_1 \cdots i_n} \quad (2)$$
$$- \sum_{k=3}^{n} \sum_{i_k=0}^{M_k} u^k_{i_k}$$

$$\text{St:} \quad \sum_{i_2=0}^{M_2} \cdots \sum_{i_n=0}^{M_n} z^n_{i_1 \cdots i_n} = 1,$$
$$\text{for } i_1 = 1, \ldots, M_1, \quad (3)$$
$$\sum_{i_1=0}^{M_1} \cdots \sum_{i_n=0}^{M_n} z^n_{i_1 \cdots i_n} = 1,$$
$$\text{for } i_2 = 1, \ldots, M_2, \quad (4)$$
$$z^n_{i_1 \cdots i_n} \in \{0, 1\}$$
$$\text{for all } i_1, \ldots, i_n.$$

The next task is to maximize the dual problem with respect to the multiplier vector $(u^3, \ldots, u^n)$, which is accomplished by using techniques of non-smooth optimization. In each step, a new multiplier vector is generated and the corresponding solution is computed which determines an index alignment for the first 2 index sets, say $\alpha = \{(i_1(j), i_2(j)) | j = 0 \cdots M_0\}$ with $(i_1(0), i_2(0)) = (0, 0)$. Here, an index alignment α is a set consisting of ordered sets with the same size, i.e., $\alpha = \{\{i^1_1 \cdots i^1_k\}, \cdots, \{i^m_1, \cdots i^m_k\}\}$. A solution Y conforms to an index alignment α if

$$y_{i_1, \cdots i_n} = \left\{ \begin{array}{l} 1, \text{ if } \{i_1, \cdots i_k\} \in \alpha \\ 0, \text{ if } \{i_1, \cdots i_k\} \notin \alpha \end{array} \right\}$$
$$\text{for all } i_1, \cdots, i_n \quad (5)$$

Based on this index alignment an $(N-1)$-dimensional assignment problem is constructed by setting $c^{n-1}_{ji_3 \cdots i_n} \equiv c^n_{i_1(j) i_2(j) i_3 \cdots i_n}$ and keeping only those arcs with the first two indices appearing in α. Its dual problem is formed in the same way we do with the original N-dimensional assignment problem by relaxing all but the last two constraint sets from. The process continues and $N-2, N-3 \cdots 2$ dimensional assignment problems are formed and alignments made. At the last stage a 2-dimensional assignment problem is formed and solved. Thus a near optimal feasible solution for the original problem is found.

This algorithm is not intended to find an optimal solution. The solution it finds is sub-optimal with the optimal value of the dual function $\Phi_n(u^3, u^4, \cdots u^n)$ being its lower bound. Denote the objective function $f(x)$ where x is a feasible solution to the original N-dimensional assignment problem. For the optimal dual solution $(\hat{u}^3, \hat{u}^4, \cdots \hat{u}^n)$ with the corresponding $\hat{x}$ being a feasible solution for the original problem, we have $f(\hat{x}) \geq \Phi_n(\hat{u}^3, \hat{u}^4, \cdots \hat{u}^n)$. The value $(f(\hat{x}) - \Phi_n(\hat{u}^3, \hat{u}^4, \cdots \hat{u}^n))/\Phi(\hat{u}^3, \hat{u}^4, \cdots \hat{u}_n))$ is an approximate measure of the **duality separation** for the problem and is an important measurement of the performance of the algorithm.

2.2 Inner Problem

In the Lagrangian relaxation algorithm we maximize the dual of the recovery problem directly after maximizing the relaxed dual. As a modification to the Lagrangian relaxation algorithm we add an intermediate step called the **Inner Problem**. Suppose we have a multiplier vector $(u^3, u^4, \cdots u^n)$ which maximizes $\Phi_n(u^3, u^4, \cdots, u^n)$ together with an index alignment for the first two frames $\alpha = \{(i_1(j), i_2(j)) | j = 0 \cdots M_0\}$. The inner problem is constructed by removing from the original N-dimensional assignment problem all the arcs with first two indices not in α. The inner problem is denoted $\Phi_n(u^3, \ldots, u^n; \alpha)$. For the same multiplier vector $(u^3, u^4, \cdots, u^n)$, we can prove that the value of the dual problem is less than the value of the inner problem. And a similar result holds for later stages. Thus the optimal value of the inner problem is a better lower bound for the final feasible solution than the optimal value of the dual problem.

Also, numerical results show that if the multiplier vector $(v^{k+2}, \cdots v^n)$ maximizes the inner problem, then $(v^{k+2}, \cdots v^n)$ is generally a better initial multiplier vector to start the maximizing of dual of the recovery problem with and leads to less running time and improved solution quality.

3 The Partial Branch-and-Bound Algorithm

3.1 Heuristic for Choosing Index Alignment

The function $\Phi_n(u^3, \ldots, u^n)$ is concave and very flat near its minimal. Since the subgradient of the dual problem corresponding to a particular multiplier is the number of times the relaxed constraint is violated, the greater the norm of the subgradient, the greater the duality separation may be. Thus, in choosing the multiplier vector to carry on the recovery procedure, both the dual function value and the subgradient norm should be considered. If two multiplier vectors have different subgradient norms, we choose the one with smaller norm, otherwise we choose the one with better objective function value.

The above discussion can be stated in this way: given two multiplier vectors each associated with a subgradient vector, the one with the smaller subgradient vector norm has a better possibility to recover a better feasible solution.

3.2 Partial Branch and Bound

The improvement in solution quality and the computation of K-near optimal solutions is accomplished via an enumerative technique that is framed within the branch-and-bound paradigm. The goal is to do a partial enumeration by selectively choosing which branches to examine.

During the non-smooth optimization procedure for stage k we come up with a set of p_k multiplier vectors where p_k is the number of non-smooth optimization iterations in stage k (in each iteration one new multiplier vector is generated and an index alignment is calculated). Associated with each multiplier vector is an index alignment. If we pursue further iterations from each of these index alignments, the possibility of finding a very good feasible solution will increase greatly.

Let Y denote all the feasible solutions for the original problem and consider the following partition:

$$Y = Y_1 \cup Y_2 \cup Y_3 \cup \cdots \cup Y_{P_1} \cup \bar{Y} \qquad (6)$$

There is an index alignment α_i of the first two frames for each Y_i. Let $\bar{Y}$ denote the set of all the feasible solutions that don't conform to $\alpha_1, \alpha_2, \cdots$ or α_{P_1}, i.e., the unlisted solutions. Because of its size, it is less likely to find a good solution from $\bar{Y}$. Thus it is not examined. For each branch Y_i, we form the recovery problem, perform the non-smooth optimization process and make partitions again. We continue with the partition process till we arrive at the 2-D assignment problem.

After one has obtained a feasible solution, it is possible to delete some branches by comparing this feasible solution value to a lower bound of the feasible solutions conforming to that branch. If the lower bound is greater than the best primal feasible solution, further computation on this branch is unnecessary, i.e, there is no chance of obtainning a better solution from this branch.

A lower bound for all the feasible solutions contained in the branch $Y_{i_1 i_2 \cdots i_k}$ is computed by forming the inner problem on this branch and finding its maximum. As stated before, the reason for introducing the inner problem is to get a better lower bound, and if it is necessary to pursue this partition, a better multiplier to start the next stage of non-smooth optimization with.

To find multiple solutions, a solution buffer S with a predetermined size is set up. It keeps the best feasible solutions ever found. In this case if S is full, deleting some branches is done by comparing the lower bound of that branch to the worst solution in S. Otherwise when S is not full we don't cut out any branches.

4 Numerical Results

4.1 Problem Generation

Our algorithm is designed for the multitarget tracking environment. All the test problems

used here are generated from a tracking simulation application. Tracking simulation consists of two parts:

1. Modeling. Random tracks and corresponding reports are generated in 3D space. Noise, misdetections and false alarms are added.

2. Filtering and Scoring. Possible tracks (hypotheses) are filtered and scored [2]. Those with a high enough likelihood ratio are kept. These are the arcs in the multidimensional assignment problem.

The number of observations varies in each scan. In most cases they are close to each other. The dimension of the assignment problem is the tracking window size.

4.2 The Index Alignment Selection Heuristic

First a six dimensional assignment problem is generated with 22,518 arcs. This is one of the more complex and yet reasonably sized problems that we have managed to create using our tracking simulator. Each scan has about 40 observations. The Lagrangian relaxation algorithm is performed with and without the index selection heuristic. Because the non-smooth optimization procedure converges slowly, we iterate only a given number of iterations before terminating.

Figure 1 shows the solution quality for both algorithms. The optimality is measured by comparing to the best objective function value we have ever computed. For the basic Lagrangian relaxation algorithm, there are two big variations after 80 steps, which may deteriorate the solution quality to as low as 94% of optimal. In the Lagrangian relaxation algorithm with alignment selection, the solution quality stays above 98% after 90 steps, and in the worst cases the solution quality stays on 97%. Of all the 26 results, 5 results of basic Lagrangian relaxation algorithm are above 99% of optimal, while for alignment selection this number is 15. 13 results of basic Lagrangian

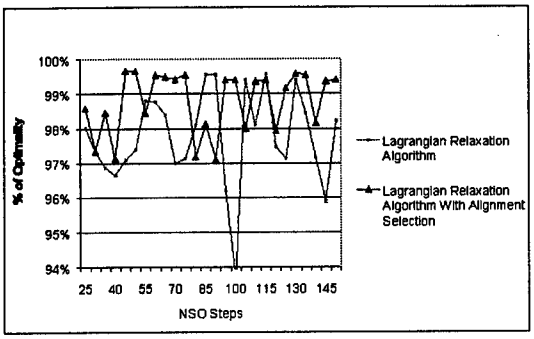


Figure 1: Comparison of Alignment Selection for Different Number of NSO Iteration Steps.

relaxation algorithm stays below 98%, while for alignment selection this number is 5, which all happens before 90 steps. This figure shows that the solution quality is improved and stablized, which justifies the validity of the index alignment selection heuristic.

It should be pointed out that during the first 20 steps of non-smooth optimization iterations, alignment selection doesn't appear to be superior to the basic Lagrangian relaxation algorithm. The reason is that when we apply our selection heuristic, we make the assumption that the dual function value is close to its optimal, so it is the subgradient vector norm that plays the major role. Indeed during the early period of non-smooth iteration the dual function value is far away from the optimal point, which misleads the alignment selection heuristic. It is suggested to allow non-smooth optimization close to the optimal point when applying the index alignment selection heuristic. Sometime this will takes lots of time. One way to deal with this is to set a large accuracy tolerance while solving the non-smooth optimization.

To further show the index alignment selection heuristic works, another six assignment problems are tested in table 2. Each of them is of 5 dimension with around 20 observations per scan. During computation we allow non-smooth optimization to converge to its optimal value. For each problem the lower bound and

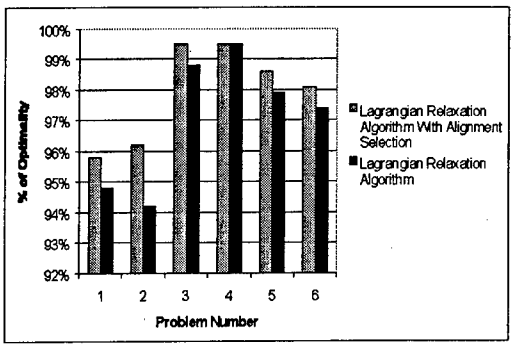


Figure 2: Comparison of Alignment Selection for Different Problems.

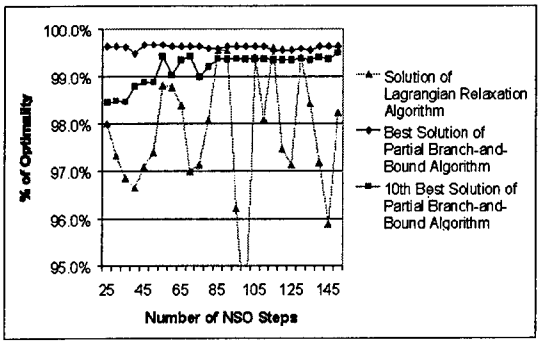


Figure 3: Results of Partial Branch-and-Bound Algorithm for Branching Number=5, Seeking 10 Best Solutions.

the two feasible solutions are solved both with and without applying the index alignment selection heuristic. The results are shown in figure 2. Alignment selection improves the solution quality for 5 out of 6 cases and for 1 it remains the same.

It is shown that if we allow non-smooth optimization to converge, the algorithm with index selection always outperforms the one without selection.

4.3 The Partial Branch-and-Bound Algorithm

There are three parameters for partial branch-and-bound algorithm. The first is the number of solutions desired. Remember that if the solution buffer is full, the bound of the branch-and-bound algorithm is set to the worst feasible solution in the solution buffer. Otherwise if the solution buffer is empty, no branches are cut. Increasing this parameter will slow down the algorithm by increasing the number of branches searched, i.e., the more solutions one desires, the slower the program runs. In this test we set the solution number to 10.

The second parameter is the branching number, which is the number of branches enumerated at each stage. The more branches enumerated, the better solutions we will find. Increasing it will cause more branches to be enumerated, which increases the running time. We will show later that the running time increases less than linearly.

The last parameter is the number of non-smooth optimization iterations performed. Performing fewer non-smooth optimization iterations tends to shorten the running time of maximizing each subproblem, but decreases the quality of their solutions, which may cause more subproblems to be solved due to the inefficient branch cutting. Like the effect of changing non-smooth iterations maybe comprehensive, i.e., it may increase or decrease the running time. Generally it is required that non-smooth iteration should be performed until no major improvement for dual problem can be found. This can be achieved by setting an adequate ϵ in the non-smooth optimization solver.

We applied the partial branch-and-bound algorithm on the complicated problem we used in figure 1. The branching number is set to 5, which is moderate for good solution quality and fast speed. The 10 best solutions are desired throughout these test cases. For different non-smooth optimization steps the results from partial branch-and-bound are shown in figure 3. The results from the Lagrangian relaxation algorithm are also shown in figure 3 for comparison. Running time comparison are shown in figure 4.

Compared to figure 1 the solution quality of partial branch-and-bound algorithm is greatly improved. The best solutions yielded from

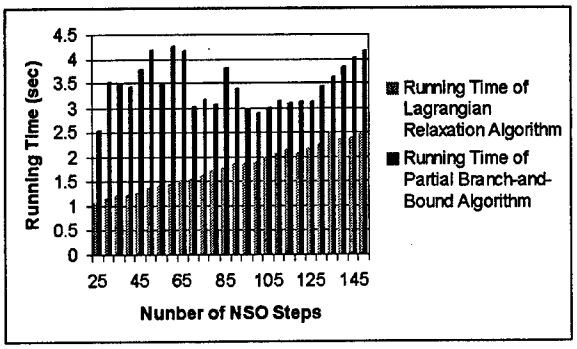


Figure 4: Running Time of Partial Branch-and-Bound Algorithm.

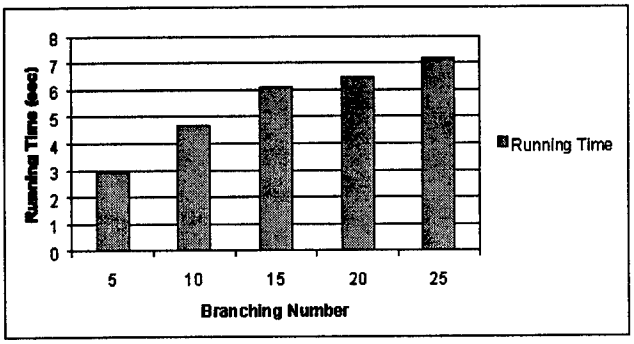


Figure 5: Running Time Comparison for Different Branching Number.

partial branch-and-bound algorithm are always better than Lagrangian relaxation algorithm. For all cases the best solutions are within 0.5% close to optimal. Quality of the 10th best solutions increases steadly as the number of non-smooth optimization iterations increases. In 21 out of all 26 cases, the 10th best solutions of partial branch-and-bound algorithm are better than the solutions from the Lagrangian relaxation algorithm.

The solution quality of partial branch-and-bound algorithm is very stable. There are no steep jumps in the solution quality as the number of non-smooth optimization varies. Also both the best solutions and the 10th best solutions have a stable increasing trend as the number of non-smooth optimization increases. It is shown in figure 4 that partial branch-and-bound algorithm may cost 50% more time than the Lagrangian relaxation algorithm, while it gives significantly better solutions to the problem.

The running time of partial branch-and-bound algorithm is not proportional to the number of non-smooth iterations taken. As stated before, this is because finding a better lower bound for the original problem will possibly cut off more branches, which helps increasing the speed. Figure 5 shows the running time of partial branch-and-bound algorithm for different branching numbers. The number of NSO iterations is set to 100. Still 10 best solutions are desired. The running time increases less than linearly as branching number increases. Its trend tends to become flatter as the branching number increases.

Here it is worth mentioning that we have tried to generate multiple solutions based on the k-best 2-D assignment solutions, which generates the k best solutions of the 2-D assignment problem encountered in each stage. It turns out to be an unsuccessful algorithm since the k best solutions from the same 2-D assignment problem are so close to each other that if one of them leads to a bad feasible solution, others seldom yield good feasible solutions.

5 Conclusion

In this paper we have presented a variant of the Lagrangian relaxation algorithm for constructing multiple quality solutions to the multidimensional assignment problem. An enumeration algorithm based on the branch and bound framework with a special selection heuristic algorithm has been developed with appropriate branching and ordering rules. Compared to continuation of iteration to reduce the duality separation, the new algorithm generates superior solutions in less time.

6 Acknowledgement

This work was supported in part by AFOSR Grant No. F49620-97-1-0273 and by ONR Grant No. N00014-99-1-0118.

References

[1] A. B. Poore and A. J. Robertson, "A new lagrangian relaxation based algorithm for a class of multudimensional assignment problems," *Computational Optimization and Applications* , 1997.

[2] A. B. Poore, "Multidimensional assignment formulation of data association problems arising form multitarget and multisensor tracking," *Computational Optimization and Applications* **3**, pp. 27–57, 1994.

[3] T. N. Barker, J. A. Persichetti, A. B. Poore, Jr., and N. Rijavec, "Method and system for tracking multiple regional objects." US Patent Number 5,406,289, issued 11 April 1995.

[4] A. B. Poore, Jr., "Method and system for tracking multiple regional objects by multi–dimensional relaxation." US Patent Number 5,537,119, issued 16 July 1996.

[5] A. B. Poore, Jr., "Method and system for tracking multiple regional objects by multi–dimensional relaxation." U.S. Patenting Pending, CIP, to issue in 1999.

[6] R. Alexander, *A Class of Lagrangian Relaxation Algorithms for the Multidimensional Assignment Problem.* PhD thesis, Colorado State University, Fort Collins, CO, 1995.

[7] K. G. Murty, "An algorithm for ranking all the assignments in order of increasing cost," *Operations Research* **16**, pp. 682–687, 1968.

On Features and Attributes in Multisensor, Multitarget Tracking*

Oliver E. Drummond, Ph. D., P. E.
Independent Consulting Engineer
10705 Cranks Road, Culver City, CA 90230 USA
E-mail: Drummond@Worldnet.ATT.Net

Abstract

An obvious use for feature and attribute data is for target typing (discrimination, classification, identification, or recognition) and in combat identification. Another use is in the data (or track) association process. The data association function is often decomposed into two steps. The first step is a preliminary threshold process to eliminate unlikely measurement-track pairs. This is followed by the second step, the process of selecting measurement-track pairs or assigning weights to measurement-track pairs so that the tracks can be updated by a filter. The primary concern of this paper is the use of feature and attribute data in the data association process for tracking small targets with data from one or more sensors.

1. Introduction

Target tracking problems can be broadly categorized into four generic classes [1], as follows: 1. sensor tracking of a single (bright) target, 2. tracking of targets that are large, 3. tracking of targets that are medium sized, and 4. tracking of targets that are small. These four classes are described in more detail in [2]. Note that the size indicated in this list is in terms of the number of resolution elements or pixels. The algorithms used in the signal, image, and track processing for each of these problems differ. A major concern in tracking small targets is the data association function.

Since each class of tracking problem poses different algorithm development issues, this paper will concentrate on only one class of tracking, namely, tracking of small targets using multiple target tracking methods. Multiple target tracking is a relatively new field. The first book dedicated exclusively to multiple target tracking was published in 1986 [3] and a number of recent books are available [4,5,6]. In addition to the numerous papers and reports in the open literature (too numerous to be listed here), there is an on-going series of annual SPIE conferences concerned exclusively with signal and data processing of small targets that started in 1989 [7]. This paper freely extracts and paraphrases material from some of the author's prior documents [1,8,9]

For this paper, a small target is characterized as one that does not provide enough data for traditional automatic target recognition (ATR) using a single frame of data [8]. In contrast, a target large enough for ATR typically extends beyond a diameter of about 15 resolution elements, for example, larger than 10 by 10 pixels square. Note that it is not uncommon to refer to all objects as targets whether they are of interest or not. Small targets of concern in this paper include point source targets and small extended targets including unresolved closely spaced objects.

A number of different theories could be used for developing algorithms for processing features and attributes. This paper uses Bayesian probability methods and addresses only track maintenance in order to limit the paper length. The primary tracking function of interest is data association and neither target typing or combat identification is addressed.

2. Features and Attributes

Although tracking small targets is a relatively new field, processing methods developed for tracking a target's trajectory, i.e., kinematic tracking, is fairly mature compared to the processing methods developed for using feature and attribute data in tracking. The term measurement (return, report, observation, or signal processing threshold exceedance) refers to all the data obtained by the signal processor or simply the measurement vector and its error covariance matrix, depending on the context.

As used here the term *feature* refers to characteristics of a target that are from continuous sample space and are obtained from sensor data that are other than the simple variables of position and its derivatives that are

* This work was supported in part by ONR Grant N00014-97-1-0570

used for kinematic tracking. Examples of features include estimated target dimensions, radar cross section, and other target signature data. Note that it may be that features are not measured directly but are computed based on a number of measured quantities. Whether the features are measured directly from the data of a signal-processing threshold exceedance or computed based on a number of measured quantities of a signal-processing threshold exceedance, in both cases the resulting feature vector and its error covariance matrix will be referred to as the *measured feature.*

By comparison, the term *attribute* will be used here to refer to characteristics of a target based on sensor data that are from discrete sample space, for example, literal, categorical, or integer parameters. Examples of attributes include target type, type of radar systems used by a target, and number of engines on an airplane. These particular definitions were chosen because, as defined, feature data and attribute data are processed differently because their uncertainties are treated differently. Using Bayesian probability methods, features can be processed based on their probability density while attributes can be processed based on their discrete probabilities or point masses.

There are a number of forms that attribute information can take and the form depends in part on how the attributes are to be processed. Due to space limitations, the approach taken for this paper is limited to using probabilistic attribute vectors of a specific type for both the processed attribute state and the measured attributes. In this form, a probabilistic *attribute vector* contains the probability or likelihood of each of the possible attributes.

What corresponds to the estimated state vector for kinematic tracking is what will be referred to as the *processed attribute state vector* that contains the *a posteriori* probabilities of each possible attribute. What corresponds to the measurement vector in kinematic tracking is what will be referred to as the *measured attribute vector* and it contains the likelihood of each of the possible attributes based on measured attributes or on attributes computed from sensor measurements of an apparent target. The likelihood for an attribute in this form is the probability of obtaining the phenomena observed (measured) by the sensor given that the apparent target exhibits that specific attribute. Note the term *apparent target* is used because what appears to be a target may actually be due to false signals, persistent clutter, or sensor phenomena not directly and completely due to a single target.

Note that some sensor processors make a hard decision for the measured attributes and that could be represented by the probability of one for the identified attribute. However, these sensor processor decisions will typically exhibit some decision errors. Assuming that an average value of probability of a decision error can be estimated empirically for a sensor, it can be used to convert a sensor processor's hard decision into a probabilistic attribute vector that contains the probability of each of the possible attributes. If the probability of a decision error is P_e then the probability of a correct decision is $P_d = 1\text{-}P_e$. Accordingly, the attribute vector would contain the value P_d for the attribute identified by the sensor processor based on measurements of an apparent target. The attribute vector would contain a value of P_e for all other possible attributes.

In addition to features and attributes, there is another class of data that has some characteristics of both attributes and features. The term that will be used here for this type of data is categorical features. *Categorical features* are from continuous sample space (possibly bounded) but they are based on known characteristics of the targets and sensors that allow classified into a finite number of classes or categories. The continuous sample space is caused by either random measurement errors or by the distribution of the inherent parameters of each type of target that cause the features that are measured, or both. An example of a categorical feature is the estimated wing span of an aircraft given there are only a few types of aircraft in the field of regard, the wing span of each type of aircraft is known *a priori*, and the sensor obtains measurements with measurement errors from which the wing span of a tracked target can be estimated based on a single look by a sensor.

Note that as with features, it could be that the categorical features are not measured directly but are computed based on a number of measured quantities. Whether the features are measured directly as some of the measured quantities of a signal-processing threshold exceedance or computed based on a number of measured quantities of a signal-processing threshold exceedance, in both cases the resulting feature vector and it covariance matrix will be referred to as a *measured categorical feature.*

In a real tracking system application, the difference between features (as first defined) and categorical features may be muddied. For example, the characteristics of a feature for most targets might be known but not known for other targets. In fact, depending on why features are processed, the

distinction between features and categorical features may have little meaning. For this paper, the term categorical feature is defined for the purpose of facilitating the discussion of how features are processed. Using Bayesian probability methods, categorical features can be processed using composite estimation that uses multiple models based on a hybrid estimation method that combines probability densities and discrete probabilities [1,5,9,10,11]

3. Preliminary Thresholding: "Gating"

The data association function can be viewed as a two-step process, namely, the preliminary thresholding in tracking is frequently the first step in the data association function in a target tracker and is sometimes referred to as *gating* or *gate processing*. This first step is followed by the second step that is the process of selecting measurement-track pairs or assigning weights to measurement-track pairs so that the tracks can be updated by a filter. The filter might be a Kalman filter or extended Kalman filter (or an equivalent filter or maybe even an approximation thereto). Note in target tracking the filter update is typically also decomposed into two steps. The first step is the time update to predict the state and the measurement to the time that the next measurement (or set of measurements) was actually observed. After the two steps of the data association, the filter measurement update is performed and the track data is stored in the track files.

3.1 Background: Gating in Kinematic Tracking

In the discussion that follows, for emphasis and clarity, a measurement used for kinematic tracking is referred to as a *kinematic measurement*. The elements of a kinematic measurement typically consists of the measurements of one or more of the following: range, azimuth, elevation and range rate plus their error covariance matrix.

In many tracking systems, the only purpose for the preliminary thresholding is to reduce the processing load. In kinematic tracking, a region in measurement space is identified that is centered at the predicted position of a target where the measurement is expected to be for that target. That region is the track gate and the size of the region can be established in a number of ways. The method used to size the gate depends on the type of information available. The size of the gate depends on the variance of both the vector of the measurement errors and the vector of the predicted target state, often just position components. For example, a 99.7 % gate would be sized so that the correct measurement for a track would be in its gate with a .997 probability. A more effective gate size could be computed using the formula of Eq. 4.7 of [3]. Only measurements that fall within the track's gate, i.e., within the identified region of measurement space, are used in the subsequent data processing for that track.

As an aside, note that there can be important computational considerations in designing the gate processing for a tracker [1]. If there are more than a few targets in the field-of-view, then the process of determining which measurements are in each track's gate (the "gate search" process) can be computationally intensive if simple brute force methods are used. With more than a few targets, simplistic gate search methods should be avoided. In addition, elliptical (ellipsoidal or hyper-ellipsoidal, as appropriate) gates are usually more effective but are also more processor intensive than are rectangular (or hyper-rectangular, as appropriate) gates.

A hyper-ellipsoidal gate process usually involves computing a chi-square statistic (or an approximation of it) of the innovations that is compared to a threshold value. Computing a chi-square statistic typically requires a matrix inversion, multiplies, and additions. In contrast, a hyper-rectangular gate process typically does not require a matrix inversion, and involves only adds, compares, and at most a few multiplies. Thus with more than a few targets, it is advisable to use two gates in series, the first is an oversized hyper-rectangular (or rectangular) gate. The measurements in that track gate, i.e., that pass this first threshold test, are then processed using a second track gate that is a hyper-ellipsoidal (or elliptical or ellipsoidal, as appropriate) gate [1].

3.2 Some Assumptions

The discussion of the gate processing for kinematic processing provides background for the preliminary threshold processing of features, attributes, and categorical features.

To facilitate the discussion, the set of attributes will be assumed to be mutually exclusive and exhaustive. The techniques that are described can be readily adapted to the more general case. The assumption for all random variable from continuous sample space is that any deviation of their probability density function from Gaussian can be neglected. Furthermore, the assumption initially is that either attributes, features or categorical features are obtained with or without kinematic measurements. If there is an attribute obtained with a kinematic measurement, the assumption is that they are statistically independent.

Also, the initial assumption is that features, attributes and categorical features are static, i.e., for a target they do not change over time except for changes due to the errors in measuring them and, if applicable, in estimating them from measurements. It is also assumed initially that the kinematic track filter does not employ multiple models. Many of these various assumptions can be relaxed and these methods adapted to handle the less restricted cases.

Since the track stage of interest is track maintenance, unless indicated otherwise, the assumption is that the *a priori* information for both the target state and all discrete alternative or hypothesis has already been incorporated into each track, in applicable.

3.3 Preliminary Thresholding of Features

Since features are much like kinematic measurements, features can be processed in much the same way. If the errors of a feature vector are cross-correlated with the errors of an accompanying kinematic measurement vector, then they should be processed as a single vector, including the filtering process. This estimated state vector should be a concatenation of the kinematic states and feature states. If filtered this way, then a properly designed Kalman (or similar) filter should provide consistent covariance matrix of the estimation errors of the estimated kinematic states and the feature parameters.

Computing the vector consisting of the predicted kinematic measurements and predicted features and also computing the covariance matrix provides (along the kinematic measurements and features) the information needed to compute both a hyper-rectangular gate and a hyper-elliptical gate for a track. Thus this processing is identical to gating in measurement space except that it is a higher dimensional space and hence involves more computationally complex processing. The threshold value is computed as discussed in Section 3.1.

If the kinematic measurements and the features are independent, then the processing can be simplified somewhat. The features can be filtered separately from the kinematic measurements. Then for the hyper-rectangular gate processing for a track and a measurement, the magnitude of each element of the kinematic-measurement's innovations vector can be tested in turn against its threshold followed by similar testing of each element of the feature innovations vector. Note that the order of the processing of these two vectors can be reversed or even interleaved, if that ordering is more effective for a tracking system application. If any element of these two vectors fails its test, then that measurement is considered not a potentially valid measurement for that track. The kinematic innovations vector is the difference between the kinematic measurement vector and its predicted vector. The feature innovations vector is the difference between the predicted feature and the measured feature vector. The threshold used for an element of an innovations vector is proportional to the standard deviation ofthat element based on the innovations covariance matrix. For the hyper-ellipsoidal gate processing for a track and a measurement, two chi-square statistics can be computed separately, one for the kinematic measurements and the other for the features. These two can then be added and compared to the appropriate threshold.

Note that a chi-square statistic is used because in kinematic tracking it is assumed that any deviation of the innovations from exhibiting Gaussian characteristics can be neglected. Furthermore, even if the true probability density of the innovations were known and were not Gaussian, then in most cases it would be too processor intensive to use the proper statistic instead of chi-square. In processing features, however, it may be that the innovations for some features are clearly not Gaussian and the above assumption should be revisited.

An elliptical (or hyper-ellipsoidal) gate is used in gate processing because it is obtained mathematically (in addition some constants) by computing minus the logarithm of the likelihood function that a specific measurement is due to the target of a specific track. Methods for computing an appropriate threshold for hyper-ellipsoidal gates have been studied extensively and are available, although there are some practical limitations [3,12]. The *a posteriori* probability that a measurement is due to the target of a specific track is not used in gating because it depends on complicated computations that involve all the measurements and tracks and so that would defeat the purpose of the gating process.

3.4 Preliminary Thresholding of Attributes.

The gate processing of attributes appears to be very different from for kinematic measurements or features. Consider a "minus log likelihood" approach to the gate processing of attributes that is analogous to the gate process used for kinematic measurements and features. Devising such an approach raises the issue of what to use for a threshold value. If the purpose of the gate process is to eliminate unlikely track-measurement pairs then if there is no rational method to compute a threshold for attributes, then there is no purpose to including attributes in the gate process.

This paper proposes approaches for computing this threshold value. First the minus log likelihood computation is described and then methods for computing the threshold are addressed.

Given a vector of attribute probabilities (the measured attribute vector) based on a single current (or recent) measurement and a track including its processed attribute state vector obtained from prior measurements, a scalar can be computed for use in the attribute gate process. The scalar envisioned is the inner product of these two vectors, namely, the inner product of the processed attribute state vector of the track and the measured attribute vector. The resulting scalar, which is the attribute likelihood is in effect

$$p[a_m(n)|A_j(n-1), j\sim m] \qquad (1)$$

where

j = track index

n = time index

m = measurement index (for time *n*)

$a_m(n)$ = phenomena of measurement *m* used to compute the measured attribute vector

$A_j(n-1)$ = phenomena of all measurements up to time *n-1* used to compute the processed attribute state vector for track *j*

and *j~m* means that measurement *m* is from the target of track *j*. The final computation is to compute minus the logarithm of this scalar to obtain the minus-log likelihood for that measurement-track pair.

An appropriate threshold is to compute a scalar that is computed in the same way as the attribute minus log likelihood except that the vector of attribute *a priori* probabilities of false signals is used instead of the processed-attribute state vector of the track. This requires that a reasonable value be obtained for the probability of each attribute for false signals on average.

If it is not practical to obtain a realistic attribute *a priori* probabilities of false signals, then there are a number of alternatives that can be considered. One alternative is to use for the attribute *a priori* probabilities of false signals a vector with all the alternative attributes equally probable. That is, if there are *k* attributes then the *a priori* probability of each possible attribute for a false signal is assumed to be simply *1/k*.

Another more conservative alternative is to use what will be referred to as the complementary probability vector. The *complementary-probability vector* used for the attribute *a priori* probabilities of false signals is computed as follows. Form a vector of ones with the same number of elements as the processed attribute state vector and subtract the processed attribute state vector from it. Then normalize the resulting vector by computing the sum of its elements and dividing each element of that vector by that sum to obtain the complementary probability vector. This vector could then be used for the attribute *a priori* probabilities of false signals to compute the threshold.

Yet another alternative is to just not include attributes in the gating process. However, there may be the need for a practical threshold value for attributes in the second step of the data association process after the gating process, so the above methods might be used for that purpose even if attributes are omitted from the gate processing.

How the threshold processing that is used depends on the type of measurement that is obtained. In most cases the measurement that provides attribute data will also provide a kinematic measurement vector. For that case, the attribute minus log likelihood can be added to the kinematic minus log likelihood (the chi-square statistic) and compared to the appropriate threshold. The appropriate threshold would then be the sum of the attribute threshold (as discussed above) and the kinematic threshold. If the sum of the minus-log likelihood functions is larger than the sum of the thresholds, then that track-measurement pair is not included in the second step of the data association processing. The processing just described is analogous to hyper-ellipsoidal gate processing.

The hyper-ellipsoidal gate can be preceded by a hyper-rectangular gate. For hyper-rectangular gate processing for a measurement-track pair, the order of the processing must first be established. The attributes could be processed before the kinematic measurements of visa versa. The most effective processing order to use and the most effective processing order of the individual measurements in the kinematic measurement vector depends on the specific characteristics of the sensors and targets.

If the kinematic measurements are processed first, then the magnitude of innovation corresponding to each element of the kinematic measurement vector would be processed in turn and compared to its threshold. If the all these innovation magnitudes are less than their threshold, then the attribute minus-log likelihood function would be tested against its threshold. Any measurement-track pair that passes all these threshold tests would then be processed using

the hyper-ellipsoidal gate.

If there are features in addition to kinematic measurements, then they too can be processed along with the kinematic measurements as discussed in Section 3.3. Thus the gate processing to handle kinematic measurements, attributes, and also features would be much like the processing just described for kinematic measurements and attributes. If on the other hand, there are only attributes for a measurement and no features or kinematic measurements, then a single attribute threshold process would serve in place of both types of gates, hyper-ellipsoidal and hyper-rectangular. The extension of the preliminary threshold processing discussed in this section to handle multiple sets of attributes, i.e., multiple attribute vectors that are independent, is straight forward.

3.5 Preliminary Thresholding of Categorical Features

In their simplest form, there are two classes of categorical features. With the simpler of these two classes, call it Class 1, the value for the feature vector (or scalar) is know *a priori* for each alternative category or hypotheses, i.e., the inherent features for a category, are fixed and deterministic. With the other class, call it Class 2, the values of the feature mean vector (or scalar) and its covariance matrix are know *a priori* for each alternative category or hypotheses. This covariance matrix is the so called "within class" (in this case "within category") covariance matrix that reflects the variation about the mean of the true feature across targets for a category.

More generally, for Class 3, the mathematical model for the measurement equation and possibly also the dynamic equation might be different for each alternative category or hypothesis. Also the categorical feature state vector need be the same length as the measured feature vector. All three of these classes of categorical feature problems can be processed using non-switching (static) multiple-model methods [1,5,9,10,11] if it is assumed that the feature characteristics do not change over time for a target.

Yet another aspect of processing categorical features is the dependence of the kinematic measurements on the categorical features. There are two distinctly difference types of dependencies. First the characteristics of the kinematic measurements may or may not depend on the feature category for a target. Alternatively, the feature category for a target could depend on the kinematic measurements or the kinematic state, but that can be even more complex and will not be discussed here due to page limits. Note that an even more complex relationship is conceptually possible where the dependency of the kinematics and the feature category for a target is in both directions.

The second type of dependency is between both the estimation errors of the kinematic state and kinematic measurement errors and the measurement errors of the measured categorical features for a target. Remember that it may be that the measured categorical features are not measured directly but rather might be computed from data obtained in conjunction with a measurement of an apparent target.

For a particular system application there are four possible combinations of these two types of dependencies and the processing method for one type may not be the best for another. Considering these four possible combinations of dependencies along with the 3 classes of categorical features could lead to 12 different processing methods to be explored.

To simplify this discussion, only three of these combinations will be addressed. First the simpler case of the errors of the measured categorical features independent of both the kinematic measurements and kinematic estimated state for a target and also independence between feature category and both the kinematic measurements and kinematic estimated state for a target will be addressed. This case will be discussed for the two simpler categorical feature classes.

For this simpler case and for all three classes of categorical features, the kinematic measurement is processed in a filter separately from the feature filtering, if applicable, for each category. Also a kinematic chi-squared statistic is computed from the kinematic innovations for a track-measurement pair independently of the categorical feature data.

For the Class 1 categorical features, a filter in not needed because the values for the inherent categorical features for a track are known *a priori* for each feature category. The difference between the measured categorical feature vector for a measurement and the *a priori* value for the feature vector for a feature category serves as the innovations vector for a feature category for a track-measurement pair. For each track-measurement pair the chi-square statistic is computed for every feature category. Note that for this case the processing can be simplified because the categorical feature chi-square statistic does not depend on the tracks, only on the measurements and the feature categories. Accordingly, the categorical feature chi-

squares can be computed for each measurement and all feature categories without using any track data. These chi-square statistics are then used along with the additional constants needed to compute the likelihood function of the features for each feature category.

These likelihood functions are used to compute the measured categorical feature vector. What corresponds to the measurement vector in kinematic tracking is the *measured feature category vector* that contains the likelihood of each of the possible feature categories based on measured features for a measurement and the feature *a priori* feature values for the feature categories. What corresponds to the estimated state vector for kinematic tracking is what will be referred to as the *processed feature category state vector* that contains the *a posteriori* probabilities of each possible feature category for a track based on processing all prior measurements.

The process that follows is like the processing of attributes. The scalar is computed, the categorical feature likelihood, that is the inner product of these two vectors, namely, the inner product of the processed feature category state vector of the track and the measured feature category vector. The final computation is to compute minus the logarithm of this scalar to obtain the categorical feature minus-log likelihood for that measurement-track pair.

The threshold used with the categorical feature minus-log likelihood is computed using a method that is similar to that used for attributes. An appropriate threshold value might be using *a priori* characteristics of false signals.

The threshold is computed in the same way as for the attribute minus log likelihood except for how the vector of feature category *a priori* probabilities of false signals and the *a priori* value for the feature vector for false signals are computed. First the difference between the measured categorical feature vector for a measurement and the *a priori* value for the feature vector for false signals serve as the innovations vector for a false signal. For each measurement the chi-square statistic is computed for every feature category. This needs to be computed only once for each measurement. These chi-square statistics for the feature categories for a measurement are then used along with the additional needed constants to compute the likelihood function for each feature category for false signals.

These likelihood functions are then used to compute the measured false signal categorical feature vector that contains the likelihood of each of the possible feature categories for a measurement-track pair. What corresponds to *measured categorical feature vector* is the *measured false signal categorical feature vector* that contains the likelihood of each of the possible feature categories based on the *a priori* false signal characteristics and a measurement. What corresponds to the processed feature category state vector is what will be referred to as the *feature category a priori probabilities vector for false signals* that contains the discrete *a priori* probabilities of each possible feature category for false signals. The threshold is computed by computing the inner product of the measured false signal categorical feature vector and the feature category a priori probabilities vector for false signals. Note that to use this method to compute the threshold requires that a reasonable value be obtained for the probability of each feature category for false signals on average and also the value of the category feature vectors for false signals for each category.

If it is not practical to obtain realistic categorical feature *a priori* information for false signals, then there are a number of alternatives that can be considered. One alternative is to use for the feature category *a priori* probabilities of false signals a vector with all the alternative categories equally probable as discussed in Section 3.4. For the chi-square values needed to compute the measured false signal categorical feature vector, the value of chi-square corresponding to cumulative probability of say 0.997 could be used, or what ever other value is appropriate for the track system at hand. Yet another alternative for the feature category *a priori* probabilities of false signals is to use the a *complementary-probability vector* for the feature category *a priori* probabilities of false signals. Thus there are a number of ways of computing the threshold value for the threshold for testing the categorical feature minus-log likelihood.

The categorical feature minus-log likelihood is processed along with the kinematic chi-square statistic (if kinematic measurements are available) in the same way as for attributes as outlined in Section 3.4 to complete the hyper-ellipsoidal gate processing. This gate processing also can be preceded by hyper-rectangular gate processing as outlined for attributes. Thus the preliminary threshold processing step for Class 1 categorical features is the same as for the processing attributes except for the computation of the measured categorical feature vector and of the threshold which do differ from the computations used for attributes.

The Class 2 category feature processing differs from the Class 1 processing because the values of inherent

features for a feature category are not deterministic. Rather, they are assumed to be characterized for each feature category by their mean and covariance matrix. There are a number of methods that can be used to process Class 2 categorical features and some are more efficient than others. The method described here is not necessarily the most efficient but is relative easy to describe so as to convey the concepts that apply. As mentioned previously, the method that follows is applicable to the case in which both the measurements and the measurement errors are independence of both the feature measurement errors and the feature category for a target.

The primary difference between the processing of Class 1 and Class 2 categorical features is that in Class 1 no processed categorical feature state vectors (one for each category) are maintained for each track but they are computed for each track for Class 2. In both classes, a processed feature category state vector is maintained for each track.

The difference between the measured categorical feature vector for a measurement and the predicted value for the processed categorical feature state vectors for a feature category for a track serves as the innovations vector for a feature category for a track-measurement pair. For each track-measurement pair the chi-square statistic is computed for every feature category. These chi-square statistics are then used along with the additional constants needed to compute the likelihood function of the features for each feature category for a measurement-track pair.

These likelihood functions are then used to compute the measured categorical feature vector that contains the likelihood of each of the possible feature categories for a measurement-track pair. The categorical feature likelihood is then computed as for Class 1 and the final computation is to compute minus the logarithm of this scalar to obtain the categorical feature minus-log likelihood for that measurement-track pair. The threshold value is computed and the hyper-ellipsoidal gate process is completed as with the Class 1 processing and can be proceeded by hyper-rectangular gate processing as for Class 1.

Finally, consider processing Class 3 categorical features with both types of dependency. That is, characteristics of the kinematic measurements depend on the feature category for a target and also both the estimation errors of the kinematic state and kinematic measurement errors depend on the measurement errors of the measured categorical features for a target. For this class of categorical feature problem the gate processing is as just described for Class 2 except that there is a single filter for each feature category for a target for both the kinematic and the categorical feature data. The kinematic and categorical features are processed simultaneously as a single vector for both the hyper-rectangular and hyper-ellipsoidal gate processes. The processing described for Class 2 applies by using the categorical feature processing that was described for both the kinematic and categorical feature measurements and similarly for the estimated states. Accordingly, this particular Class 3 problem could be processed as a non-switching multiple model problem in which the state is composed of both the kinematic state elements and the categorical feature state elements.

For gate processing, a few of the many kinds of problems that can involve kinematic, features, attribute, and categorical feature have been discussed. Of course there are more combinations that deserve attention than those discussed. Also the probabilistic derivations that are the basis for the processing methods described have not been presented due to space limitations. Finally, some of the simplifications that could further reduce the processing have not been discussed nor have the adaptation of these methods to less restricted problems in which some of the assumptions are relaxed.

4. Data Association, Step 2

The second step of the data association function is to select measurement-track pairs or assign weights to measurement-track pairs so that the tracks can be updated by a filter. There are a variety of algorithms for this process [1,2,3,5,6,7], including both single frame and the more complex multiple frame processing, such as multiple hypothesis tracking. In addition, there are hard decision approaches, such as (independent) nearest neighbor and most probable hypothesis tracking and there are soft decision approaches that are also called probabilistic or Bayesian method. While these approaches are all different, they can be classified for the purpose of this discussion into two groups. In their data association decision or weighting process, one group uses the minus-log likelihood function for each track-measurement pair. For the other group, the likelihood functions are used. Many single frame methods, for example, that make hard decisions use the minus log likelihood function. By contrast, soft decision methods use the likelihood functions for each track-measurement pair that survives the gate processing.

Given the minus-log likelihood for a measurement-track pair and the appropriate accompanying

constants, then the likelihood for that measurement-track pair can be computed. Accordingly, all the information needed for the second step of the data association process has already been addressed in Section 3 for the specific types of problems that were discussed. In many cases the hyper-elliptical gate threshold values discussed in Section 3 are also needed in the second step of the data association processing.

In addition to the data association function in track maintenance is the filtering function. If the processing included either attributes or categorical features, then an additional process function is needed to supplement the filter process. That additional function is to update the processed attribute state vector and the processed feature category state vector, if applicable. This processing can be accomplished using a straight forward application of Bayes rule. The other track maintenance functions are track promotion-demotion logic and track management which should not require any major modifications (at least conceptually) beyond those used for processing kinematic data to accommodate features, attributes, and/or categorical features.

5. Conclusions

In this paper, the types of measurement data used for multiple target tracking with data from multiple sensors has been classified into four types, namely, kinematic, feature, attribute, and categorical features. The motivation for this classification scheme was to partition the types of data according to how it might be process in a tracker because different processing methods are required depending on the characteristics of the data. Processing approaches have been outlined that illustrate how the processing might differ if features, attributes or categorical data were available in addition to kinematic data. The form of the state that corresponds to each of these data types was also shown to depend on the data type.

The paper introduces methods for computing the threshold for the gate processing for attributes and categorical features that are substantially different from methods used for kinematic and feature measurements. Material left for subsequent documentation include the derived equations for the processing methods presented, identified methods to further simplify the processing, describing the processing for other combinations of the types of data, and extension of the processing methods to accommodate relaxation of the assumptions used for the purpose of this paper.

Acknowledgements

Discussions with Dr. Rabinder Madan were very helpful in establishing the direction for this study.

References

1. Drummond, O. E., *Multiple Target Tracking Lecture Notes*, (a sequence of editions since 1985) Technical Book Company, 2056 Westwood Blvd., Los Angeles, CA 90025, 1998.

2. Drummond, O. E., Target Tracking, *Wiley Encyclopedia of Electrical and Electronics Engineering*, Vol. 21, pp. 377-391, 1999.

3. Blackman, S. S., Multiple Target Tracking With Radar Applications, Denham, MA: Artech House, 1986.

4. Bar-Shalom, Y. and X. R. Li, *Estimation and Tracking: Principles, Techniques and Software*, Boston: Artech House, 1993.

5. Bar-Shalom, Y. and X. R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques,* Los Angeles: OPAMP Tech. Books, 1995

6. S. S. Blackman and R. F. Popoli, *Design and Analysis of Modern Tracking Systems,* Norwood, MA: Artech House, 1999.

7. Drummond, O. E., (ed.), *Signal and data processing of small targets 1998*, *Proc. SPIE,* Vol. 3373, 1998, (also proceedings for earlier years starting in 1989).

8. Drummond, O. E., (ed.), Introduction, *Signal and Data Processing of Small Targets 1997*, *Proc. SPIE,* 3163, p. ix, 1997.

9. Drummond, O. E., Target Tracking With Retrodicted Discrete Probabilities, *Signal and Data Processing of Small Targets 1997*, *Proc. SPIE,* 3163: pp. 249-268, 1997.

10. D. T. Magill, D. T., Optimal adaptive estimation of sampled stochastic processes, *IEEE Trans. Autom. Control,* AC-b (4), pp. 434-439, 1965

11. Sims, F. L. and D. G. Lainiotis, Recursive algorithm for the calculation of the adaptive Kalman filter weighting coefficients, *IEEE Trans. Autom. Control,* AC-14 (2), pp. 215-217, 1969.

12. Drummond, O. E., Gate Size for Sequential Most Probable Hypothesis Tracking, *Signal and Data Processing of Small Targets 1994*, *Proc. SPIE,* 2235, pp. 670-682, 1994.

Best Linear Unbiased Estimation Fusion

Yunmin Zhu *
Department of Mathematics
Sichuan University
Chengdu, Sichuan 610064, P. R. China
ymzhu@scuu.edu.cn

X. Rong Li † (Corresponding Author)
University of New Orleans
New Orleans, LA 70148
Phone: 504-280-7416, Fax: 504-280-3950
xli@uno.edu

Abstract—A general version of the best linear unbiased estimation (BLUE) fusion rule is developed. It has the least mean-square (LMS) error among all linear unbiased estimation fusion rules. It is very general — it relies only on two assumptions: (1) the local estimators are unbiased and (2) the error covariance matrix C_k of all local estimates at each time k is known. Not only does it include existing fusion results as special cases, but it is also valid for many more general cases, including (1) coupled measurement noises across sensors; (2) sophisticated network structures or communication patterns; (3) different local dynamic models or estimator types; and (4) efficient fusion of asynchronized estimates. First, we formulate the problem of distributed estimation fusion in a general, i.e., BLUE, setting, which is the key to the other contributions of the paper. In this setting, the fused estimator is a weighted sum of local estimates with a matrix weight. We show that the set of weights is optimal if and only if it is a solution of a matrix quadratic optimization problem subject to a linear equality constraint. Secondly, we present a general solution to the above optimization problem, which depends only on the covariance matrix C_k. We also give the unique solution of the optimization problem, along with a necessary and sufficient condition for it to hold. Thirdly, we present an explicit formula of the optimal weights for the case in which C_k is nonsingular. We also discuss the generality and usefulness of the BLUE fusion formulas developed. Finally, we provide an off-line recursion of C_k for a class of multisensor linear systems with coupled measurement noises.

Key Words: fusion, distributed estimation, best linear unbiased estimation

*Supported in part by The National Key Project and Natural Science Foundation of China.

†Supported in part by ONR via grant N00014-97-1-0570, NSF via Grant ECS-9734285, LEQSF via Grant (1996–99)-RD-A-32.

1 Introduction

Modern estimation/tracking systems often involve multiple homogeneous or heterogeneous sensors that are spatially distributed to provide a large coverage, diverse viewing angles, or complementary information. If a central processor receives all measurement data from all sensors directly and processes them in real time, the corresponding processing of sensor data is known as centralized estimation. This approach has several serious drawbacks, including poor survivability and reliability, and heavy communications and computational burdens.

An alternative approach is the so-called *distributed* or *decentralized approach*. In this approach, also known as *sensor level estimation*, each sensor maintains its own estimation file based only on its own data and possibly messages received. These local estimates are transmitted to and fused in a central processor to form a fused estimate that is superior to the local estimates in some sense. In addition to better survivability and reliability and usually a lower communication load, this approach has the advantage of distributing the computational load.

This distributed approach has two major components (or steps): sensor level estimation and estimation fusion. Like most other work on distributed estimation, this paper deals only with the second component: optimal distributed estimation fusion. Specifically, a general version of optimal distributed estimation fusion in the best linear unbiased estimation (BLUE) sense is developed.

Not only does this general version include existing results on distributed estimation fusion known to the authors as special cases (for example, the two-sensor track fusion of [2, 3] and the distributed tracking by Chong et al. [8, 10, 11]), but it is also perfectly valid for many more general and realistic cases, such as those with

- coupled measurement errors across sensors;

- sophisticated network structures or communication patterns, including feedback;
- distinct local dynamic models;
- heterogeneous local estimators; and
- efficient fusion of asynchronized local estimates.

Applications of this BLUE fusion to the optimization of distributed networks in terms of reliability and survivability are also discussed.

First, we formulate the problem of distributed estimation fusion in a general setting of best linear unbiased estimation (BLUE), also known as linear unbiased least mean-square (LMS) estimation. For unbiased local estimators, the linear, unbiased fused estimator of the smallest mean-square error is their weighted sum with a matrix weight. We show that for most practical problems, the set of weights is optimal if and only if it is a solution of a matrix quadratic optimization problem subject to a linear equality constraint. This differs from the prevailing approach to estimation fusion based on the equivalence between distributed and central estimation under the linear Gaussian assumption. In other words, we approach the estimation fusion problem from a point of view that is theoretically more fundamental and convenient. This enables us to employ more powerful mathematical tools to achieve more general and fundamental results.

Then, we present a general solution of the above optimization problem. It depends only on the covariance matrix C_k of the stacked vector of all unbiased local estimates, which can be calculated off-line provided that it is known and the covariance matrix of each local estimate can be calculated off-line. The unique solution of the above optimization problem is given, along with a necessary and sufficient condition for the uniqueness. As such, both general and unique optimal BLUE fusion rules are obtained, together with a necessary and sufficient condition. We also present an explicit formula of the optimal weights for the special case in which the covariance matrix C_k is nonsingular.

We also discuss the generality and usefulness of the results obtained. For example, some potential applications are pointed out. The usefulness of the nonunique optimal fusion rules is also discussed in terms of survivability, reliability, and communication requirements.

Finally, an off-line recursion of C_k is given for a class of multisensor linear systems.

The remaining of the paper is organized as follows. In Sec. 2, we formulate the BLUE fusion as a matrix quadratic optimization problem subject to a linear-equality constraint. In Sec. 3, we present both the general solution and the unique solution of the optimization problem. Sec. 4 is dedicated to discussions of the generality and usefulness of the results obtained. An off-line recursive formula is presented in Sec. 5 for the computation of the covariance matrix of the stacked local estimators, a key quantity in BLUE fusion. Conclusions are provided in Sec. 6 and a proof that an existing fusion formula is a special case of the BLUE fusion is given in Appendix.

2 Formulation of BLUE Fusion as an Optimization Problem

Consider a distributed estimation system. Denote by $\{\mathbf{x}_k\}$ an N-dimensional state sequence of a dynamic system to be estimated, and by $\{\hat{\mathbf{x}}_k^{(i)}\}$ the corresponding sequence of local unbiased estimates of the state sequence based on all received data at the ith sensor. Assume that the following error covariance matrix is available

$$C_k = \begin{bmatrix} C_k^{(11)} & \cdots & C_k^{(1l)} \\ & \cdots & \\ C_k^{(l1)} & \cdots & C_k^{(ll)} \end{bmatrix}$$

where l is the number of local estimates to be fused and

$$C_k^{(ij)} = E[(\hat{\mathbf{x}}_k^{(i)} - \mathbf{x}_k)(\hat{\mathbf{x}}_k^{(j)} - \mathbf{x}_k)']$$

Note that $C_k^{(ij)}$ can be calculated recursively in many cases (see [2] and Sec. 5 below).

Given a set of unbiased "local estimates" $\{\hat{\mathbf{x}}_k^{(1)}, \ldots, \hat{\mathbf{x}}_k^{(l)}\}$[1], we want to find an optimal fused estimator in the BLUE sense:

$$\hat{\mathbf{x}}_k = B_k + W_k' X_k \tag{1}$$

where B_k, W_k do not involve X_k and

$$X_k = \begin{bmatrix} \hat{\mathbf{x}}_k^{(1)\prime} \\ \vdots \\ \hat{\mathbf{x}}_k^{(l)\prime} \end{bmatrix}, \quad W_k = \begin{bmatrix} W_k^{(1)\prime} \\ \vdots \\ W_k^{(l)\prime} \end{bmatrix}, W_k \in R^{(lN)\times N}$$

That is, $\hat{\mathbf{x}}_k$ has the minimum mean-square error among all choices of B_k and W_k that guarantee unbiasedness. Taking expectation of (1) yields, by the unbiasedness of the fused and local estimators,

$$E[\mathbf{x}_k] = B_k + W_k' E(X_k) = B_k + \sum_{i=1}^{l} W_k^{(i)\prime} E[\mathbf{x}_k]$$

In order for this equation to hold for every possible $E[\mathbf{x}_k]$, a necessary and sufficient condition is

$$B_k = 0, \quad \sum_{i=1}^{l} W_k^{(i)\prime} = I$$

or

$$B_k = 0, \quad AW_k = I, \quad \text{with } A = [I \cdots I]$$

[1] Although the term "local estimates" is used, this set of estimates is not necessarily obtained based on all *distinct* sensor data. This will be clear later.

Thus $\hat{\mathbf{x}}_k = W_k' X_k$ and the BLUE fusion problem becomes one of a matrix quadratic optimization problem subject to a linear equality constraint:

$$W_k = \arg\min E[(W'X_k - \mathbf{x}_k)(W'X_k - \mathbf{x}_k)'] \quad (2)$$

subject to

$$AW_k = I \quad (3)$$

for some $W_k \in R^{(lN)\times N}$. Note that $W_k^{(i)}$ is the matrix-valued weight for the estimate $\hat{\mathbf{x}}_k^{(i)}$ and the above implies that the fusion weight satisfies the linear-equality constraint $AW_k = I$.

The error covariance of the fused estimate associated with a weighting matrix W_k is

$$\begin{aligned} P_k &= E[(W_k'X_k - \mathbf{x}_k)(W_k'X_k - \mathbf{x}_k)'] \\ &= E[(W_k'X_k - W_k'A'\mathbf{x}_k)(W_k'X_k - W_k'A'\mathbf{x}_k)'] \\ &= W_k'C_kW_k \end{aligned}$$

Substituting this equation into (2) yields

$$W_k = \arg\min_{AW=I} W'C_kW \quad (4)$$

$$\min P_k = \min_{AW_k=I} W_k'C_kW_k \quad (5)$$

In other words, a necessary and sufficient condition of BLUE fusion is that the optimal weight is a solution of the above quadratic optimization problem subject to a linear-equality constraint.

It can be shown that previous results on distributed estimation fusion (see e.g., [4, 2, 3, 9, 10, 11]) satisfy the linear equality constraint $AW_k = I$. For example, the fusion equations presented in [9] are given by, using their notations,

$$P^{-1}(k|k) = P^{-1}(k|k-1) + \sum_{i=1}^{M}[P_i^{-1}(k|k) - P_i^{-1}(k|k-1)]$$

and

$$P^{-1}(k|k)\hat{\mathbf{x}}(k|k) = P^{-1}(k|k-1)\hat{\mathbf{x}}(k|k-1)$$

$$+\textstyle\sum_{i=1}^{M}[P_i^{-1}(k|k)\hat{\mathbf{x}}_i(k|k) - P_i^{-1}(k|k-1)\hat{\mathbf{x}}_i(k|k-1)]$$

That is, the weighting matrix is

$$W_k = P(k|k)\Big[P^{-1}(k|k-1), P_1^{-1}(k|k), \ldots, P_M^{-1}(k|k), -P_1^{-1}(k|k-1), \ldots, -P_M^{-1}(k|k-1)\Big]$$

which clearly satisfies $AW_k = I$ if we let $l = 2M+1$ and define

$$\begin{aligned} X_k &= [\hat{\mathbf{x}}_k^{(1)\prime} \ \cdots \ \hat{\mathbf{x}}_k^{(l)\prime}]' \\ &= \Big[\hat{\mathbf{x}}(k|k-1)', \hat{\mathbf{x}}_1(k|k)', \ldots, \hat{\mathbf{x}}_M(k|k)', \\ &\qquad \hat{\mathbf{x}}_1(k|k-1)', \ldots, \hat{\mathbf{x}}_M(k|k-1)'\Big]' \end{aligned}$$

Note that the "local estimates" do not have to be obtained from different local filters or from entirely different set of data provided they are unbiased estimates of the same quantity, not to mention the data used for these estimates could be coupled.

3 Optimal Fusion Weights

It should be recognized that the optimization problem (2)–(3) is actually a matrix linear least-squares problem subject to a linear equality constraint. A number of solution methods and algorithms are available for such problems (see e.g., [14, 6, 17, 12, 16] and the references therein). Some of them are precise and others are approximate. Some are numerically more efficient than others.

We now derive the most general version of the optimal weights given by (4) without any assumption using a method we developed recently for the general linear least-squares problem with linear constraints, presented in [17]. This method is based on the pseudoinverse technique and the perfect square method.

Theorem 1. The general solution of (4) is

$$W_k = \frac{1}{l}(I + (PC_kP)^+)A' + PZ \quad (6)$$

where $P = I - \frac{1}{l}A'A$ and $Z \in R^{(lN)\times N}$ is an arbitrary matrix satisfying $C_k^{\frac{1}{2}\prime}PZ = 0$.

Proof. It is well-known (see [5]) that

$$W = A^+ + P\xi, \qquad \forall \xi \in R^{(lN)\times N} \quad (7)$$

is the general solution of the following matrix equation

$$AW = I$$

It is straightforward to show from the basic properties of the matrix pseudoinverse directly [5] that

$$C_k^{\frac{1}{2}\prime}P(PC_k^{\frac{1}{2}}C_k^{\frac{1}{2}\prime}P)^+PC_k^{\frac{1}{2}}C_k^{\frac{1}{2}\prime}P$$

$$= C_k^{\frac{1}{2}\prime}P(C_k^{\frac{1}{2}\prime}P)^+(PC_k^{\frac{1}{2}})^+PC_k^{\frac{1}{2}}C_k^{\frac{1}{2}\prime}P = C_k^{\frac{1}{2}\prime}P$$

Substituting the last three equations above into the right side of (4), we have

$$[(A^{+\prime} + \xi'P)C_k^{\frac{1}{2}}][(A^{+\prime} + \xi'P)C_k^{\frac{1}{2}}]'$$

$$= [\xi' - A^{+\prime}C_k^{\frac{1}{2}}C_k^{\frac{1}{2}\prime}P(PC_k^{\frac{1}{2}}C_k^{\frac{1}{2}\prime}P)^+]PC_k^{\frac{1}{2}}C_k^{\frac{1}{2}\prime}P$$

$$[\xi' - A^{+\prime}C_k^{\frac{1}{2}}C_k^{\frac{1}{2}\prime}P(PC_k^{\frac{1}{2}}C_k^{\frac{1}{2}\prime}P)^+]'$$

$$+A^{+\prime}C_k^{\frac{1}{2}}C_k^{\frac{1}{2}\prime}A^+$$

$$-A^{+\prime}C_k^{\frac{1}{2}}C_k^{\frac{1}{2}\prime}P(PC_k^{\frac{1}{2}}C_k^{\frac{1}{2}\prime}P)^+PC_k^{\frac{1}{2}}C_k^{\frac{1}{2}\prime}A^+$$

Clearly, minimizing the quadratic objective function amounts to making

$$[\xi' - A^{+'}C_k^{\frac{1}{2}}C_k^{\frac{1}{2}'}P(PC_k^{\frac{1}{2}}C_k^{\frac{1}{2}'}P)^+]PC_k^{\frac{1}{2}}C_k^{\frac{1}{2}'}P$$
$$\cdot[\xi' - A^{+'}C_k^{\frac{1}{2}}C_k^{\frac{1}{2}'}P(PC_k^{\frac{1}{2}}C_k^{\frac{1}{2}'}P)^+]' = 0.$$

Using the pseudoinverse technique (see e.g., [17]), it can be shown that

$$P(PC_kP)^+ = (PC_kP)^+P = (PC_kP)^+$$

Hence, we should take

$$\xi = (PC_kP)^+C_kA^+ + Z,\ \forall Z : C_k^{\frac{1}{2}'}PZ = 0 \quad (8)$$

The theorem thus follows from (7), (8), and $A^+ = \frac{1}{l}A'$. ■

Now we present the unique optimal weight W_k, along with a necessary and sufficient condition for it to hold.

Theorem 2. The optimal weight [i.e., the solution of (4)] is given uniquely by

$$W_k = \frac{1}{l}(I + (PC_kP)^+C_k)A'$$
$$= \frac{1}{n}[I + V(V'C_kV)^{-1}V'C_k]A' \quad (9)$$

if and only if $\begin{pmatrix} A \\ C_k^{\frac{1}{2}'} \end{pmatrix}$ has column full rank Nl, where V is a full-rank square-root matrix of P: $VV' = P$.

Proof. Note first that only W_k of (6) with $PZ = 0$ can be a unique solution, otherwise W_k of (6) with αPZ for any real number $\alpha \neq 0$ would be a distinct solution since $\alpha C_k^{\frac{1}{2}'}PZ = 0$. When $PZ = 0$, (6) gives a unique solution due to the uniqueness of the M-P pseudoinverse. Thus the constrained optimization problem has a unique solution iff $PZ = 0$. A necessary and sufficient condition for $PZ = 0$ when $C_k^{\frac{1}{2}'}PZ = 0$ is that the vector PZ is in the row space of $C_k^{\frac{1}{2}'}$. Since $P = I - \frac{1}{l}A'A$ is a projector onto the orthogonal complement of the row space of A, the above necessary and sufficient condition holds iff the row space of $C_k^{\frac{1}{2}'}$ is the orthogonal complement of the row space (i.e., subspace spanned by the row vectors) of A, which is equivalent to $\begin{pmatrix} A \\ C_k^{\frac{1}{2}'} \end{pmatrix}$ having full column rank Nl. The second equation in (9) follows from the first one because it can be shown (see [17]) that

$$(PC_kP)^+ = V(V'C_kV)^{-1}V'$$

■

When W_k is unique, from the definition of A and Theorem 2, we have a clear expression for each of its elements $W_k^{(i)}$ immediately. Denote by $M_{ij}^{(+)}$ the (i,j)th sub-block $(N \times N)$ matrix of $\frac{1}{l}(I + (PC_kP)^+C_k) = \frac{1}{n}[I + V(V'C_kV)^{-1}V'C_k]$. Then

$$W_k^{(i)} = \sum_{j=1}^{l} M_{ij}^{(+)} \quad (10)$$

Moreover, if C_k has full rank, we have a more explicit expression of $W_k^{(i)}$ that depends only on C_k^{-1}.

Theorem 3. If C_k has full rank, we have the following explicit expression of each element $W_k^{(r)}$:

$$W_k^{(r)} = \sum_{j=1}^{l} C_{rj}^{(-1)} \Big(\sum_{i,j=1}^{l} C_{ij}^{(-1)} \Big)^{-1}$$

where $C_{ij}^{(-1)}$ is the (i,j)th submatrix of C_k^{-1}.

Proof. First of all, we know $AC_k^{-1}A'$ is nonsingular since A has full row rank. Then we can easily verify the following two identities

$$\arg \min_{AW=I} W'C_kW$$
$$\overset{(4)}{=} \arg \min_{AW=I} [(W_k - W)'C_k(W_k - W) + W_k'C_kW_k]$$
$$= \arg \min_{AW=I} [(W_k - W)'C_k(W_k - W) + (AC_k^{-1}A')^{-1}]$$

where

$$W_k = C_k^{-1}A'(AC_k^{-1}A')^{-1} \quad (11)$$

and

$$AW_k = AC_k^{-1}A'(AC_k^{-1}A')^{-1} = I$$

Finally, the theorem follows from the product of the two block matrices in (11). ■

When $l = 2$, using the well-known inverse of the (nonsingular) partitioned matrix, as shown in the Appendix, the two-sensor track fusion formula presented in [3] is a special case of Theorem 3.

The above fused estimators is optimal only if $E[\mathbf{x}_k]$ is not known. If it is known and

$$C_X = E[(X_k - EX_k)(X_k - EX_k)']$$
$$C_{X\mathbf{x}} = E[(X_k - EX_k)(\mathbf{x}_k - E\mathbf{x}_k)']$$

are also known, using the well-known linear unbiased LMS estimation result (by treating X_k as the measurement in the standard formulation) (see, e.g., [7]), the BLUE fused estimator is given by

$$\hat{\mathbf{x}}_k = (I - W_k'A')E[\mathbf{x}_k] + W_k'X_k$$

where the optimal weighting matrix is the unique solution of (2) without the linear equality constraint $AW_k = I$, given by

$$W_k = C_X^+C_{X\mathbf{x}}$$

The covariance associated with the fused estimator is given by

$$P_k = \text{cov}(\mathbf{x}_k) - W_k' C_X W_k$$

Note that in most practical situations, unfortunately, either $E[\mathbf{x}_k]$, $C_{X\mathbf{x}}$, or C_X is not known.

4 Discussions

In this section, we discuss mainly the generality and usefulness of the general BLUE fusion results obtained above.

First, it should be emphasized that the above BLUE fusion results rely on two assumptions: (1) the local estimators are unbiased and (2) the covariance matrix C_k is known. Both assumptions are fairly reasonable. Theorems 1 through 3 also assume that the unconditional expectation of the state is not known.

It is hard to image how we can have an unbiased estimator by fusing biased local estimators unless the biases are known perfectly. If, however, the biases are known, then we can obtain unbiased local estimators in the first place. Nevertheless, the above unbiasedness assumption can still be relaxed to some degree. For example, if $E[\mathbf{x}_k]$ and $E[\hat{x}_k^{(i)}]$ are known, then the constraint $AW_k = I$ should be replaced by $W_k' E[X_k] = E[\mathbf{x}_k]$ (setting $B_k = 0$). Exactly the same approach can be followed to yield the BLUE fusion rule except that the final result is somewhat more complicated. This result could be theoretically superior to the one by debiasing each local estimator.

If the measurement noises are independent across sensors, the covariance C_k is not (block) diagonal only because the same *random* state is estimated by all local estimators, for example, because of the common process noise in the system dynamics on which each local estimator is based. C_k in this case and in some more general and coupled cases can be easily obtained (see Sec. 5 below). In general, C_k quantifies the coupling among local estimators. It, or its equivalent, is needed for optimal fusion. This availability assumption for C_k basically guarantees that our BLUE fusion results are optimal given the coupling among local estimators. When this coupling is neither known nor obtainable, the above fusion results are not applicable directly but may facilitate the development of the corresponding optimal fusion.

Note that in the above BLUE fusion, $\hat{x}_k^{(i)}$ could be any unbiased estimate of x_k. For example, one could have

$$X_k = [\hat{x}_{k|k-1}^{(1)\prime}, \hat{x}_{k|k}^{(1)\prime}, \hat{x}_{k|k-2}^{(2)\prime}, \hat{x}_{k|k}^{(2)\prime}, \ldots, \hat{x}_{k|k-1}^{(m)\prime}]'$$

where the superscript j denotes the estimate by the jth "local" estimator. With this understanding, most advantages of the above BLUE fusion results presented below can be easily appreciated.

Comparing with existing results on estimation fusion, the above BLUE fusion formulas have at least the following advantages:

- They are valid for cases with coupled observation noises across sensors. This is useful in practice when the dynamic process is observed in a common noisy environment, such as when a target is taking an electronic countermeasure (ECM), e.g. noise jamming, or when the sensor noises are coupled because of, say, their dependence on the target state. A class of systems that fall into this category is given in Sec. 5. Another important application area is the fusion of estimates based on observations obtained over different time periods. The fusion-based optimal smoothing using measurements corrupted by autocorrelated noise is a good candidate for application of our new results. Almost all previous fusion results assume that the sensor observations are conditionally independent given the target state/signal to be estimated. A formula was mentioned in [15] that is valid for fusing local estimates based on not necessarily disjoint observation data, which is a limited yet useful form of data dependence, still a special case.
- The fused estimator depends on the network structures or communication patterns only through C_k. Consequently, the fusion rule is invariant no matter if there is feedback or not, if the network has a parallel, tandem, tree or general structure, or what the communication bandwidths are. This means that in practice all we need to obtain is the coupling between every pair of local estimates. If there is no fusion center, such as for the network structure considered in [13], each sensor can use our BLUE fusion formulas to obtain the best estimator based on its own observations and any information received from other sensors.
- The local estimators does not have to use the same dynamic system model. The local estimates may be obtained based on different dynamic models and then be fused using our BLUE fusion formulas. The use of different dynamic models for the local estimators may be necessary or more effective. It is necessary, for example, when state augmentation is needed for some local estimators with autocorrelated sensor measurement noise. It may be more effective, for example, when multiple models are used for the process to be estimated. Of course the key to a successful application of our results in such cases lies in the determination of the covariance matrix C_k. In order to have the best fusion performance, the choice of the local models is of course not arbitrary

(see, e.g., [1]).

- There is no requirement on synchronism of the local estimates provided they can be converted to the estimates of the state at the same time (based on observations up to different times). In reality, it is difficult to synchronize local estimates. However, many well-known results are valid only for synchronized local estimates and thus require artificial synchronization for the local estimates. This either increases the computational burden considerably or degrade the fusion performance. Our BLUE fusion formulas provide a more convenient and efficient framework for fusing asynchronized local estimates. For example, our BLUE fusion formulas are applicable (optimally) after we simply convert all local estimates $\hat{\mathbf{x}}^{(i)}(t_i|t_j)$ to $\hat{\mathbf{x}}^{(i)}(k|t_j)$, that is, the same point k in time at which the state is to be estimated by the fusion center, where the conversion could either be prediction if $k > t_i$ or smoothing (retrodiction) if $k < t_i$.
- The local estimators need not be of the same type. For example, our results are valid if some local estimators are MMSE (minimum mean-square error) estimators while others are MAP (maximum a posteriori) estimators, provided they are unbiased. This flexibility is useful for some applications.

The authors are not aware of any existing fusion results that are valid for such cases.

It is quite possible in practice for C_k to be singular (or more precisely, for $\begin{pmatrix} A \\ C_k^{\frac{1}{2}\prime} \end{pmatrix}$ not to have full rank). Intuitively, this may be the case if there is no independent parts between any two sensor observation noises (see the dynamic system with $w_k^{(i)} \equiv 0$ given in Sec. 5). The optimal W_k is not unique in this case. The general solution (6) is particularly useful in this case. A special solution (i.e., a special set of weights) in this case may be chosen based on some other considerations, such as survivability, reliability, and communication requirements. For instance, we may choose the solution from the set of optimal solutions that is the best among all cases where a given number of sensors are lost. In some cases, we may be able to choose a solution in which some of the optimal weights vanish, which implies that the performance of the fused estimator will not deteriorate without the corresponding local estimates. In other cases, we may want to put these local sensors/estimators in a "stand-by" mode since their removal incurs no degradation of system performance.

Theorem 2 is computationally more efficient than Theorem 3 mainly because the matrix $V'C_kV$ to be inverted has a lower dimension than C_k. However, Theorem 3 is in a form that has a greater resemblance to the existing fusion results.

5 Recursive Computation of Covariance C_k

It can be easily seen that the optimal weighting matrix W_k, given by Theorems 1, 2, and 3, depends only on the covariance matrix C_k and the computational burden of W_k relies mostly on the computation of C_k (and its inverse). In many practical situations, C_k may depend only on the system coefficient matrices and known noise covariances. Hence, C_k and thus W_k can be calculated off-line. An off-line recursion for $C_k^{(i,j)}$ is presented in [2] assuming that the measurement noises are independent across sensors. In this section, we extend that result to a class of linear systems having dependent measurement noises with known correlations between any two sensors.

Consider a linear dynamic process

$$\mathbf{x}_{k+1} = F_k\mathbf{x}_k + v_k$$

with additive zero-mean white noise

$$E[v_k] = 0, E[v_k v_j'] = Q_k\delta_{kj}$$

and noisy measurement

$$y_k^{(i)} = H_k^{(i)}\mathbf{x}_k + w_k^{(i)} + e_k^{(i)}, \ \forall i \le l$$

where the measurement noise is the sum of two zero-mean white noises $w_k^{(i)}$ and $e_k^{(i)}$, uncorrelated with the process noise:

$$\begin{aligned} &E[w_k^{(i)}] = 0, \quad && E[w_k^{(i)}w_j^{(i)\prime}] = \sigma_k^{(i)}\delta_{kj} \\ &E[e_k^{(i)}] = 0, \quad && E[e_k^{(i)}e_j^{(i)\prime}] = S_k^{(ii)}\delta_{kj} \\ &E[v_k w_j^{(i)\prime}] = 0, \quad && E[v_k e_j^{(i)\prime}] = 0 \end{aligned}$$

However, while $w_k^{(i)}$'s are independent across sensors, $e_k^{(i)}$'s are coupled across sensors:

$$E[w_k^{(i)}w_k^{(j)\prime}] = 0, \qquad E[e_k^{(i)}e_k^{(j)\prime}] = S_k^{(ij)}$$

Clearly, this system reduces to the one with independent measurement noise when $e_k^{(i)} \equiv 0$. As explained before, this model may be useful e.g., when a target is generating noise jamming or when the sensor noises are dependent on the target state.

Similar to the derivation in [2], it can be shown using Kalman filter formulas for the above system that we have the following recursive formulas, for $k = 1$ and assuming $S_0 = 0$,

$$\begin{aligned} C_1^{(ij)} = &(I - K_1^{(i)}H_1^{(i)})Q_0(I - K_1^{(j)}H_1^{(j)})' \\ &+K_1^{(i)}S_1^{(ij)}K_1^{(j)} \\ &+K_1^{(i)}\sigma_1^{(i)}\delta_{ij}K_1^{(j)\prime} \quad i,j = 1,...,l \end{aligned} \tag{12}$$

and for any $k > 1$,

$$C_k^{(ij)} = (I - K_k^{(i)} H_k^{(i)}) Q_{k-1} (I - K_k^{(j)} H_k^{(j)})' + (I - K_k^{(i)} H_k^{(i)}) F_{k-1} S_{k-1}^{(ij)} F_{k-1\prime} (I - K_k^{(j)} H_k^{(j)})' + K_k^{(i)} S_k^{(ij)} K_k^{(j)} + K_k^{(i)} \sigma_k^{(i)} \delta_{ij} K_k^{(j)\prime}, \quad i, j = 1, ..., l \tag{13}$$

where K is the Kalman filter gain.

Let

$$\lambda_k^{(i)} = (I - K_k^{(i)} H_k^{(i)}) F_{k-1}$$
$$\Lambda_k = \text{diag}\{\lambda_k^{(1)}, ..., \lambda_k^{(l)}\}$$
$$K(k) = \text{diag}\{K_k^{(1)}, ..., K_k^{(l)}\}$$
$$\Sigma_k = \text{diag}\{\sigma_k^{(1)}, ..., \sigma_k^{(l)}\}$$
$$S_k = \begin{bmatrix} S_k^{(11)} \cdots S_k^{(1l)} \\ \cdots \\ S_k^{(l1)} \cdots S_k^{(ll)} \end{bmatrix}$$

$$M_k = [Q_{k-1}^{\frac{1}{2}\prime} (I - K_k^{(1)} H_k^{(i)})' \cdots Q_{k-1}^{\frac{1}{2}\prime} (I - K_k^{(l)} H_k^{(l)})']'.$$

Then, an off-line recursion of C_k is obtained by rewriting (12)–(13) in the matrix form as

$$C_k = \Lambda_k C_{k-1} \Lambda_k' + K(k)(\Sigma_k + S_k) K(k)' + M_k M_k'$$

which can be initialized by

$$C_1 = M_1 M_1' + K(1)(\Sigma_1 + S_1) K(1)'$$

6 Conclusions

A general version of best linear unbiased estimation (BLUE) fusion has been developed that has the least mean-square estimation error among all linear unbiased fusion rules. This BLUE fusion has been formulated as a matrix quadratic optimization problem subject to a linear equality constraint. Both the most general solution and the unique solution of this optimization problem, along with a necessary and sufficient condition for the uniqueness, have been presented. The fusion rule depends only on the grand covariance matrix of a stacked vector of all local estimators. The generality and usefulness of the fusion formulas developed have been discussed, with an emphasis on cases with coupled measurement noises among sensors, sophisticated network structures, different local dynamic models, and asynchronized local estimates. Examples have been given, which demonstrate that this BLUE fusion rule includes existing fusion results as special cases. An off-line recursion of the grand covariance matrix has been presented for a class of multisensor linear systems with coupled measurement noises.

Appendix

Corollary 1. For the two-sensor case, if C_k and $C_k^{(11)}$ both have full rank, then $W_k^{(i)}$ in Theorem 3 reduces to the same form as the one given in [3].

Proof. Using the assumption on C_k and $C_k^{(11)}$ and the well-known inverse of the partitioned matrix, we have, denoting $C_k^{-(ij)} = (C_k^{(ij)})^{-1}$,

$$C_{11}^{(-1)} = (C_k^{(11)} - C_k^{(12)} C_k^{-(22)} C_k^{(21)})^{-1}$$
$$C_{12}^{(-1)} = -[C_k^{(11)} - C_k^{(12)} C_k^{-(22)} C_k^{(21)})^{-1} C_k^{(12)} C_k^{-(22)}$$
$$C_{21}^{(-1)} = -C_k^{-(22)} C_k^{(21)} (C_k^{(11)} - C_k^{(12)} C_k^{-(22)} C_k^{(21)})^{-1}$$
$$C_{22}^{(-1)} = C_k^{-(22)} + C_k^{-(22)} C_k^{(21)} \cdot (C_k^{(11)} - C_k^{(12)} C_k^{-(22)} C_k^{(21)})^{-1} C_k^{(12)} C_k^{-(22)}$$

Therefore,

$$\sum_{j=1}^{2} C_{1j}^{(-1)} = (C_k^{(11)} - C_k^{(12)} C_k^{-(22)} C_k^{(21)})^{-1} (I - C_k^{(12)} C_k^{-(22)}) \tag{14}$$

$$\sum_{i,j=1}^{2} C_{ij}^{(-1)} = C_k^{-(22)} + (I - C_k^{-(22)} C_k^{(21)})(C_k^{(11)} - C_k^{(12)} C_k^{-(22)} C_k^{(21)})^{-1} \cdot (I - C_k^{(12)} C_k^{-(22)}) \tag{15}$$

Using well-known identity of matrix inverse and the following equation

$$C_k^{(11)} + C_k^{(22)} - C_k^{(12)} - C_k^{(21)} = (C_k^{(22)} - C_k^{(12)}) C_k^{-(22)} (C_k^{(22)} - C_k^{(21)}) + C_k^{(11)} - C_k^{(12)} C_k^{-(22)} C_k^{(21)} \tag{16}$$

we have

$$(\sum_{i,j=1}^{2} C_{ij}^{(-1)})^{-1} = C_k^{(22)} - (C_k^{(22)} - C_k^{(21)}) \cdot (C_k^{(11)} + C_k^{(22)} - C_k^{(12)} - C_k^{(21)})^{-1} (C_k^{(22)} - C_k^{(12)})$$

Using Theorem 3, (14), and (16),

$$W_k^{(1)} = \sum_{j=1}^{2} C_{1j}^{(-1)} (\sum_{i,j=1}^{2} C_{ij}^{(-1)})^{-1} = (C_k^{(11)} - C_k^{(12)} C_k^{-(22)} C_k^{(21)})^{-1} (C_k^{(22)} - C_k^{(12)}) - (C_k^{(11)} - C_k^{(12)} C_k^{-(22)} C_k^{(21)})^{-1} (I - C_k^{(12)} C_k^{-(22)}) \cdot (C_k^{(22)} - C_k^{(21)}) (C_k^{(11)} + C_k^{(22)} - C_k^{(12)} - C_k^{(21)})^{-1} \cdot (C_k^{(22)} - C_k^{(12)}) \tag{17}$$

The first term on the right side of the above equation can be rewritten as

$$(C_k^{(11)} - C_k^{(12)} C_k^{-(22)} C_k^{(21)})^{-1} \cdot (C_k^{(11)} + C_k^{(22)} - C_k^{(12)} - C_k^{(21)}) \cdot (C_k^{(11)} + C_k^{(22)} - C_k^{(12)} - C_k^{(21)})^{-1} (C_k^{(22)} - C_k^{(12)}) \tag{18}$$

It follows from (16), (17), and (18) that

$$W_k^{(1)} = (C_k^{(11)} + C_k^{(22)} - C_k^{(12)} - C_k^{(21)})^{-1}(C_k^{(22)} - C_k^{(12)})$$

Noticing $W_k^{(1)} + W_k^{(2)} = I$ or using the argument similar to the above, we have

$$W_k^{(2)} = (C_k^{(11)} + C_k^{(22)} - C_k^{(12)} - C_k^{(21)})^{-1}(C_k^{(11)} - C_k^{(21)})$$

■

As a matter of fact, the fusion formula given in [3] must be a special case of the general BLUE fusion formulas of this paper, in particular Theorems 2 and 3. The reasons are that the fusion rules are unique in this case and that they rely on the same assumptions: (1) the local estimators are unbiased, (2) C_k is known, and (3) the mean of the state is unknown. The above proof just shows this explicitly.

References

[1] A. T. Alouani and J. D. Birdwell, "Distributed Estimation: Constraints on the choice of the local models," *IEEE Trans. Automatic Control*, vol. AC-33, pp. 503–506, May 1988.

[2] Y. Bar-Shalom, "On the Track-to-Track Correlation Problem," *IEEE Trans. Automatic Control*, vol. AC-26, pp. 571–572, Apr. 1981.

[3] Y. Bar-Shalom and L. Campo, "The Effect of the Common Process Noise on the Two-Sensor Fused-Track Covariance," *IEEE Trans. Aerospace and Electronic Systems*, vol. AES-22, pp. 803–805, Nov. 1986.

[4] Y. Bar-Shalom and X. R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: YBS Publishing, 1995.

[5] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*. New York: John Wiley & Sons, 1974.

[6] A. Bjorck, *Numerical Methods for Least Squares Problems*. Philadelphia, PA: SIAM, 1996.

[7] H. F. Chen, *Recursive Estimation and Control for Stochastic Systems*. New York: Wiley, 1985.

[8] C. Y. Chong, "Hierarchical Estimation," in *Proc. Second MIT/ONR Workshop on C3*, (Monterey, CA), July 1979.

[9] C. Y. Chong, K. C. Chang, and S. Mori, "Distributed Tracking in Distributed Sensor Networks," in *Proc. 1986 American Control Conf.*, (Seattle, WA), June 1986.

[10] C. Y. Chong, K. C. Chang, and S. Mori, "Tracking Multiple Targets with Distributed Acoustic Sensors," in *Proc. 1987 American Control Conf.*, (Minneapolis, MN), June 1987.

[11] C. Y. Chong, S. Mori, and K. C. Chang, "Distributed Multitarget Multisensor Tracking," in *Multitarget-Multisensor Tracking: Advanced Applications*, (Y. Bar-Shalom, ed.), pp. 247–295, Norwood, MA: Artech House, 1990.

[12] R. Farebrother, *Linear Least Squares Computations*. New York: Marcel Dekker, 1988.

[13] S. Grime and H. F. Durrant-Whyte, "Data Fusion in Decentralized Sensor Networks," *Control Engineering Practice*, vol. 2, no. 5, pp. 849–863, 1994.

[14] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*. Philadelphia, PA: SIAM, 1995.

[15] M. Liggins, C. Y. Chong, I. Kadar, M. G. Alford, V. Vannicola, and S. Thomopoulos, "Distributed Fusion Architectures and Algorithms for Target Tracking," *Proceedings of the IEEE*, vol. 85, pp. 95–107, Jan. 1997.

[16] D. A. Wismer and R. Chattergy, *Introduction to Nonlinear Optimization*. New York: Elsevier North-Holland, 1978.

[17] Y. M. Zhu and X. R. Li, "Recursive Least Squares with Linear Constraints," *IEEE Trans. Information Theory (submitted)*, 1999.

MOTION DETECTION AND TRACKING FOR HUMAN ACTIVITY MONITORING

Cina MOTAMED

Université du Littoral Côte d'Opale
Laboratoire LASL :
195 av P.L.King 62228 CALAIS FRANCE
tel : 03 21 46 36 91
motamed@lasl-gw.univ-littoral.fr

Abstract

A video-surveillance system for human activity monitoring based an a distributed architecture is presented.
The first stage concerns the motion detection process. It can tolerate low change of illumination and can be tuned with respect to the application context. The second stage is the object tracking process containing a local and a distributed level. The local approach is the motion analysis in the field of each sensor improved with a belief revision approach. The distributed level of the tracking permits to operate over a wide area.
The third stage is the global interpretation which for our application represents the recognition of specific activity.

Keywords : Video-surveillance, Motion detection, tracking, Distributed sensor.

I. Introduction

An automatic Video surveillance generally contains three main hierarchic stages. The motion detection, the tracking unit and finally the high level motion interpretation. The main objective is to observe, recognize activity or detect incident.
The video-surveillance is one of the tasks with a high degree of dependence to the context. In fact we must help the analysis by pointing out what is really important. In many applications a top-down approach has to be considered by modeling what we want to observe.
In general, the contextual knowledge is naturally integrated into over computer vision application. Since a few years the notion of context is formalized over the High level computer vision community and AI related area.

Our application field concerns human activity monitoring. It concerns the human detection and tracking in order to recognize specific behaviors [1][2][3].

A specificity of our system is that we are interested in the monitoring of wide areas. This constraint means that numerous sensors have to be distributed in space and have to cooperate in order to obtain a global interpretation (Fig.1) [4][5].

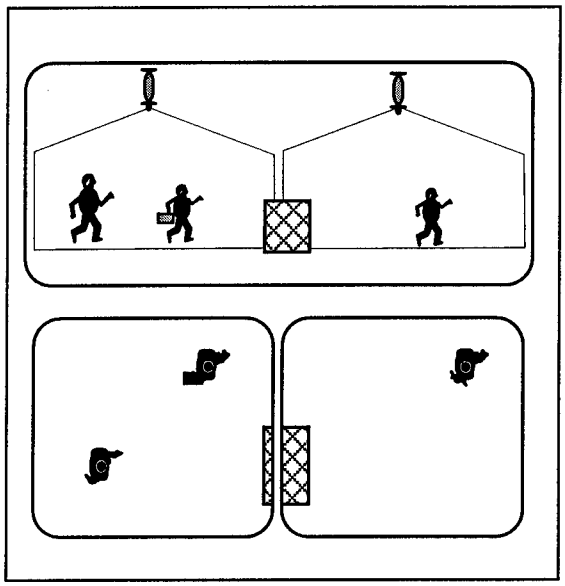


Figure 1: Distributed surveillance

In such a system, the second stage, representing the object tracking process, contains a local and a multi-sensor level. At the first level Each sensor has to interpret its own field of view. At the multi-sensor level, sensors have to help each other to match their observations from a sensor to another. These two levels are called respectively the local and the distributed tracking. The last stage concerns the global interpretation of the scene using both the local

and distributed information. At this stage, numeric and symbolic information have to be manipulated.

The paper from this point is divided into three parts :
In the first part, we present the motion detection process. It is based on a difference between the current image and a reference image. This reference image represents the background of the scene without moving objects.
In the second part we explain our approach concerning the tracking process at the local and distributed level.
The third part is focused on the development of a global interpretation for a video-surveillance purpose. It concerns the real time surveillance of humans behavior by two cameras.

II. Motion Detection Process

Cameras are fixed to the infrastructure or set into a moving platform. For fixed cameras the common motion detection approach is to model stationary background. In this case the moving object representing the foreground can be easily extracted by a simple difference from the background. For a moving sensor the motion detection has to operate independently from the flow induced by it self.
We have focused our work toward fixed camera. For our system each sensor control its own detection process. The motion detection as any low level processing stage, affects directly upper level of the interpretation. It forces us to take into account all external parameters hindering its functioning. Working outdoors or with ambient lightning, the motion detection becomes particularly sensitive. The principal difficulty is the variation of illumination inducing false detection. When these variations are slow with respect to object motion, a continuous updating of the background can solve most of ambiguities.

$$D^k(P) = \begin{cases} 1 \leftarrow if \left| R^k(P) - I^k(P) \right| > Td \\ 0 \leftarrow else \end{cases} \quad (1)$$

$D^k(P)$ is the detection image that highlight moving regions at instant k. $R^k(P)$ and $I^k(P)$ represent respectively the background and the current image at sequence k. Td is the detection threshold adjusted by an operator.

Background updating

The reference image R is constructed by updating a model at each new image acquisition I. We use an recursive filter which limits image storage. This updating is computed as :

$$R^{k+1}(P) = a.R^k(P) + (1-a)I^k(P) \quad (2)$$

Classical updating consider that this task have to be uniform in the whole image and have to take into account illumination variability only.

However there are also cases where objects come into the scene and remain. Initially objects belong to the foreground, but over time we might want to include them as part of the background. This operation is called background integration. The opposite procedure is also useful, a car leaving its park area for example. We consider that an important parameter of an integration process is time. This integration time depends on the context. In our application we want a specific and constant time value at different locations of the scene Ti(P) defined by an analyst.

The integration procedure can be made either thanks to a high level interpretation or directly inserted into low level background updating. Our work focus on the second category. The dynamic adjustment of the constant 'a' (represented by $a^k(P)$) permits to control efficiently objects integration, it takes the following form (for each pixel P) :

$$R^{k+1} = a^k.R^k + (1-a^k).I^k \quad (3)$$

$$a^{k+1} = \delta^k.V_1.a^k + (1-\delta^k).((1-V_2) + V_2.a^k) \quad (4)$$

$a^k(P)$ is a coefficient that takes values for each pixel in the range [0.. 1]. It is directly linked to a measure of the temporal stability of the reference image at the location of pixel P. A high value of $a^k(P)$ indicates that the pixel P of the reference image is reliable and effectively present in the stationary background. The $a^k(P)$ coefficient that we also called background quality indicator, depends on the value $\delta^k(P)$ which has been used to compute the 'minimal'

motion detection image. It takes the following form :

$$\delta^k(P) = \begin{cases} 1 \leftarrow if \left| R^k(P) - I^k(P) \right| > T \\ 0 \leftarrow else \end{cases} \quad (5)$$

This image represents the minimal change that can detect the system (T < Td).

The values of V_1 and V_2 have to be tuned in order to control temporally the $a^k(P)$ coefficient. V_1 depends on the integration time Ti. The V_2 parameter operates at the the increasing stage of $a^k(P)$ when the object is considered to be integrated to the background with an ε relative error ($V_2 << V_1$). For details see [6].

$$V_1 = e^{\frac{2*\ln(\varepsilon)}{T_i*(T_i+1)}} \quad (6)$$

ε : % of relative error of the detector

Finally a sequence of erosion , dilatation and erosion removes any fracture in foreground image. Figure 2 illustrates a motion detection result for a real sequence at 2 fps (Frame/second).

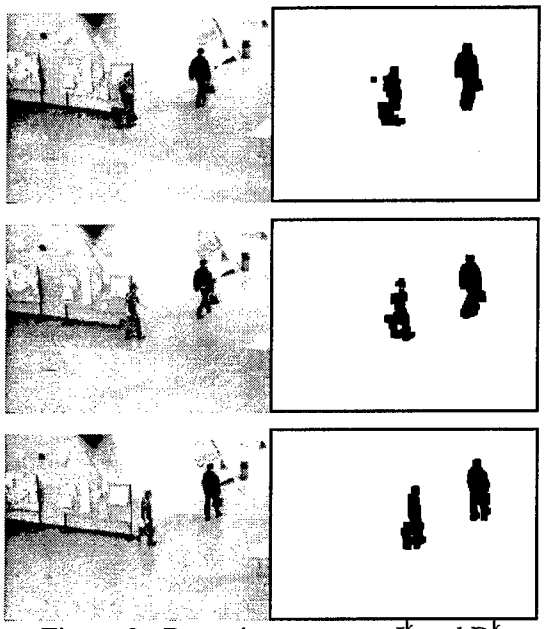


Figure 2 : Detection process : I^k and D^k

III. Tracking

The object tracking can be achieved by a wide variety of techniques. Two principal categories of constraints are used. The first concerns the object rigidity and the second depends on how the object movement can be modeled. For non rigid object tracking, dynamic shape models of objects are used. Their use imply favorable conditions of acquisition i.e. high resolution and low temporal variability. The efficiency of the motion model permits to help the tracking process by improving the prediction.

Human tracking in a large area induces several difficulties when:

- objets are non rigid
- no accurate motion model is available
- for small sized objects no enough efficient visual information is available
- the field of view is large and several distributed sensors have to be associated to obtained the global path.

Our approach as for any visual surveillance devices, takes into account contextual information in order to help the tracking process [7][8]. The contextual information concerns static and behavioral knowledge associated with the observed scene.

Two main hypotheses are used, the first concerns the size conservation of each object, and the second assumes that objects have behavioral limitation (speed, acceleration, direction variability...). We consider that a human has a typical calibrated dimension (Size_min to Size_max). These measurements can be obtained by learning from the observed scene. An isolated detected region with this typical size is assumed to be a human being. Larger size can be interpreted as the fusion of several people. For illustration an histogram of region size during 10 minutes is presented (Fig 3).

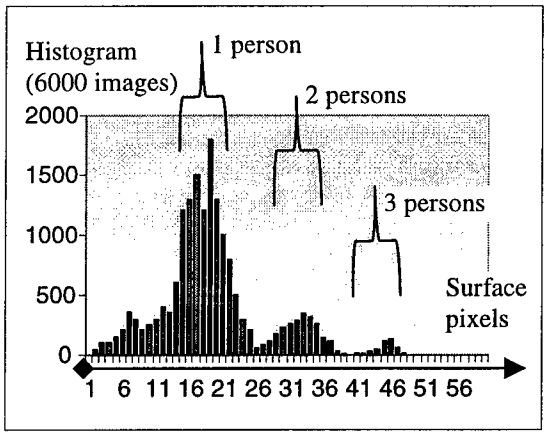


Fig. 3. Detected surface histogram during a sequence of 6000 images by tracking of 95 persons from 10 minutes at 10 fps.

The aims of the local tracking process is to match a detected region from a temporal sequence to another and to take into account the splitting and merging phenomenon during objects motion. The distributed tracking process has to match a region from a sensor to another, using local tracking information.

Local tracking algorithm

The local tracking process is based on a first order prediction of region displacement and an overlapping degree Od between the prediction and a current region.

$$Od\ (a,b,k) = \frac{O\_Area\ (Sp_a^k, S_b^k)}{Area\ (S_a^k)} \quad (7)$$

S^k_i represents region i at sequence k. Sp^k_i is the prediction of S^k_i. The O_Area and Area functions represent respectively the common (overlapping) area and the area of regions in pixels, see illustration Fig 4.

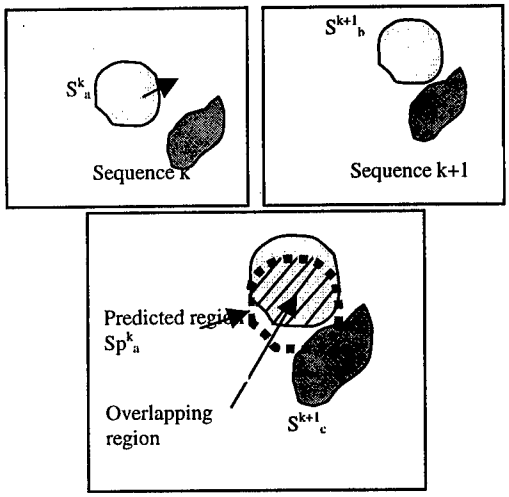


Figure 4 Prediction and overlapping area.

Another component of local tracking is the integration of a belief revision mechanism associated with each detected object, reflecting an instantaneous quality of its tracking.

We consider that the belief can increase when
- The tracked object follows a continuous path
- the object size is stable
- no ambiguity appears (other target meeting : target splitting, lost of target ..)

Otherwise the belief decreases. Two examples of belief revision is presented in Fig 5and 6.

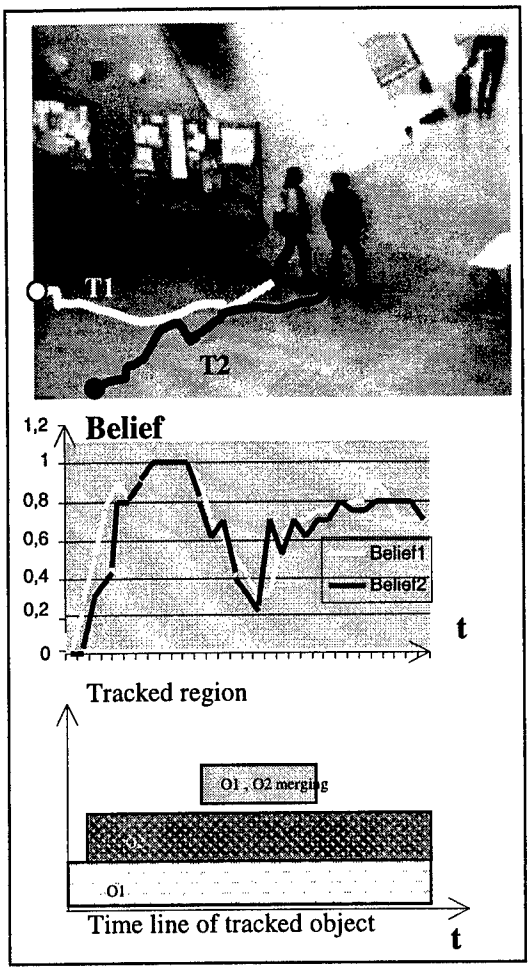


Figure 5. Belief revision in the presence of merging region

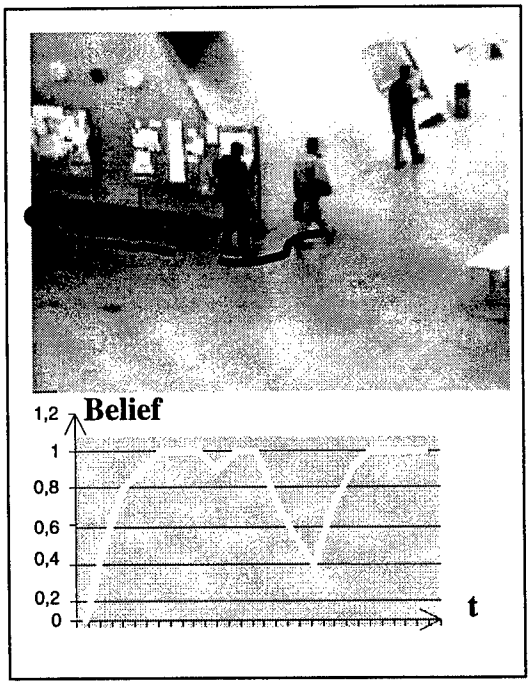


Fig 6. Belief revision while loss of region

The first case concerns the behavior of the belief in the presence of merging situation (Fig. 5). The second illustrates loss of target during 3 frames in Fig. 6.

When working in real time (near 10fps for example) Generally, intersection between two consecutive regions exists, and so Od is strictly positive. When a target is lost for any reason the best region verifying constraint (size and proximity) is chosen. It may occur in some situations (occlusion or no detection ..) that no candidate region can be found. In this case the prediction is followed while tracking belief is positive. A null value of belief induces termination of the track.

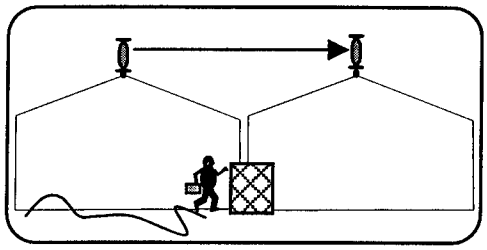
Distributed tracking

We are interested in the interpretation of wide areas, this constraint means that numerous sensors have to be distributed in space. With the aim of tracking objects over a wide scene, it is necessary to recognize each object when it appears in the field of each sensor.

We have proposed a distributed approach based on the cooperation of sensors in order to interpret globally the scene [9]. Generally, it is not possible to cover the whole scene, so the sensors are separated into blind zones, for which we do not have any observation. One of the principal difficulties is to ensure a robust recognition of the mobile objects perceived by the different sensors from different points of view at different moments.

We pointed out in [9] that temporal information modeled by a fuzzy curve allows to roughly predict the possible arrivals of an object in front of the closest sensors likely to perceive it. Then, thanks to this information, each sensor is able to match the perceived objects with the expected ones. This approach has been tested in a highway environment in order to obtain a global interpretation of traffic flow. Meanwhile, with large blind area, this approach can be used efficiently provided we have enough additional visual information permitting objects discrimination.

Above, we have mentioned that for our human activity monitoring, objects are small in size and lack visual characteristics that can be extracted. So we imposed that sensors fields of view of be closed to each other in order to reduce uncertainties induced from the prediction over large blind areas. In this situation we can secure a reliable matching from a sensor to another.
This distributed tracking level is implemented in a decentralized architecture where each sensor is autonomous and the cooperation is based on a message passing procedure. Messages contain information concerning object (region + belief) crossing from a field of a sensor to another. The sensor that has tracked an object indicates to its neighbor the possibility of its appearance, see fig. 7. Our leading argument favoring a decentralized approach is the distribution low level image processing and to avoid image transmission to a central node. The other argument as modularity and survivability [4] of the system can also be taken into account.



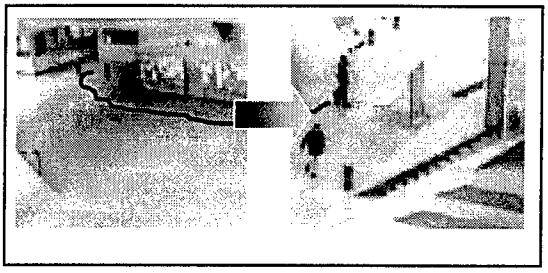


Fig 7 cooperative tracking by distributed sensors

IV application : human activity monitoring

We have tested these different components in real scene. Two vision systems were geographically distributed in order to cover the whole interested scene.
For each local vision system, we have implemented the motion segmentation and tracking on a Pentium II (350 Mhz). The algorithms run at 10 frames rate on 240x180 images, with 5 maximum simultaneous tracked objects.

Currently, the activity recognition only uses off-line data (temporally indexed trajectories) obtained by the two local vision systems. It permits to detect some specific activities. The model of activity that we use are based on fuzzy temporal graphs [10].

V Conclusion

A distributed approach of human activity, tracking and recognition has been presented, The originality of our approach is the control of the low and mid level of interpretation with respect to their instantaneous decisions. At the low level, it represents the adaptation of the background in order to tolerate illumination variation and new background object integration. At the tracking level the algorithm take into account several difficulties : objects splitting, merging and loss of target. The use of an instantaneous belief permits to reduce part of the ambiguities by taking into account past behaviors. These improvements will naturally help the quality of the final interpretation.

References

[1] N.Johnson, D.C.Hogg.Learning the distribution of object trajectories for event recognition in Pycock, E (Editor) BMVC 95, vol 2, pp583-592, 1995.

[2] L.Davis,S.Fejes, D.Harwood, Y.Yacoob, I.Hariatoglu, M.J.Black. Visual Surveillance of human activity.Proc ACCV 98 Munbai-India, 1998.

[3] A. Bobick .Movement, activity, action : the role of knowledge in perception of knowledge. Proc Workshop on knowledge-based vision in Man and Machine, London, England 1997.

[4] B.S.Y.Rao, H.F.Durrant-White J.A.Sheen. A fully decentralized multi-sensor system for tracking and surveillance. The International journal of robotic Research. Vol 12, N°1, p20-44. 1993.

[5].Hutber D., Moisan S., Shekhar C. and Thonnat M., Perception interfacing via multi-sensor data fusion and multi-program supervision for the prolab2 road vehicle. In the 7th IFAC/IFORS Symposium on Transportation Systems: Theory and Application of Advance Technology, vol 3, 1994.

[6] P.Vannoorenberghe, C.Motamed . and J.G.Postaire.,, updating a reference image for detecting motion in urban scenes (in french), Revue de traitement du signal, vol 15 n°2 p139-148, 1998.

[7] F. Brémond, & M. Thonnat. A context representation for surveillance systems. In Proc. of the Workshop on Conceptual Descriptions from Images at The European Conference on Computer Vision (ECCV), Cambridge. 1996

[8] C.Motamed, P.Vannoorenberghe. ' Behavioural knowledge for video surveillance''Actes du workshop ''Dynamic scene recognition from sensor data'' organisé par L'ONERA CERT de Toulouse et L'AFCET, Juin 1997.

[9] O. Wallart, C. Motamed & M. Benjelloun, Temporal knowledge for Cooperative Distributed Vision, seventh IEE Conference on Image Processing and its applications, Manchester ,July 1999.

[10] V.Eude, B.Bouchon-Meunier, E.Collain. Reconnaissance d'activités à l'aide de graphes temporels flous. Proc LFA'97 Logique Floue et Applications, Lyon France p 91-98, 1997.

Session RB3
Emerging Applications I
Chair: Xue-Gong Zhang
Tsinghua University, Beijing, P. R. China

Estimate Traffic with Combined Neural Network Approach

Edmond Chin-Ping Chang, Ph.D., P.E
Oak Ridge National Laboratory,
Oak Ridge, TN 37831-6207 USA E-mail: ecc2005@ornl.gov

Abstract - Many operating agencies are currently developing computerized freeway traffic management systems to support traffic operations as part of the Intelligent Transportation System (ITS) user service improvements. This study illustrates the importance of using simplified data analysis and presents a promising approach for improving demand prediction and traffic data modeling to support pro-active traffic control. This study found that the proposed approach of combining advanced neural networks and conventional error correction is promising for improved ITS applications.

Keywords: Intelligent Transportation Systems, Numerical Data Analysis, Traffic Prediction, Neural Networks.

1. Introduction

Many operating agencies are currently developing Freeway Traffic Management Systems (FTMSs) and the Intelligent Transportation System (ITS) to improve traffic control and operations along major urban freeway corridors. Often, operating agencies must design and implement various traffic response plans to provide needed system control strategies during normal, congested, incident, dangerous conditions, and pre-scheduled special events. To improve the implementation of real-time control strategies, control centers must be able to identify traffic demand pattern changes quickly based on massive amount of real-time, up-to-date traffic measures [1][2][3]. Therefore, effective decision-making supports are essential to these Traffic Management Centers (TMC) in order to integrate traffic operations, environmental measures, roadway control, and motorist information in the shortest time possible. It is especially important since many traffic management centers have been significantly expanded operations as part of the Intelligent Transportation System (ITS) user service improvements.

Significant developments have been made in applying computer models and numerical analysis techniques for evaluating traffic control alternatives and assisting the traffic system improvement analyses. With proper data calibration, either macroscopic or microscopic traffic models can be used to assist traffic operational analysis of traffic control strategies. However, the macroscopic models cannot accurately represent system behavior, while microscopic models are often too computationally intensive and unsuitable for real-time applications due to the design complexity. With the increasing applications of ITS systems, developing pro-active control strategies, and designing accurate demand prediction capabilities quickly using real-time traffic data available are essential [3][4].

Simplified traffic prediction analyses not only are essential for detecting non-recurring incidents, but are also important for identifying daily traffic patterns for practical, day-to-day operations. New traffic control system software designs can take advantage of the increasing real-time surveillance capabilities currently being installed in most freeway traffic management systems. This paper examines a practical study approach of combining both advanced neural networks and conventional error correction techniques to improve freeway traffic operational behavior analysis based directly on real-world traffic measures. In this way, this system can minimize traffic demand calibration for improved system operations.

2. STUDY BACKGROUND

The traffic control industry is moving rapidly toward real-time, proactive traffic control. It is essential to develop a practical system software design that allows efficient traffic demand prediction algorithms that can be performed automatically in real-time. This section summarizes the theoretical background, the neural network formulation, and the general neural network training procedure being used. Several numerical analysis methods, often used for

the time-series data prediction, are compared.

The automatic data reduction and analysis process should allow the user to identify traffic demand and flow pattern changes more accurately. This study will enhance traffic demand prediction functions, and examine the effectiveness of its usage in conjunction with an on-line error correction algorithm to provide improved adaptability for traffic demand prediction. Neural networks have become an emerging research area in engineering field. A neural network basically emulates the biological reasoning functions of a human brain in order to interpret and solve complicated, numerical, and pattern recognition problems. Neural network approach can especially organize massive information for improved pattern recognition. A neural network consists of three layers; each layer may contain several neurons with each neuron acts as function, such as using a sigmoid function [5].

2.1 Neural Network

A neural network training procedure includes data collection, paradigm selection, structure selection, parameter setup, and result testing. However, a large amount of relevant data should be collected in order to provide successful results. Most data are used in training, while reserving other data for later testing. Several paradigms are currently available, such as self-organizing map, backpropagation, adaptive resonance theory, and recurrent backpropagation [5]. Different paradigms have characteristics and strengths to represent specific numerical functions. For instance, self-organizing map is unsupervised and feedforward; backpropagation is supervised and feedforward; adaptive resonance theory is unsupervised and feedback; recurrent backpropagation is supervised and feedback. After training, the performance (prediction) of neural nets must be tested. Once the prediction results are satisfactory, the neural nets can then be used.

2.2 Numerical Analysis

Various numerical analysis methods have been used to analyze time series behaviors and perform prediction, including Exponential smoothing techniques, Kalman Filter, Box-Jenkins technique. These methods are limited to specific problems.

3. System Design

This study is to examine the effectiveness of the proposed neural network approach and the enhancements for the potential use of real-time traffic observations to improve traffic modeling based on real-world traffic observations. Several subtasks are designed to evaluate the promising approach as illustrated in Figure 1 "Overall Design Approach."

The system development process includes the problem analysis, neural network training, error correction, system evaluation, and application subtasks.

To assist the real-time traffic demand prediction, a neural network and an error correction algorithm were devised to provide a needed heuristic adjustments to the neural network model. The design considerations for an accurate pattern identification and allows enhancements for future automatic heuristic adjustments after the neural network have been developed, that can take advantages of numerical analysis techniques, are very important. To support this system design, this study uses the neural network training, error correction, and practical system application.

3.1 Neural Network Training

After the freeway traffic volume model is established, the pre-processed real-world freeway data are used to train neural nets. After proper network training, different neural net configurations are tested against the traffic data for appropriateness. Only successfully trained neural nets can be accepted.

3.2 Error Correction Technique

R1, a heuristic, historical, numerical based error correction algorithm, developed at Texas Transportation Institute (TTI) in 1960s, can be used to improve the on-line data prediction results by smoothing the results obtained from the neural nets as developed from the historical data records collected previously.

As illustrated in Figure 2 "Error Correction Algorithm," R1 algorithm is a heuristic-based error correction algorithm, based on the exponential smoothing concept, to improve the effectiveness of neural network traffic prediction and thereby increase its effectiveness of the proactive traffic control.

The R1 algorithm then adjusts the next prediction depending on the direction of error measured at the current observation. If the error is greater than 0, which represents an insufficient correction, the next prediction should be decreased by a certain amount. If the error is less than 0, which represents an over-correction from the prediction, the next prediction correction amount should be increased. The amount of error correction, as developed, allows the users to adjust the sensitivity of the neural network as developed through the heuristic observation, according to the subjective measures, such as the quality of the detector data. Therefore, proper correction amount can be obtained to smooth sharp prediction. In this way, errors can be minimized to predict closer to real-world traffic observations.

3.3 Practical System Application

After satisfactory performance evaluation, the neural nets and error correction algorithms are used for prediction. Since most freeway traffic control software is implemented in conventional environment, the data interfaces between the neural nets and error correction algorithms are further designed and the program was implemented the conventional C program languages.

After training, the trained neural nets can be embedded into or integrated with these traffic control applications. In this way, the user can better interact with freeway systems and monitor system traffic responses for an entire freeway. Appropriate pro-active traffic control strategies, according to the users' confidence on the quality of the detector data and the level of control strategies, can then be applied to improve freeway traffic control prediction capabilities based on real-time traffic measures.

4. Study Results

Several numerical analyses and neural net modeling experiments were performed at TTI, using I-35W freeway traffic data collected from the Fort Worth District of the Texas Department of Transportation (TxDOT).

Based on the real-time traffic volume data were obtained in the 5-minute intervals from each freeway lane and ramp, different types of data analyses were performed to examine the operational sensitivity of various data smoothing techniques, characteristics of different traffic lanes, weekday/weekend variations, and effects of seasonal variations.

As shown in Figure 3 "Traffic Flow, I-35W Study Site," a section of the interstate freeway I-35W passes through Hattie, Rosedale, Allen, Morningside, Barry, Ripy, and Seminary streets in Fort Worth, Texas. In all cases, the system was able to predict reasonably well at these locations during April 27 and 28, 1993.

5. Conclusions and Recommendations

Many operating agencies are currently developing computerized systems to improve computerized traffic management as part of the Intelligent Transportation System (ITS) user service improvements. To facilitate the prediction, diagnosis, and control decisions from the uncertain information available in most ITS systems, it is important to develop automated decision-making support techniques that can provide improved automatic traffic prediction and support proactive traffic control through simplified but practical data analysis techniques. In addition, self-learning, automatic adjustment, and human interface functions, being designed, can later be integrated into the ITS system data warehouse to provide automatic system tuning and calibration based on the real-time traffic measures as these systems expand in the future.

This study found that the proposed combined approach of neural networks and error correction algorithm is promising for traffic prediction and proactive control. Once the neural net models are successfully trained, the system can quickly pick up demand trends for pro-active traffic demand management. The error correction algorithm can further smooth out errors that may be caused by sharp neural net prediction. The error correction algorithm can also provide human interactions after the neural network has been developed; therefore, can improve traffic system prediction.

Further study is also recommended to use traffic data from other more heavily loaded freeways for additional analysis using this technique. In addition, further traffic estimation algorithm evaluation is recommended to examine the prediction capability using traffic observations from freeways located at different areas.

6. References

[1] Stephanedes, Y.J. and Chang, K.K, "Optimal Control of Freeway Corridors," J. of Trans. Eng., Vol. 119, No. 4, 1993.

[2] Davis, G.A., "Estimating Freeway Demand Patterns and Impact of Uncertainty on Ramp Controls," J. of Trans. Eng., Vol. 119, No. 4, 1993.

[3] Chang, Edmond C., and Kunhuang Huarng, "Freeway Incident Detection using Advanced Technologies," 1992 Annual Transportation Research Board Meeting, 01/1993.

[4] Chang, Edmond C. "A Neural Network Approach to Freeway Incident Detection," Proceeding. 3rd International Conference on Vehicle Navigation & Information Systems (VNIS'92 RTI/IVHS), Norweigian Society of Chartered Engineers, Oslo Plaza, 09/1992.

[5] Chang, Edmond C. "Traffic Estimation for Proactive Freeway Traffic Control." (990164), 1999 Transportation Research Board Meeting, Washington, D.C. 01/1999.

[6] BrainMaker User's Guide and Reference Manual, 5th Ed., CSS, CA, 1990.

AUTHOR BIOGRAPHY - Edmond Chin-Ping Chang, Ph.D., P.E. is the ITS Program Manger, Energy Division, Oak Ridge National Laboratory after working as an Research Engineer of the Texas Transportation Institute at the Texas A&M University System. Dr. Chang has over 24 years of Transportation Engineering, Transportation Management, and System Engineering experience, and authored over 400 publications. Dr. Chang is a Registered Professional Engineer, Certified Civil Engineer in Texas and in Taiwan, R.O.C., the Senior Advisor of United Nations Development Program's (UNDP) STAR and TOKTEN program. Dr. Chang is also members of many active international ITS standards development professional organizations, including ITS America Standards & Protocol (S&P) Committee, ITS America System Architecture (SA) Committee, AASHTO NTCIP Ramp Metering (RMC) Working Group, U.S. Technical Advisory Group (USTAG) to ISO TC204 Committee. Dr. Change's professional research interests include: Computerized Freeway and Arterial Traffic Management Systems, Signal Timing Optimization, Urban Transportation Systems Analysis, Traffic Flow Theory, Real-Time Traffic Simulation, Microcomputer Applications, Expert Systems, Artificial Intelligence applications, and Intelligent Transportation Systems and NTCIP.

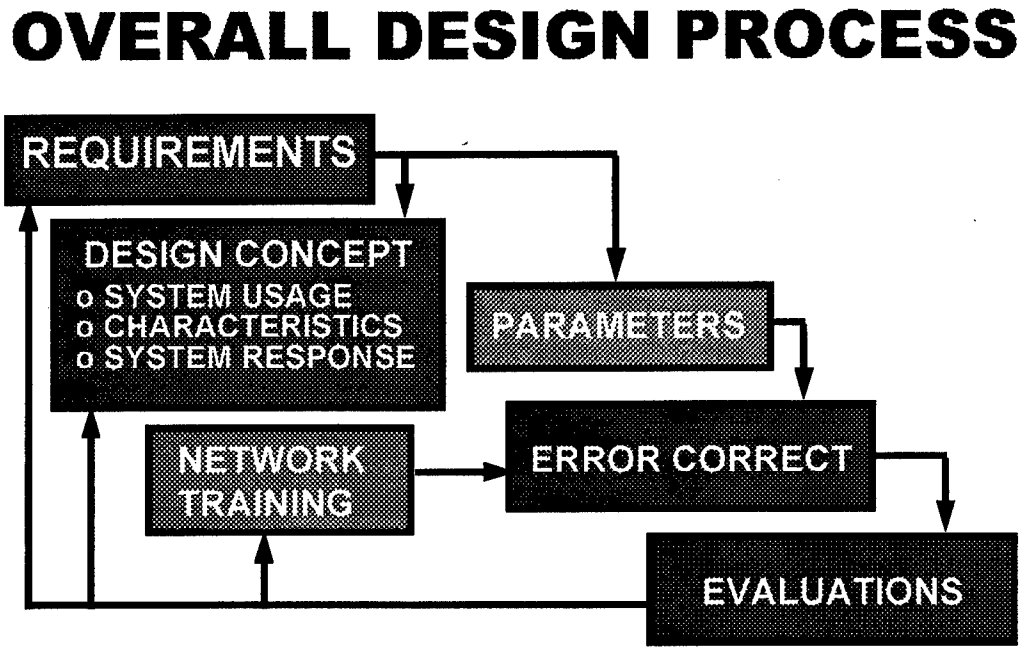


Figure 1. Overall Design Approach.

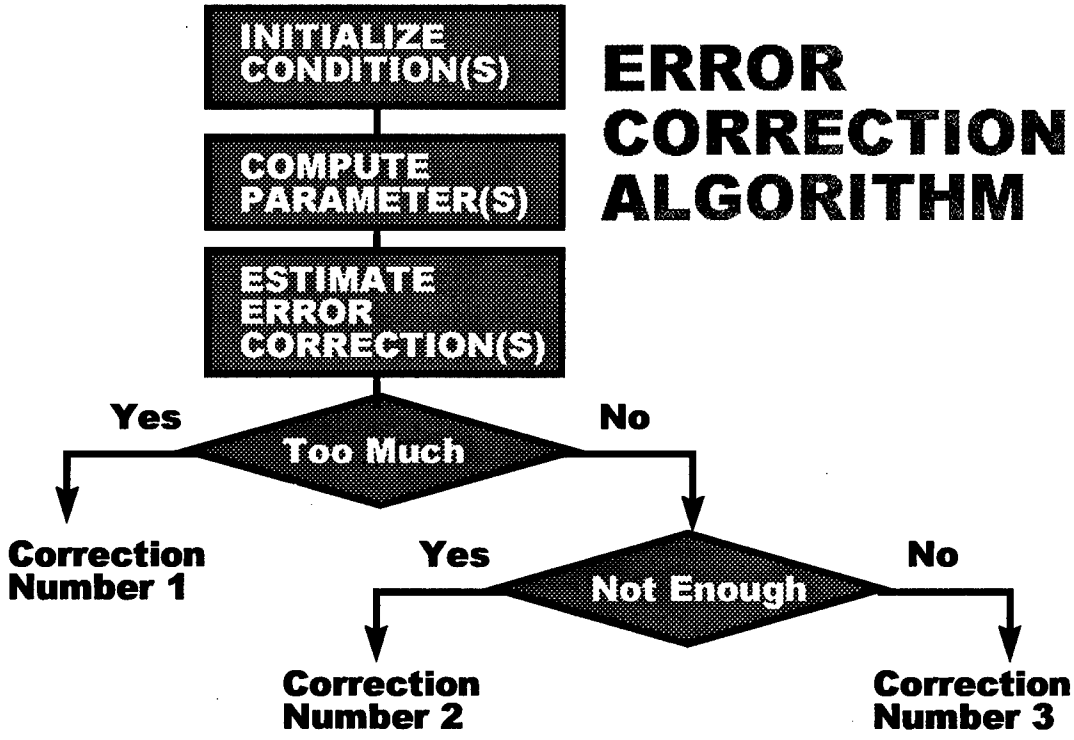


Figure 2. Error Correction Algorithm,

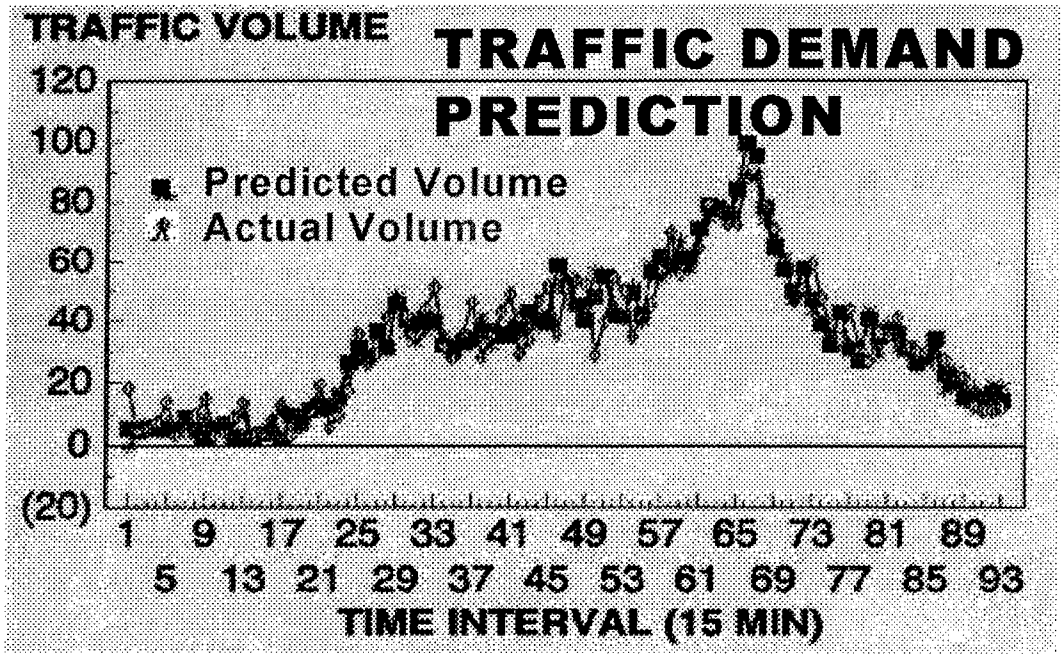


Figure 3. Traffic Flow, I-35W Study Site.

Combining Multiple Biometric Person Authentication Systems

W. Liu, N. Mukawa and H. Sakano
Research and Development Headquarters
NTT Data Corporation
Tokyo, Japan
{wjliu, mukawa, sakano@rd.nttdata.co.jp}

Abstract *In recent years, varieties of person recognition systems based on biometric characteristics, such as fingerprint and signature, have been developed. These systems can be divided into identification systems (classifiers) and authentication systems (verifiers). In the paper the latter is discussed. Since a verifier is evaluated with performance indexes different from that of a classifier, existing methods of combining multiple classifiers should be adapted for combining multiple verifiers. A method is proposed to estimate the performance for the combined system, and it is suggested that the combination is implemented through choosing the combination rule and adjusting the thresholds of all individual verifiers based on the estimated performance. The method is compared with methods based on logical formalism and Bayesian formalism. In an experiment to combine three biometric authentication systems, it shows improved results.*
Keywords: Biometric person authentication, Verification, Combination

1 Introduction

Biometric person authentication refers to recognition of an individual based on his/her physiological or behavioral characteristics, such as face, voice, fingerprint and signature. Though face and voice seem to have been used naturally for person authentication from ancient times, and fingerprint and signature have also been researched for more than dozens of years, implementation of an automatic biometric person authentication system on a machine was proved a difficult task[9]. But in recent years, according to the development of sensor devices and recognition algorithms, a variety of commercial products for automatic biometric person authentication have been released. These products have been based on face, facial thermogram, fingerprint, hand geometry, hand vein, iris, keystroke, retinal pattern, signature, voice and so on, and have been applied to secure physical access control, computer logon, voting, and other fields[7].

In the paper, we discuss how to combine two or more such biometric person authentication systems. It is thought there are at least two advantages with a combination: 1. Every biometric characteristic has its limitation in applications, for example, fingerprint is hardly recognized for dry or oil skins, and face works only under suitable illumination conditions. So multiple characteristics may be necessary in practice to ensure all users can be accepted. 2. To decrease recognition errors, improving the performance of a system based on one characteristic may be costlier at present than combining multiple existing systems based on different characteristics.

It is noticed that methods of combining multiple *classifiers* have been proposed in recent years as a new direction for the development of highly reliable character recognition [3, 5, 8, 10] and biometric person identification[1, 4, 8] systems. But the concept of combining multiple *verifiers*, to which the biometric person authentication systems belong, has not yet been

well studied. In fact we can divide biometric person recognition systems into two alternative categories according to their configurations, which we denote *classifiers* and *verifiers* in the paper. They are described as follows:

- A *classifier* identifies one person by comparing a biometric trait against a database of previously stored biometric traits of *many* people. Since its matching process between trait measurements and their templates is one-to-many, it is often denoted as *person certification* or *person identification.*

- A *verifier* validates a person's identity by comparing a biometric trait against *his own* previously stored biometric trait. Since its matching process between trait measurements and their templates is one-to-one, it is often denoted as *person verification* or *person authentication.*

Unlike a classifier, a verifier requires inputs of not only the feature measurements but also the label (an entry number or identity number, etc.) of an individual to be recognized, while its output is about if the individual should be accepted or rejected. The indexes to evaluate the performance of a verifier include the error rates with respect to the *Acceptance* and the *Rejection*. Because after combining multiple verifiers the whole system performs still as a verifier, it is reasonable to evaluate a combination method through estimating the performance indexes for the combined verifier.

In the paper, we consider multiple verifier combination at decision level, where each individual verifier to be combined is required to output a dicision of either *Acceptance* or *Rejection* according to its own information. Another kind of combination is at score level where each individual verifier outputs a real value to indicate a degree to *Acceptance* or *Rejection*. Obviously the score level combination can be converted to the decision level combination because each individual verifier can be forced to make a decision based on its score.

It is intended to build a framework of combination for verifiers. A method, which is denoted as *combination based on optimum formalism*, is proposed to estimate the performance for the combined system, and it is suggested that the combination is implemented through choosing the combination rule and adjusting the thresholds of all individual verifiers based on the estimated performance.

The paper is organized as follows. In Section 2 the performance indexes used to evaluate a verifier are introduced. In Section 3, three combination methods based on logical formalism, Bayesian formalism and optimum formalism are described respectively. The third one represents our proposal. The three methods are investigated with an experiment to combine three biometric authentication systems in Section 4. Finally Section 5 shows short conclusions.

2 Performance Indexes of Verifiers

In research and practice, a verifier is usually evaluated by two indexes, namely *FAR(False Acceptance Rate)* and *FRR(False Rejection Rate)*. As indicated in Figure 1, FAR and FRR are both functions of a threshold t, which is compared to a similarity (or a distance) measure between a feature measurement and its template.

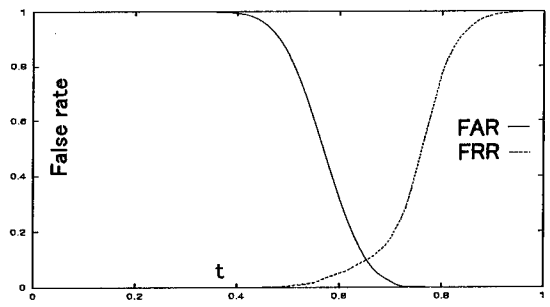


Figure 1: FAR and FRR of a verifier

Without losing generality, a person to be verified is represented by a d dimensional fea-

ture vector X, which represents a point of the feature vector space V. Any decision rule with a threshold t divides V into two exclusive parts of $V_A(t)$ and $V_R(t)$: $V_A(t) + V_R(t) = V$. If $X \in V_A(t)$, the decision is *Acceptance*, otherwise *Rejection*. At the case, $FAR(t)$ and $FRR(t)$ are defined in theory as follows:

$$FAR(t) = P(R) \int_{V_A(t)} p(X|R)dX, \quad (1)$$

$$FRR(t) = P(A) \int_{V_R(t)} p(X|A)dX, \quad (2)$$

where, $P(A)$ and $P(R)$ represent the a priori probabilities for *Acceptance* and *Rejection*, respectively, and satisfy $P(A) + P(R) = 1$. $p(X|A)$ and $p(X|R)$ represent the conditional probability density functions(*pdf*) of X under conditions of *Acceptance* and *Rejection*, respectively.

But in practice, the $FAR(t)$ and $FRR(t)$ curves are seldom calculated using the equations (1) and (2) because the calculation of the multiple dimensional *pdf*'s and the integration are very difficult. Instead the curves are usually obtained based on counting error rates of a number of testing samples.

If t corresponds to a similarity measure, $FAR(t)$ is a monotonic decreasing function while $FRR(t)$ a monotonic increasing function of t. When t varies, it is certain that one of the two indexes becomes better while the other becomes worse than before. So a trade-off between $FAR(t)$ and $FRR(t)$ must be taken to evaluate a verifier. One usual choice is to minimize the so-called *EER(Equal Error Rate)*(or called *cross-over error rate*), which means the value of $FAR(t)$ at the t where $FAR(t) = FRR(t)$. It is noticed that minimizing EER is equivallent to minimizing $max\{FAR(t), FRR(t)\}, \forall t$. Another often used trade-off is to minimize $FRR(t)$ (or $FAR(t)$) when $FAR(t)$ (or $FRR(t)$) is not greater than a specified value, which is equivallent to the $Neyman - Pearson$ test[2] if the *Acceptance* and the *Rejection* are considered to be two usual classes.

3 Combination Methods

Assume there are k verifiers to be combined. The ith$(i = 1, \sim, k)$ verifier is with the output e_i, $e_i = A$ or $e_i = R$ (from now on, *Acceptance* and *Rejection* are simplified to A and R respectively in mathematical expressions). All the e_i's are inputs to a combination rule and generate a combined output e, $e = A$ or $e = R$.

3.1 Combination Based on Logical Formalism

Because the outputs of a verifier have binary-state values, the simplest combination rules are implemented according to some logical operators, such as AND and OR operators. With respect to AND, the combined system outputs *Acceptance* if and only if all individual verifiers output *Acceptance*, so it is strict to get a combined output of *Acceptance*. In othor words, the AND rule ensures FAR of the combined system less than any of the individual verifiers, but makes FRR even worse than the worst of the individual ones. On the other hand, with the OR rule, the combined *Acceptance* is obtained if there is at least one individual verifier to output *Acceptance*. As the OR rule outputs *Rejection* only when all individual verifiers output *Rejection*, it makes FRR better but FAR worse than any individual verifiers.

Otherwise the *Majority* rule is also widely used. At the case, the combined system outputs *Acceptance* when majority of the individual verifiers output *Acceptance*. The *Majority* rule is in a sense between the AND rule and the OR rule, which gives less improvement to one of FAR and FRR meanwhile less deterioration to the other.

3.2 Combination Based on Bayesian Formalism

The combination method based on Bayesian formalism is proved effective to combining multiple classifiers[10]. We show below how to adapt it to verifiers.

When the output e_i of the ith verifier is obtained, its decision error can be described by its confusion matrix

$$C_i = \begin{bmatrix} n_{aa}^{(i)} & n_{ar}^{(i)} \\ n_{ra}^{(i)} & n_{rr}^{(i)} \end{bmatrix}, \quad (3)$$

where the two rows correspond to the *Acceptance* and the *Rejection* categories respectively, and the two columns correspond to the events of $e_i = A$ and $e_i = R$ respectively. Then $n_{aa}^{(i)}$ denotes that $n_{aa}^{(i)}$ samples from *Acceptance* category have been assigned *Acceptance* by the ith verifier, and $n_{ar}^{(i)}$ denotes $n_{ar}^{(i)}$ samples from *Acceptance* category been assigned *Rejection*.

When a verifier is given with its $FAR(t)$ and $FRR(t)$ curves, the confusion matrix then can be described without additive testing, that is,

$$C_i = \begin{bmatrix} [1 - FRR(t)]N_A & FRR(t)N_A \\ FAR(t)N_R & [1 - FAR(t)]N_R \end{bmatrix}, \quad (4)$$

where, N_A and N_R are the equivalent numbers of testing samples for *Acceptance* and *Rejection* respectively.

Based on the confusion matrix, probabilities of $e_i = A$ or $e_i = R$ under conditions of *Acceptance* or *Rejection* are estimated by

$$P(e_i = A|A) = 1 - FRR(t), \quad (5)$$

$$P(e_i = R|A) = FRR(t), \quad (6)$$

$$P(e_i = A|R) = FAR(t), \quad (7)$$

$$P(e_i = R|R) = 1 - FAR(t). \quad (8)$$

When k independent individual verifiers $e_i (i = 1, \cdots, k)$ are combined, the probabilities of $P(A|e_1, \cdots, e_k)$ and $P(R|e_1, \cdots, e_k)$ are calculated based on Bayes theorem, that is,

$$P(A|e_1, \cdots, e_k) = \frac{P(A)\prod_{i=1}^{k} P(e_i|A)}{\prod_{i=1}^{k} P(e_i)}, \quad (9)$$

and

$$P(R|e_1, \cdots, e_k) = \frac{P(R)\prod_{i=1}^{k} P(e_i|R)}{\prod_{i=1}^{k} P(e_i)}. \quad (10)$$

Thus the combination rule based on Bayesian formalism is as follows (assume $P(A) = P(R)$),

$$\begin{aligned} \text{if} \quad & \prod_{i=1}^{k} P(e_i|A) > \prod_{i=1}^{k} P(e_i|R) \\ \text{then} \quad & e = A, \\ \text{else} \quad & e = R. \end{aligned} \quad (11)$$

3.3 Combination Based on Optimum Formalism

Neither the combination based on logical formalism nor that based on Bayesian formalism considers definitely the FAR and FRR of the combined system. Thus it is not ensured that the combined system is an optimal verifier. Generally there is a common and intuitive assumption that the combination of multiple verifiers must improve performance, because "surely more information is better than less information". But on the other hand, a different intuition suggests that if a stong verifier is combined with a weaker one, the resulting decision environment is in a sense averaged, and the combined performance will be degraded from the performance that would be obtained by relying solely on the stronger one. There is truth in both intuitions. The key to resolving the apparent paradox is that when two verifiers are combined, one of the resulting error rates (FAR or FRR) becomes better than that of the stronger one, while the other error rate becomes worse even than that of the weaker one. If the two verifiers differ significantly in their power, and each operates at its own optimum working state, then combining them may give significantly worse performance than relying solely on the stronger one.

To make the combined system optimal, that is, to make the trade-off between FAR and FRR of the combined system satisfying a prespecified objective, it is necessary at first to estimate the FAR and FRR. Then based on the estimated FAR and FRR some related factors should be optimally chosen. We indicate that the related factors include both the combination rules and the thresholds of all individual

verifiers. As have shown in the previous subsections, there exist various combination rules with different properties, so a procedure for selecting the one best suitable to a particular case is necessary. Meanwhile by adjusting a threshold, which determines the working state of a verifier, we can make an individual verifier to contribute its most to the combined system.

The optimization is implemented under the so-called Behavior-Knowledge Space(BKS)[5]. The concept is proposed originally to consider combinations of non-independent classifiers based on a number of well selected testing samples. We adopt the concept but make a different use of it. At our case, as shown in Figure 2, a BKS is composed of k dimen-

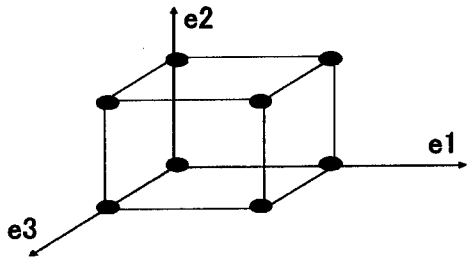


Figure 2: The three dimensional BKS

sionals, where each dimensional consists of values of the possible outputs of one verifier. As there are just *Acceptance* and *Rejection*, a k dimensional BKS consists of 2^k nodes. Each node represents a possible input to the combination rule. A combination rule may assign a value of either *Acceptance* or *Rejection* to a node. Thus there exist 2^{2^k} possible combination rules. But according to suitable reasoning the rules can be organized into groups. For instance, a combined *Acceptance* decision should remain unchanged when an individual decision varies from *Rejection* to *Acceptance*, while a combined *Rejection* decision should remain unchanged when an individual decision varies from *Acceptance* to *Rejection*. Under such conditions the number of combination rules for k verifiers are greatly decreased to $\sum_{i=1}^{k}[\sum_{j=1}^{C_k^i} C_{C_k^i}^j]$, where $C_k^i = \frac{k!}{i!(k-i)!}$. In detail, when $k = 3$ there are 15 rules, and $k = 4$, 94 rules, and $k = 5$, 2107 rules. So it becomes possible to compare all rules exhaustively when k is not big.

At the mth$(m = 1, \cdots, 2^k)$ node, $FAR_m(t)$ and $FRR_m(t)$ are calculated by

$$FAR_m(t) = \prod_{i=1}^{k} P(e_i|R), \tag{12}$$

$$FRR_m(t) = \prod_{i=1}^{k} P(e_i|A), \tag{13}$$

where, $P(e_i|A)$ and $P(e_i|R)$ are defined in equations (5)$\sim$(8). Since any combination rule divides the BKS into either the *Acceptance* or the *Rejection* categories, its corresponding $FAR(t)$ and $FRR(t)$ are obtained by

$$FAR(t) = \sum_{m \in M_A} FAR_m(t), \tag{14}$$

$$FRR(t) = 1 - \sum_{m \in M_A} FRR_m(t), \tag{15}$$

where, M_A is the node set of *Acceptance* category.

Since it can be easily proved that

$$\sum_{m \in M} FAR_m(t) = 1, \tag{16}$$

$$\sum_{m \in M} FRR_m(t) = 1, \tag{17}$$

where M represents all nodes in a BKS, the $FAR(t)$ and $FRR(t)$ of equations (14) and (15) can also be obtained using the nodes in *Rejection* category.

Based on the estimated $FAR(t)$ and $FRR(t)$ in equations (14) and (15), all possible combination rules as well as all possible setting for the individual thresholds are compared. And the optimal ones according to a specified trade-off can be determined.

3.4 Comparison of Bayesian Formalism and Optimum Formalism

When the equation (11) is compared to the equations (12)$\sim$(15), it is noticed that the

combination rule based on Bayesian formalism minimizes $FAR(t) + FRR(t)$ of the combined system. Since the minimization is different from the usual objectives to a verifier, it is certain that the combination rule based on Bayesian formalism is not a general optimal combination rule.

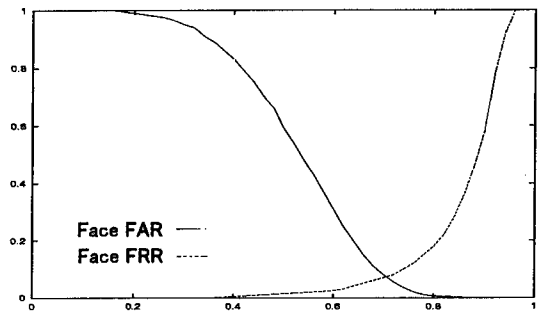
Furthermore, it is noticed that $FAR(t) + FRR(t)$ of a verifier corresponds to the Bayes' recognition error[2] if the *Acceptance* and *Rejection* categories are taken as two usual classes in the *classifier* researches.

4 Experiments

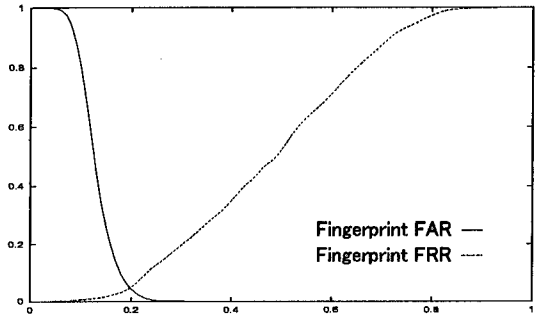
Figure 3 shows the FAR and FRR curves of three verifiers, where (a) a face and (c) a voice based verifiers are under development in our research group, and (b) a fingerprint based verifier is adapted from [6]. Thresholds of the three verifiers have been scaled to [0,1]. When the three thresholds are assumed, the FAR and FRR of the combined system can be calculated from these curves for any combination rules including those based on logical formalism and Bayesian formalism.

The procedure of combination algorithm is as follows. Assume a setting of thresholds of all individual verifiers, the corresponding FAR and FRR can be obtained from the known curves. From equations of (12) and (13), FAR and FRR at all nodes of the BKS are calculated. Possible combination rules are organized into groups, and each group gives one division of the BKS into *Acceptance* and *Rejection* categories. From equations (14) and (15), the FAR and FRR for the combined system can be estimated. By investigating all rule groups and threshold settings we can get the experiment results.

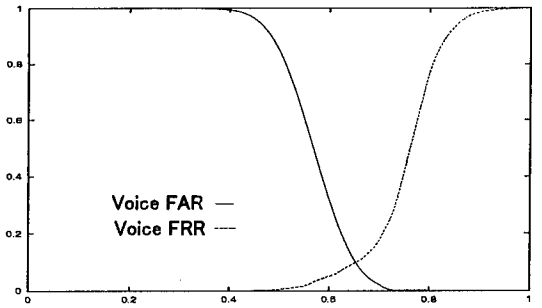
Two experiments are executed. In the first the objective function is set to minimize EER. When the three verifiers are assumed to work with their EER corresponding thresholds (Table 1), the FAR and FRR results of the combination methods based on logical AND, logical OR, logical *Majority*, and Bayesian formal-



(a) Face



(b) Fingerprint



(c) Voice

Figure 3: FAR and FRR curves of three verifiers

Table 1: $EER(= FAR = FRR)$ of individual verifiers

	face	fingerprint	voice
threshold	0.71	0.20	0.66
$EER(\%)$	7.13	4.66	10.14

ism are given in Table 2. But we can not ensure

Table 2: Results after combination(%)

	FAR	FRR
AND	0.03	20.43
OR	20.43	0.03
Majority	1.46	1.46
Bayesian	1.46	1.46
Optimal	1.23	1.23

that the individual EER corresponding thresholds are the best setting to make the EER of the combined system minimum. According to our proposed method based on optimum formalism, it is found that when the thresholds are set to (0.71, 0.18, 0.68), a majority rule show optimal FAR and FRR result (Table 2).

In the second experiment, the objective function is set to minimize FRR under the condition of $FAR \leq 0.5\%$. The corresponding thresholds, FAR and FRR of the individual verifiers are set as shown in Table 3. The combined results are shown in Table 4, where the

Table 3: FRR of individual verifiers($FAR = 0.5\%$)

	face	fingerprint	voice
threshold	0.81	0.25	0.72
$FAR(\%)$	0.45	0.30	0.33
$FRR(\%)$	20.0	12.88	24.29

Table 4: Results after combination(%)

	FAR	FRR
AND	0	47.24
OR	1.08	0.63
Majority	0.004	9.31
Bayesian	1.08	0.63
Optimal	0.48	1.23

final result is found when the thresholds are adjusted to (0.86, 0.26, 0.72) and an OR rule is tested.

It is noted that if without adjusting thresholds, the above Optimal results are actually the same as that of the *Majority* and *Bayesian* in the first experiment, and the OR and *Bayesian* in the second. It shows that when logical combination methods are applied, alternative kinds should be investigated according to the system requirement. Meanwhile, the *Bayesian* method is proved to be effective when few verifiers with similar strength are combined.

5 Conclusions

Biometric person recognition systems are divided into identification systems (classifiers) and authentification systems (verifiers). Since the verifiers are evaluated with a trade-off between the *False Acceptance Rate*(FAR) and the *False Rejection Rate*(FRR), which is different from the performance indexes used for classifiers, existing methods for combining multiple classifiers should be adapted or improved.

To implement combination of verifiers effectively, the FAR and FRR of the combined system should be estimated. We have proposed a method for the estimation from the FAR and FRR curves of the individual verifiers. The method is built based on probability theory. With the estimated FAR and FRR, it is suggested that both the combination rules and the thresholds of all individual verifiers are optimally adjusted simultaneously.

The proposed method is compared to the traditional logical operation methods and a method adapted from combining classifiers through analyses and experiments, and it shows better results.

References

[1] R.Brunelli and D.Falavigna, "Person identification using multiple cues," *IEEE Trans. on PAMI*, vol.17, no.10, pp.955-966, Oct.1990.

[2] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, Inc., 1972.

[3] K. Ho, J. J. Hull, and S. N. Srihari, "Decision combination in multiple classifier systems," *IEEE Trans. on PAMI*, vol.16, no.1, pp.66-75, Jan. 1994.

[4] L. Hong and A. Jain, "Integrating faces and fingerprints for personal identification," *IEEE Trans. on PAMI*, vol.20, no.12, pp.1295-1307, Dec. 1998.

[5] Y. S. Huang and C. Y. Suen, "A method of combining multiple experts for the recognition of unconstrained handwritten numerals," *IEEE Trans. on PAMI*, vol.17, no.1, pp.90-94, Jan. 1990.

[6] A.Jain, L.Hong and R.Bolle, "On-line fingerprint verification," *IEEE Trans. on PAMI*, vol.19, no.4, pp.302-313, Arp.1997.

[7] A. Jain, R. Bolla, and S. Pankanti, eds., *Biometrics: Personal Identification in Networked Society*, Norwell, Mass.: Kluwer Academic Publishers, 1999.

[8] J. Kittler, M. Hatef, R. P. W. Duin and J. Matas, "On combining classifiers," *IEEE Trans. on PAMI*, vol.20, no.3, pp.226-239, Mar. 1998.

[9] R. M. Stock and C. W. Swonger, "Development and evaluation of a reader of fingerprint minutiae," Cornell Aeronautical Laboratory, Technical Report CAL No.XM-2478-X-1:13-17, 1969.

[10] L. Xu, A. Krzyzak, and C. Y. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition," *IEEE Trans. on SMC*, vol.22, no.3, pp.418-435, May/Jun. 1990.

Collaboration of Information from Different Sources for Petroleum Reservoir Prediction[1]

Xuegong Zhang
Dept. of Automation, Tsinghua University, Beijing 100084, China
email: xzhang@simba.au.tsinghua.edu.cn

Abstract - *Analyzing potential subsurface petroleum reservoirs and predicting their spatial distribution is a difficult problem, for the targets are usually thousands of meters deep in the earth and people almost have no way to directly observe them. Prospecting seismic data and well data are the two major categories of data that can be obtained for the task. They are different in mechanism and have their own characteristics, and both are imprecise and incomplete. The third information source is knowledge and know-how of human experts. In this paper, an information fusion scheme is presented for predicting potential reservoirs by the collaboration of information from different sources. Neural networks are applied in both supervised and unsupervised mode in the scheme, and human-computer cooperation is also involved for this complicated task. Practical applications have shown that this information-fusion approach is very powerful.*

Keywords: data fusion, neural networks, petroleum reservoir analysis, indeterminate information, SOMA

1. Introduction

Analyzing potential subsurface petroleum reservoirs and predicting their spatial distribution is a difficult problem, for the targets are usually thousands of meters deep in the earth and actually there is almost no method to directly observe these targets (for reasonable cost). In today's petroleum industry, most observation data that can be obtained about subsurface reservoirs are of two major categories: prospecting seismic data and well data. Well data are recorded in actual drilling holes, so they can be regarded in certain sense as a kind of direct and precise information. But they are available at only sparse locations in the investigated area, and are usually far from being able to provide adequate information about the spatial distributions of potential reservoirs. This is especially true for areas at early stages of exploration, because of the high expense and risk of drilling. On the other hand, seismic data are recorded reflection signals from subsurface interfaces, which are observed on the surface, and are relatively less expensive. So seismic data can be widely available for the whole area, but the information that they can provide about subsurface reservoirs are very indirect, inaccurate and indeterminate. The third source of information for reservoir analysis is human judgement based on experts' knowledge and experience, which usually plays a key role in the final decision-making. However, humans are not good at directly analyzing the large amount of observation data, and the rules behind human judgement are mostly quite ambiguous and often differ from case to case, which makes the efforts for coding them into some machine systems unfruitful till now. So the only choice seems to be designing some systems that can efficiently collaborate all these information from different sources by taking the advantages of their respective characteristics and overcome their shortcomings.

In this paper, we present such a system, which has already been proved powerful in several practical cases. The system utilizes neural networks of both supervised and unsupervised kind for data analysis, which can fit better for cases when known well data are too scarce to be applied directly as the supervisors. For the unsupervised part of the work, we developed a novel approach called SOMA[1], based on the SOM neural network model. The standard MLP neural network model with BP learning algorithm was adopted for the supervised task. Human judgement is invited into the system so that the mathematically derived results by unsupervised analysis can be better evaluated and be assigned to proper

[1] This work is supported by NSFC, the National Science Foundation of China.

physical meanings. These results then can be further be applied in supervised learning so that some details about the analyzed reservoir can be estimated. With this system, the possible distributions of petroleum reservoirs can be predicted from seismic data and very limited well data, and some important lithological parameters can also be predicted quantitatively. By the collaboration of seismic data, well data and human's analysis, the insufficiency of information in the problem can be largely compensated. One of the practical application cases will also be briefly introduced in the paper. In fact, the ideas behind this system can also be applied in other similar problems with multiple information sources of different characteristics and resolutions.

2. Supervised Analysis

Although traditionally seismic data were used only for deriving information about subsurface strata structures, it has been shown in recent decades that lithological information of potential reservoirs can also be extracted from the data, since they penetrate through these reservoirs and the differences in reservoir properties do leave "fingerprints" on the data[2][3]. The problem is that till now, people still haven't succeeded in finding some determinate models describing the relationship between seismic data and reservoir properties, which also differs for different areas and different geological environments. Thus our problem falls into the domain of estimating unknown dependencies from observations.

Since at locations of wells, it can be regarded that the information of subsurface reservoirs are known, well data can play the role of supervisors or training examples for our problem. The general idea can be illustrated by the diagram of Figure 1, which can be viewed as a standard supervised analysis system. In the system, well data are used to training some learning machine so that it can estimate the dependency relationship between seismic data and the desired reservoir properties at well locations. This relationship will then be used for predicting the reservoir properties for locations where only seismic data are available. The prediction result can be qualitative (such as whether the target stratum possibly contains oil/gas at certain locations) and quantitative (some lithological parameters such as sand-percentage, average porosity, etc. of the prospective reservoirs). The learning machine can be any of the popular ones such as the MLP neural network model with BP learning algorithm.

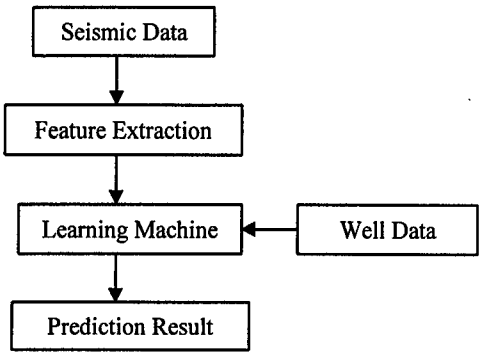


Figure 1. The basic diagram of supervised reservoir analysis

This straightforward system does succeeded in some applications. However, certain conditions must be met. Among them, one important condition is that the well data to be applied as training examples must be representative in the investigated area, and the amount of training wells must be sufficient. In fact, these conditions are very restrictive for many practical cases.

Indeed, the task is to predict the possible distribution of reservoirs, so that better decisions can be made about where should the wells be drilled. It is quite obvious that one will not drill many wells until the task is fulfilled. That's why the condition about sample size can usually not be met. On the other hand, the fact that one always wants the wells to be productive tends to make the available well data not so representative. Even if there are many wells in some area, the proportion of "negative" samples are usually much smaller than the actual probability that the target stratum is not prospective at certain locations.

3. Unsupervised Analysis and SOMA

If it can be assumed that the target stratum consists of a couple of different types with respect to the reservoir properties (such as, for simplicity, prospective and non-prospective), and if the features extracted from seismic data are well-related with the types, unsupervised learning techniques can be applied to the seismic data alone to find clusters in the data that represent these types. This is usually true for most of the practical cases.

For this purpose, we developed a novel unsupervised analysis method that we called *SOMA* (short for *SOM Analysis*). It is based on the SOM neural network model by Kohonen[4]. Comparing with most traditional clustering algorithms, SOMA is designed with less assumption about the style of the data distribution, and it does not ask users to set or estimate *a priori* the number of clusters. Since no training information is involved in the procedure, unsupervised methods can only give mathematical results, which need to be further interpreted. SOMA provides a straightforward way for users to interactively combine their knowledge and understanding of the investigated area into the clustering results [1][5].

The kernel of SOMA is the idea of SOM density map, and example of which is shown in Figure 2. The map is obtained by simply cumulating the number of samples in the data set that are being projected onto each neuron after the self-organizing learning procedure. The densities on the map can "optimally" represent the distribution of the data in the original space in certain sense, so that they can be used for making decisions on clustering in the data set. To enhance this effect, the standard SOM learning procedure should be modified slightly[5].

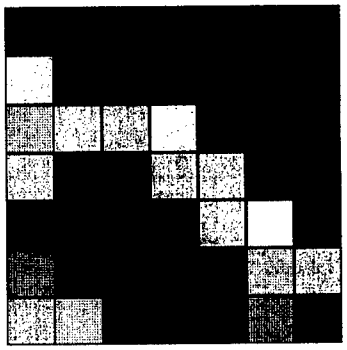


Figure 2 An example of SOM density map.

Both the proper number of clusters and their classification boundaries can be decided according the SOM density map. And even when there are no clustering relations in the data set (in which cases most traditional clustering will still blindly make the classification), the density map can still be utilized as a tool for analyzing the similarity relations between the samples in the data set. Although automatic algorithms can be developed for clustering based SOM density map, in our specific problem of petroleum reservoir analysis, we invite human experts to interpret this map manually, and by this, their knowledge and know-how can be well integrated into our whole scheme. This idea can be described by Figure 3.

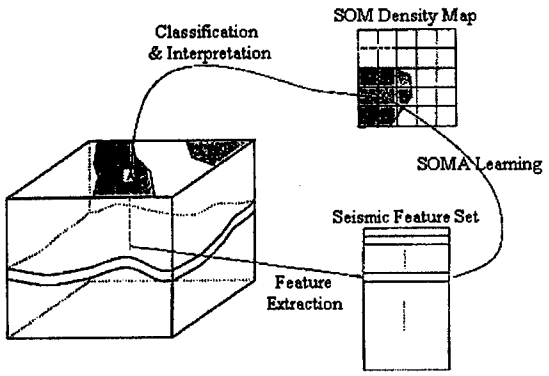


Figure 3. Diagram of SOMA reservoir analysis.

4. Supervised Learning Again

After applying SOMA, a rough understanding of possible reservoir distributions can be obtained. With this result, the restrictive limitations of applying supervised analysis can be overcome in large. The idea is, pseudo-examples can be selected according to the prediction by SOMA, so that there can be more training samples for the learning machine, and the samples can be selected to be more representative for the investigated area. In this way, some quantitative results can be obtained such as the estimated distribution of sand percentage and

thickness in the target stratum, averaged porosity and oil saturation. The whole procedure can be described by the diagram in Figure 4. The input to the whole system is the sparse and incomplete well data, the inaccurate and indeterminate seismic data of relatively low resolution, and the ambiguous expertise. The output is the prediction of possible reservoir distributions and the estimations of its lithological parameters.

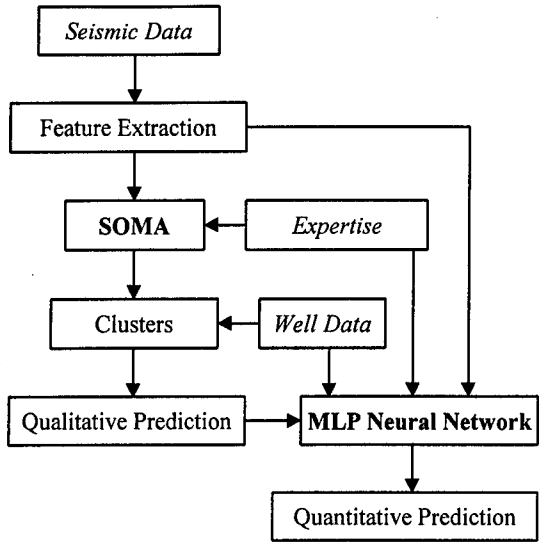


Figure 4. Collaboration of seismic data, well data and experts' knowledge and know-how for reservoir analysis and prediction, by a neural-network-based information fusion system.

5. A Practical Application Example

The above-described scheme has already been applied successfully in several practical cases. In one of them[6][7], the investigated area is about 50km^2, with only three wells (named TH-A2, TH-A5 and TH-A6). For reasons discussed in Section 2, the result of directly applying supervised methods is not acceptable. Figure 5 is the predicted reservoir distribution map at the target stratum, obtained with our SOMA technique. Using the three wells and pseudo-examples selected according to this map, a MLP neural network was trained to predict the lithological parameters. As an example, Figure 6 shows the map of the estimated average porosity at the target stratum. According these predictions, two new wells were suggested (TH-A51 and TH-A52 on the maps). They were both drilled later and were both highly productive. Table 1 is a comparison of the predicted sand thickness in the target stratum at the well location and the true values measured after the wells were drilled.

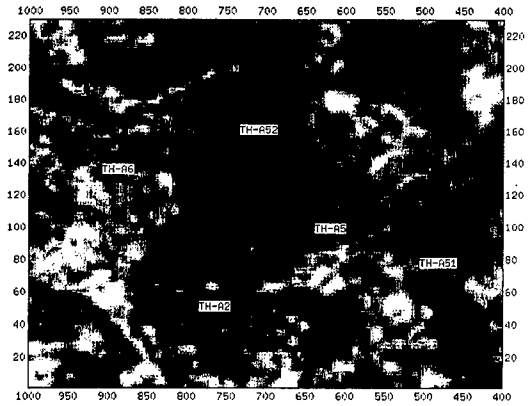


Figure 5. The predicted reservoir distribution obtained by SOMA. Dark places indicate more prospective areas. Wells TH-A51 and TH-A52 were designed according to the prediction results.

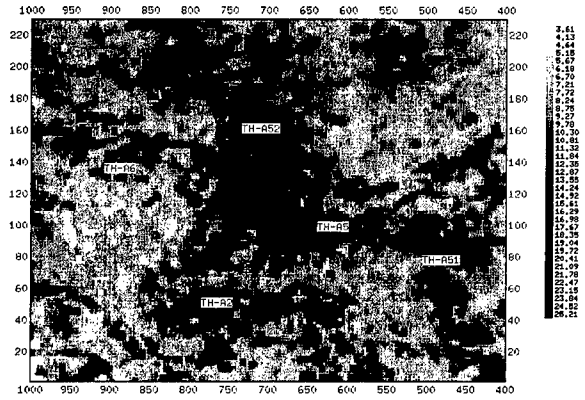


Figure 6. The estimated porosity map obtained by MLP neural network with the help of SOMA result. Gray level represents the value of average porosity, as shown in the scale at right (in percent).

Table 1. Comparison of the predicted values and the actual values of total sand thickness in the target stratum at well locations.

Wells	Target Stratum(m)	Total sand thickness	
		Actual	Estimated
TH-A2	1233-1410	35.7	*known*
TH-A5	1242-1420	51.9	*known*
TH-A6	1239-1416	21.8	*known*
TH-A51	1264-1441	42.1	42.4
TH-A52	1261-1440	45.6	42.8

6. Conclusion and Discussion

In this paper, we described a scheme of fusing information from different sources for petroleum reservoir prediction and analysis. Each kind of information alone can not give reliable and determinate results for the whole area. However, by the collaboration of these information, taking each one's advantages in their characteristics and making up their disadvantages, we can get a better comprehensive understanding of the potential subsurface reservoirs, including both qualitative predictions of reservoir distribution and quantitative estimations of the lithological parameters. The discussed scheme has already been succeeded in several practical cases, results from one of which were also presented in this paper.

The key fusing methods in the scheme are neural networks. Both supervised MLP neural network and unsupervised SOMA methods are applied in the scheme in a cooperation mode. And human experts' knowledge and know-how is also utilized to compensate for the insufficiency and inaccuracy of those "hard" data. We believe that in those large practical problems, providing human experts with the feasibility to interact with those "automatic" algorithms is very important.

Although the described scheme is designed for the special task of petroleum reservoir analysis, the ideas and general procedures can be well adapted to other similar problems, where multiple information sources exist and each one alone is incomplete and indeterminate.

References:

[1] Xuegong Zhang & Y. Li, Self-organizing map as a new method for clustering and data analysis, **Proc. of IJCNN'93**, pp. 2448-2451, Nagoya, 1993

[2] B. J. Raflpour, Seismic response for reservoir-fluid evaluation, **SPE Formation Evaluation**, March 1989, pp. 45-48

[3] R. F. Kubichek & E. A. Quincy, Statistical modeling and feature selection for seismic data pattern recognition, **Pattern Recognition**, **18**(6), 1985, pp. 441-448

[4] T. Kohonen, The self-organizing map, **Proc. of IEEE**, **78**(90) , pp. 1464-1480, 1990

[5] Xuegong Zhang, Neural Network Approaches for Pattern Recognition with Aplication to Petroleum Exploration (in Chinese), Ph.D. Dissertation, Tsinghua University, 1993

[6] Xuegong Zhang et al., Early-stage reservoir analysis with SOMA - a neural network approach, **SEG 65th Annual Meeting**, Houston, USA, 1995

[7] Xuegong Zhang et al., Estimating reservoir's lithological parameters from seismic data using neural network, **SEG 65th Annual Meeting**, Houston, 1995

Petroleum Reservoir Framework Prediction by Information Fusion[1]

Wenkai Lu, Xuegong Zhang, Yanda Li,
Department of Automation, Tsinghua University, Beijing 100084, P.R. China

Abstract: *Reservoir framework prediction is a key problem in the exploration and production of oil and gas. A novel technique is proposed in the paper, in which well-logging data and knowledge of geologists are integrated by dynamic programming and genetic algorithm. Experiment results on outcrop data showed that the technique is quite promising for practical applications.*

Keywords: reservoir framework prediction, genetic algorithm, dynamic programming, outcrop

1. Introduction

Reservoir framework prediction is a key problem in the exploration and production of oil and gas. Although in the production phase of a survey there is usually a large amount of data we can use, it is a difficult problem to fuse various information to predict the reservoir framework with high certainty. There are two major information sources that we can use. One is the mass data collected by various exploration techniques, such as well data and seismic data. They carry a lot of information of the underground reservoirs, but among them the information carried by any single type of data is quite limited and is not enough to predict the reservoir framework. Another information source is the qualitative knowledge of geologists. To get a description of the reservoir's detailed framework structure with high certainty, these information from different sources must be integrated in an efficient and reasonable approach. In recent years, much research work has been done in this field. Since it is almost impossible to know the "true" answer of any detailed subsurface structure exactly, people are studying their methodologies on outcrop data, which are actually the "subsurface" structure that goes out of the surface so that can provide a direct way to verify the performance of the proposed methods.

Most of the methods applied in today's industry for predicting reservoir framework from wells are various kinds of statistical simulation methods. They can simulate many possible answers to the problem, but fail to utilise geologists' experience and cognition about the reservoir well, so that the obtained results are too uncertain. This paper reports our research in this direction. A novel technique is proposed, in which well-logging data and knowledge of geologists are integrated by dynamic programming and genetic algorithm. Experiment results on outcrop data showed that the technique is quite promising for practical applications.

2. The Basic Ideas

In our methods, dynamic programming is used to correlate the strata between two wells. The distance of two specific strata is given according to the cognition of geologists about the investigated area. Changes of the reservoir strata from one well to another are simulated as a genetic course. Strata are treated as the populations. They have the following features: thickness, position and lithology of the stratum. Some geologic statistical results are derived from the well-logging data. The percentage of the various lithological strata and the histogram of thickness of the strata are used to form the objective function in the genetic algorithm. The evolution choice is controlled by two factors besides the objective function: (1) the cognition by geologists (including the direction of the stratum, the variation extent of the stratum, the width-to-thickness ratio of a special kind of lithological stratum, the distribution of a special lithology and the sedimentary environment), and (2) the result of stratigraphic correlation obtained

[1] This work is supported by NSFC, the National Science Foundation of China.

by dynamic programming algorithm. An uninterrupted stratum will never disappear, and only the position and thickness of the stratum change. The mutation is suitable for simulating the appearance of new strata and the disappearance of existing strata according with the object function. The whole procedure is designed as a recursive procedure: the mid-point of the two wells is predicted first, which is then treated as a new well, and then the new mid-points are predicted. The procedure will not stop till all points are predicted.

3. Stratigraphic Correlation Using Dynamic Programming

Dynamic programming is an optimization algorithm based on the back propagation of the cost. No matter how the initial stage and state are, only the last stage and state are need to decide the optimal path of the current stage.

Let there be M states in the $k+1$th stage, the traditional dynamic programming course is defined as:

$$E_{k,n} = \min_{m \in M} \left\{ L_{k,k+1,n,m} + E_{k+1,m} \right\}$$

where $E_{k+1,m}$ is the optimal cost of state m in stage $k+1$, $L_{k,k+1,n,m}$ is the cost from state n in stage k to state m in stage $k+1$, $E_{k,n}$ is the optimal cost of state n in stage k.

Dynamic programming was used in stratigraphic correlation widely [1].

In our research, the cost function is defined as the function of the thickness, position and lithology of the strata. The cost function will be modified if any priori knowledge about the strata can be used. If two strata are connected, the distance between them is set to zero. The distance between one of the two strata and any stratum else is set to a value large enough to forbid their connection. If one stratum connects with a gap (no connection exists between it and any other stratum), the distance between it and any other stratum is set to a value large enough to forbid their connection too.

4. Reservoir Frame Prediction Using Genetic Algorithms

Genetic algorithms are searching algorithms based on the mechanics of natural selection and natural genetics [2]. They combine the mechanism of survival of the fittest among string structures with a structured yet randomized information exchange strategy to form a searching algorithm with some of the innovative flair of human searching. Because the algorithm simultaneously evaluates many possible (high fitness) points in the parameter space, it is more likely to reach the global minimum (or maximum). Genetic algorithms differ from most optimization techniques by searching from one group of solutions to another, rather than from one solution to another, and it is this fact that makes them uniquely suited to multi-objective optimization problems.

In our method for reservoir framework prediction, changes of the reservoir strata from one well to another are simulated as a genetic course. Strata are treated as the populations.

4.1 Coding

Although traditional genetic algorithms map problems to strings of binary bits and manipulate these encoding, it is more natural and therefore preferable[2] to represent each solution by one three-dimensional arrays, corresponding to the lithology, the thickness, and the position of stratum respectively.

4.2 Initial Population

The initial population is generated by the information from the two wells. All strata in the two boreholes are included.

4.3 Fitness Evaluation

There are two objectives, i.e., the percentage of various lithological strata and the histogram of thickness of the strata, that are used to form the fitness function in the genetic algorithm. They can be got as statistics from all the well-logging

data if there are enough wells in the survey area. If the survey area is a new one and only a few wells can be used, we can adopt the statistics from other areas nearby. Besides the fitness function, the evolution choice is also controlled by two factors: (1) the cognition of geologists (including the direction of the stratum, the variation extent of the stratum, the width-to-thickness ratio of a specific lithological stratum, the distribution of a special lithology and the sedimentary environment), and (2) the result of stratigraphic correlation obtained by dynamic programming algorithm.

4.4 Operations

The general structure of genetic algorithms is as following:

```
Initialize population
  Calculate the fitness function
      of each individual
  Selection
  REPEAT
  Crossover
  Mutation
  Calculate the fitness function
      of each individual
  Selection
  UNTIL (termination condition satisfied)
```

The *selection* operator is based on the principle of "survival of the fittest". A specific individual having a fitness value above the average level will have more chance of being selected than those individuals having fitness value below the average level. In our research, the two individuals forming the parents come from the two boreholes at the same time. If one stratum of a strata pair (the strata is connected) is chosen as one of the parents, the other must be one of the parents too.

The *crossover* operator is used to produce new offspring from their parents, at the mean time exchanging the information between them. The *crossover* operation is done in the three dimensions. Here one can see that if the parents are a strata pair, only the thickness and the position of the child stratum change, and the lithology is kept the same as its parents.

The *mutation* operator is used to bring about new information at the bit level, so that the genetic algorithms can search new areas otherwise not accessible when searched using only selections and crossovers. In our case, the mutation operation is used to estimate the appearance of new strata and the disappearance of existing strata according with the object function.

The whole procedure is designed as a recursive procedure: the mid-point of the two wells is predicted first, which is then treated as a new well, and the new mid-points are predicted. The procedure will not stop till all points are predicted.

5. An Example

One of our outcrop sections (about 1000 meters wide and 200 meters deep) is used to verify the efficiency and performance of the proposed technique. Figure 1 shows the result: Fig.1(a) is the original section (the correct answer), Fig.1(b) is the three boreholes used for prediction (the known condition of the system), and Fig.1(c) is the predicted section according to the data in Fig.1(b) and some qualitative knowledge about the strata. Deferent colors in the figure (reproduced as gray levels in the proceedings) represent deferent kinds of lithology, and the largest interval between the boreholes in our experiment is 408 metres. It can be seen that keeping the sparseness of know data in mind, this prediction is rather accurate. And the accuracy can be further improved as the amount of information available is being increased.

6. Conclusion

In the proposed technique, the effective fusion of quantitative information from well-logging data and qualitative information from the cognition of geologists greatly reduced the uncertainty in reservoir prediction. Processing on outcrop data showed an encouraging result.

In fact, if there are other information available (such as seismic data), they can also be

integrated into the system in the future for better results.

References

[1] Smith, T.F. and Waterman, M.S., New stratigraphic correlation techniques. Journal of Geology, vol.88, no.4, pp. 451-457, 1980

[2] D.E. Goldberg, Genetic Algorithms in Search, Optimization & Machine Learning. Addison Wesley, 1989

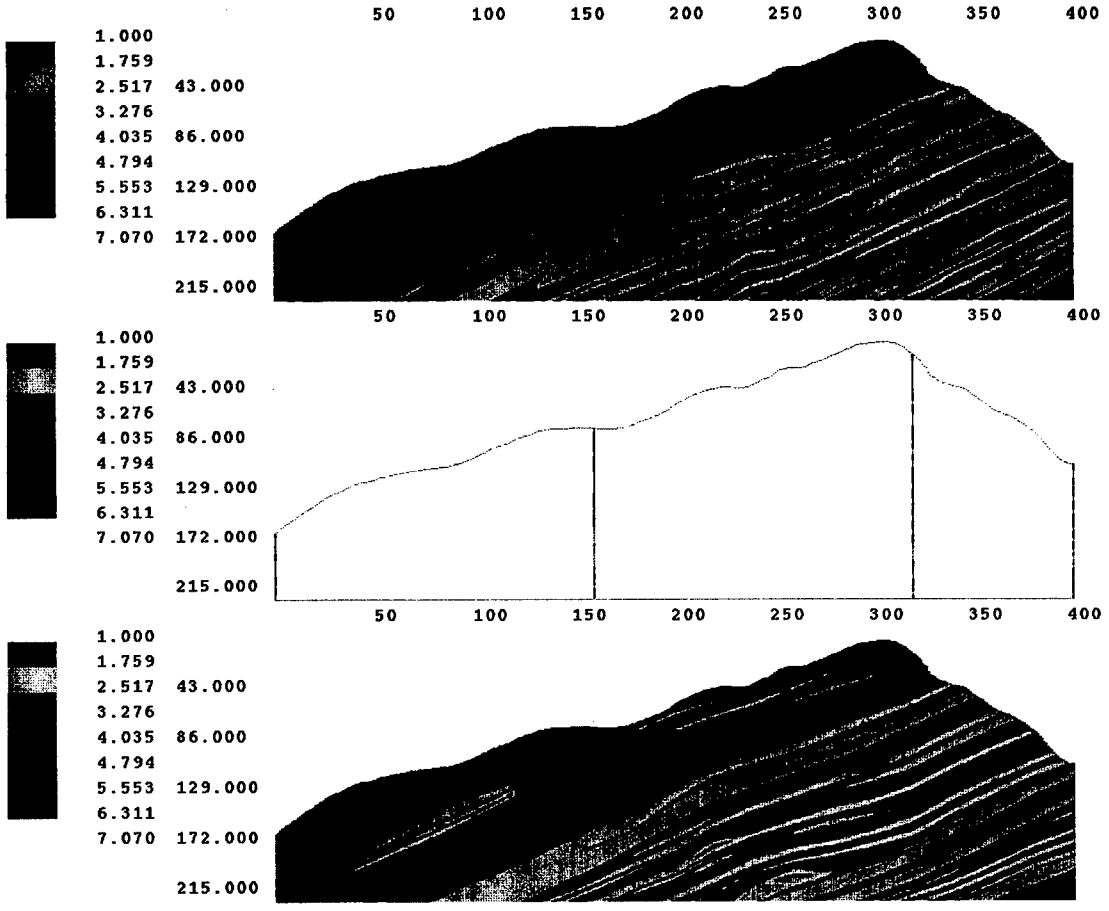


Figure 1. An experiment result of the method described in the paper on some outcrop data. (a) Top: the original section of an outcrop data; (b) Middle: the four boreholes extracted from (a); (c) Bottom: the predicted section according the data of (b) by methods described in this paper. Colors (grey levels for printing) in the figures represent different types of lithology.

Multivariate Sensor Fusion by a Neural Network Model

Hans-H. Bothe, Martin Persson, Lena Biel, Magnus Rosenholm
Applied Autonomous Sensor Systems Laboratory (AASS)
Örebro University, Fakultetsgatan 1
S-70182 Örebro, Sweden

Abstract. This paper describes a hierarchically organized technical system performing auditory-visual sound source localization and camera control. Desired applications are in the field of mobile robotics or multimedia. The measurement set-up uses four microphones and one video camera. Starting points are functionality and signal processing in the auditory and visual pathways of the central nervous system in mammals, performed with the help of neural networks. Sound and vision estimates of an intentional target are fused in order to control a virtual fovea within the vision system. For this purpose, optical signals of the CCD of the video camera are collected in macropixels which determine the grid of foveal attention control. Pre-processed sound signals are interpreted as spike-train coded action potentials to be accumulated in the neurons soma. Spike signals which arrive approximately synchronously activate an output action potential. This enables the system to perform a correlative input selection as to be used in echo cancellation, for instance. A respective technical system is designed and implemented on an industrial mobile robot. Experimental results of the behavior of the overall system are presented.

Key Words: sound source localization, neuro, audio-visual, fovea, virtual camera

1 Introduction

Goal of this work is to find improved paradigms for a smart human-machine communication and to define efficient algorithms on base of biological examples of auditory and visual perception. A particular goal of this paper is the design of a technical system for auditory-visual sound and target localization by attention control. Here, *perception* is - in distinction to sensing - understood as a high-level process to gain knowledge from sensory signals, and *attention control* is understood as an unconscious low level process which is performed to select a particular area of interest within the visual input data space.

The technical system is implemented on a mobile industrial robot and may serve for intelligent speech, gesture, and lip movements recognition with possible human instructors or supervisors. Since the robot may be considered as a severely disabled person with impairments particularly in hearing and vision, the new paradigms may also be used for improvements in the field of rehabilitation engineering. Another possible technical application is automatic camera control for videoconferencing.

The integrated use of vision and audition is a basic skill of many animals. Whereas sound information enforces the creature to focus its vision system towards specific spatial areas of interest, vision allows for association of these sounds with respective visual sound sources. The resulting low level sensor fusion enables high level recognition or identification processes of, for instance, discrete sounds or speech utterances or of specific visual objects in a complex and noisy environment.

Many traditional AI approaches to the above task integrate multimodal sensor information on a symbolic level after a separate extraction of

symbolic information from the acoustic and optic input signals. Neuro-physiological research results point out that locally distributed topological representations of the perceptional space are efficiently used in animals for uni- or multimodal auditory-visual localization or fusion tasks as well as for attention control. Exemplary work is presented in [1], [2], [3], [4], [5], and [6].

In the following, some important properties of the auditory, the vision system and of biologically inspired neural signal processing will be introduced.

2 Spatial Hearing

Main performance and properties of auditory signal processing in animals are frequency analysis, directional hearing, and pattern recognition or identification. One of the primary ways of separating sounds as a preprocessing step for subsequent recognition or identification is by identifying their spatial location.

2.1 Biological Example

Auditory sensory organs are the ears. They basically consist of pinna and outer ear channel, middle ear, and the basilar system of the inner ear. The pinna collects and amplifies external sound signal stimuli, the middle ear acts as a subsequent block for impedance transformation and amplification, the inner ear with basilar membrane and hair cells performs a *tonotopic* separation of the sound waves by means of a frequency analysis [7]. The mechanical properties of pinna and outer ear channel create a direction and frequency depending echo signal which is fed to the cochlea. The determination of the accumulated signals allows for three-dimensional sound localization, i.e. calculation of azimuth and elevation angle [8]. Cochlear models are essentially functional complete. Exemplary work can be found in [9] or [10].

Tonotopic information is processed through the central auditory pathway which consists of a serial array of connected nervous nuclei or, speaking in technical terms, processing blocks.

The first binaural interactions occur in the olivary complex which is believed to perform a non-linear crosscorrelation between the pre-processed signals of left and right ear. The neurons of the olivary nuclei synapse to the colliculus complex in which a twodimensional map is generated and stored. Activations of this map are believed to represent the spatial sound source positions calculated on base of delay times and amplitude variations of the binaural signals. Besides, the colliculus gets also visual input from specific ganglion cells, each of which reacting on stimuli in a relatively large visual area and detecting visual novelties.

Neurons of the nuclei in the auditory pathway synapse to cortical areas in which high-level sound and especially speech processing, recognition and identification, is performed.

2.2 Technical Implementation

According to the duplex theory (see [10]), interaural time differences (ITD) between the sound signals arriving at the two ears with specific delays are depending on the relative position of the sound source to the head coordinate system. They are good for horizontal localization in case of acoustic clicks up to frequencies of approximately 1.5 [kHz]. With a distance d between the acoustic sensors, a sound source with azimuth angle $\alpha \in [-\pi/2, \pi/2]$, and speed of sound waves in air c=343 [m/s], and the sound source being known to be in the frontal hemisphere, the mapping between α and ITD is approximately given by the pseudo-fixed relation

$$ITD = d/2c\ (\alpha + \sin\alpha).$$

This formula holds for particular propertiesof the sound and in the absence of echoes. The evaluation of complex sound signals can be derived from ITD if the signals are decomposed into sequences of onset clicks. This procedure is of particular significance since signal onsets are known to be a basic source of information for sound localization in humans and animals [10], often described as the *precedence effect.*

The elevation angle can be derived from i) the diffraction of the waves by the head, which generates significant differences in the intensity level (IID) especially for higher frequencies (*head-shadow effect*) , or from ii) the direction and frequency depending head-related transfer function of the information channel between the

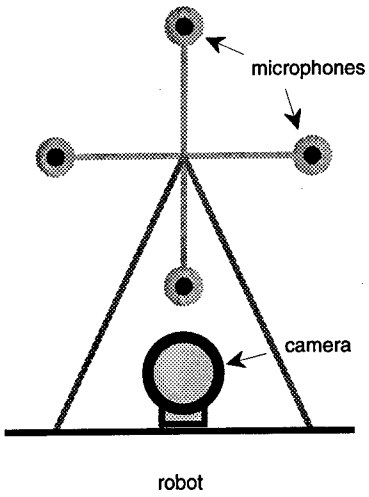


Figure 1. Sensor head for sound source localization.

sound source and the sensors. The latter procedure considers effects of echo production in pinnae and outer ear channels and leads to a large look-up table which is regarded as a numerical pulse response function. Complex spatial sound source models are derived by applying a convolution of the sound signals with the head related transfer functions.

Extending mechanisms of binaural sound recording in humans, we have designed a sensor head with two orthogonal sets of microphones, each of which detecting azimuth or elevation angle only with the help of ITD evaluation. A primary advantage of this design is that time consuming short-time convolution algorithms can be avoided. A schematic figure of the sensor head with four microphones and one video camera, is shown in figure 1.

At the moment, the technical realization of the cochlea functionality uses a simple filter bank for the convenience of short calculation time instead of a more complex model as described, for instance, in [9] or [10]. Nevertheless, the system can easily be extended. We use three 4th order Chebytchev type bandpass filters with edge frequencies of

filter 1: 1000 - 1778 Hz,
filter 2: 1778 - 3162 Hz,
filter 3: 3162 - 5623 Hz.

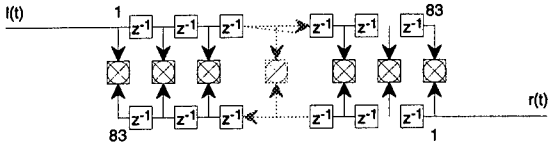


Figure 2. Coincidence detection for binaural sound source detection.

Representations of the two filtered binaural sound signals $l_{ae}^{(i)}(t)$ and $r_{ae}^{(i)}(t)$ of filter (i) for the determination azimuth or elevation angle are fed into a coincidence detector (see figure 2) which consists of single detectors and discrete delay blocks. The two corresponding signals are taken from the same filter range in order to determine the ITD according to Jeffress' model of binaural sound localization [11]. Each detector multiplies it's two input signals which are the delayed signals $l_{ae}^{(i)}(t)$ and $r_{ae}^{(i)}(t)$. If the sound is a simple click, the delay in the signal amplitudes of $l_{ae}^{(i)}(t)$ and $r_{ae}^{(i)}(t)$ determines the position. Thus, the activation of a particular detector represents the position. The relationship is non-linear. Each coincidence detection step uses a number of 83 data samples recorded with a sample frequency of 16.7 kHz. Assuming a total range of interest of 50 degrees of the angle, this leads to a theoretically possible resolution of 15 angular segments and 16 possible positions.

Instead of feeding the original signals to the coincidence detector, they are at first decomposed into onset signals $on\_l_{ae}^{(i)}(t)$ and $on\_r_{ae}^{(i)}(t)$ since they are usually by far more complex than pure clicks, containing speech, background noise, or echos. The onset signals are fed to the coincidence detector. In many practical cases this leads to a sharp peak of the correlation function once an onset signal enters the delay lines. When the onsets are too short, it is necessary to double one of the signals in order to guarantee matching inside the detector. With respect to a proposal in [16], the onset signals on_l/r(t) are calculated using a recursive envelop function en_l/r(t) of the signals l/r(t), yielding

$$en\_l/r(t) = \max \{\beta\ en\_l/r(t-1), abs(l/r(t))\}$$

$$on\_l/r(t) = \max \{0, en\_l/r(t) - en\_l/r(t-1)\}.$$

$\beta \in [0,1]$ is a heuristic parameter which has to be

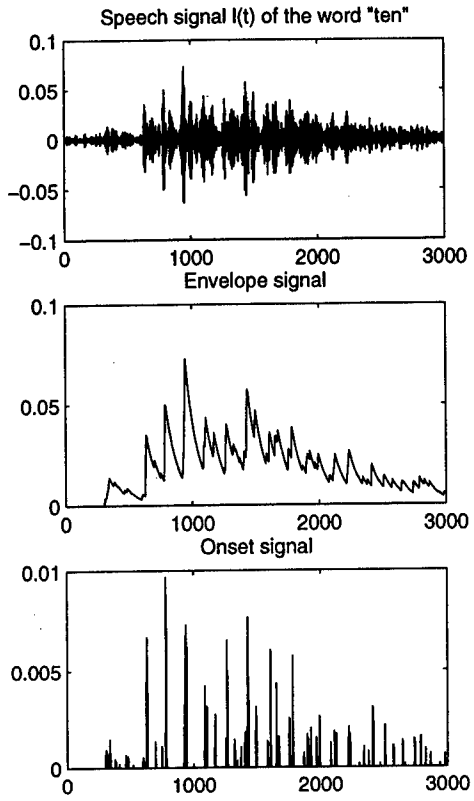


Figure 3. Top down: original speech signal l(t) of the word "ten", envelop signal en_l(t), onset signal on_l(t).

tuned according to the properties of the sound spectrum. For an increasing signal abs(l/r(t)), en_l/r(t) follows abs(l/r(t)), and in case of a decreasing signal abs(l/r(t)) it might soften the decrease with a speed that is depending on β. The onset signal on_l/r(t) is derived from the envelop en_l/r(t) by bounded differentiation.

If there is an echo in the signal, the amplitude will be smaller than the one of the direct signal. Thus, in many practical situations echo signals will be eliminated by decomposing the original signals into onset signals. The construction of the onset on_l/r(t) from the original signal l/r(t) is shown in figure 3.

The next processing unit in the auditory pathway is the colliculus complex. The topological maps in the colliculus complex are modelled by a set of homogeneous 16×16 matrices, each of which representing one frequency band of the cochlea. These maps are the basis for auditory-visual sensor fusion for fovea position control. The size of the maps corresponds to the desired precision of the attention control.

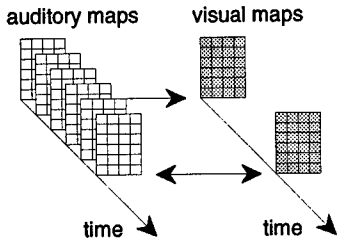


Figure 4. Time discretization of the sequences of auditory and visual topological maps.

The calibration function of the acoustic positioning is nonlinear with respect to the sound source position and the map representation when the visual image matrix is taken as a reference. This nonlinearity will be compensated by the map building process itself.

The map building generates a sequence of topological maps of size 16×16 with auditory estimates for the sound source localization. They are related to the set points of the maps of the visual estimates as shown in figure 4.

Coincidence detection for azimuth and elevation result in separate estimates for sound source location with a resolution of 83 possible positions, i.e. outputs of the detection cells. A simple way of combining both estimates to a twodimensional map ios to apply of any kind of 'AND' operation, for instance a simple multiplication. The resulting 83×83 map would be too large for direct match of the two modalities, and by means of the shifting procedure of the signals in the coincidence detector this would require the generation of a very large number of auditory maps between two subsequent visual maps. Another possibility to create a topological auditory map is to apply a one-dimensional fuzzy partition with 16 membership functions to the outputs of the coincidence detectors for azimuth and elevation, or a radial basis function network, respectively, and multiply the outputs of the 16+16 collecting receptive fields. Activation of the receptive fields should then be calculated by summing up the influences of the single onset signals, whereas the radial basis functions reflect the weight distribution for this weighted summation. The receptive fields act similar to the soma of a spiking neuron model in which the dendritic reactions on incoming spikes and spike trains are accumulated. Using a trainable neural network for map building

reduces the virtual reflections that occur when azimuth and elevation vectors contain more than one serious estimate for a sound source position and when they are simply correlated.

The implemented ANN is of MLP type with the output neurons ordered in a 16x16 array and no hidden layer. Each neuron can be interpreted as a twodimensional extension of an Adaline, performing a linear superposition of the input values. We used this approach as a first working solution since it is well known that Adalines are suited for echo cancellation.

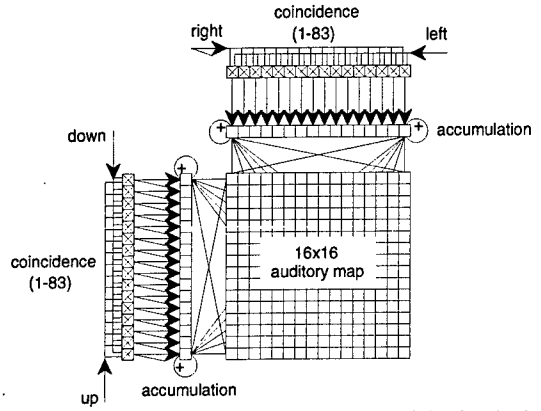


Figure 5. Auditory map building with the help of a neural network.

Some extensions concern the input vector configurations. A schematic diagram is shown in figure 5. The sample train of the two signals is devided into double sets of 83 subsequent samples, the complete discrete signals are fed into the coincidence detector, and the outputs of the detector cells are accumulated after each shift. The two resultung coincidence vectors are used as the input vectors for the MLP. Target output is the map with the neuron at the desired target position activated and all other neurons non-activated. The supervised training of the weights is performed by an improved back-propagation algorithm with momentum term and adaptive learning rate.

Besides the suppression of virtual echos, the ANN approach for auditory map building has the advantage of adapting the coordinate systems of the auditory and visual space to each other. Whereas the CCD camera recording of the images represent objects associated to planar surfaces in space, the cells of auditory maps are derived from a spherical coordinate system. The homomorphic mapping between them is non-

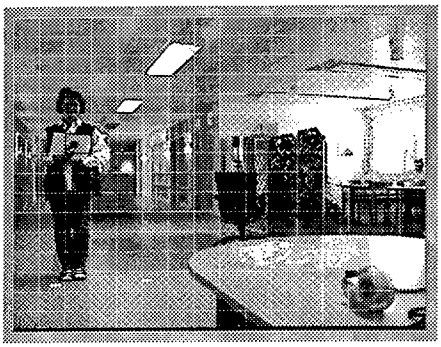


Figure 6. Calibration schene for auditory map building.

linear. In order to calibrate the auditory map building process and to match the two modalities, the auditory training data of the target positions are related to an equidistant grid in the respective images by eye-balling. A typical image of the calibration process is shown in figure 6. The training data set is composed by target positions on the visual grid, together with the respective microphone measurements. Figure 7 shows a typical predicted auditory map after training of the neural network, the target map, and the average distances of the prediction error for training data and a set of 50 untrained data.

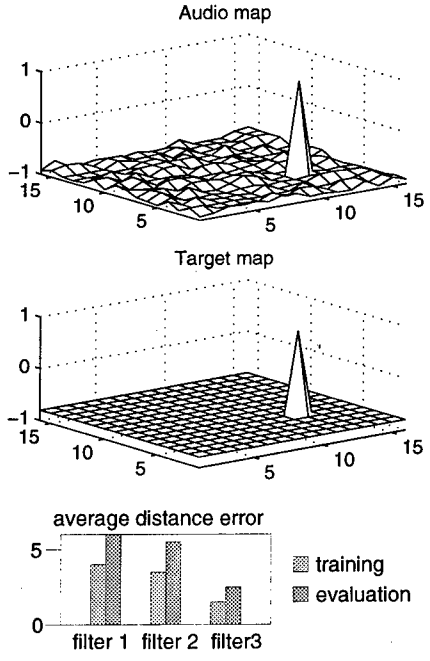


Figure 7. Exemplary auditory map after training, target map, and prediction errors for the filters.

High-level sound analysis algorithms with a similar functionality to those which are performed

in the auditory cortex were developed in previous projects [12], with special adaptation to speech recognition and speaker identification tasks. They are subject to later integration.

3 Monocular Vision

Biological systems use vision extensively. Properties of the visual pathway in animals have been studied extensively from the retina to the visual cortex, providing detailed knowledge about several subsequent steps of visual data processing. Main performances and properties are detection of edges, corner, motion, color, temporal and spatial novelty detection as well as complex visual feature binding, visual pattern recognition or identification, target tracking. Results of respective research work are, for instance, described in [13] or [14]. All but threedimensional vision tasks can be performed with one sensor only, whereas threedimensional vision requires binocular vision.

3.1 Biological Example

Visual sensory organs are the eyes. Visual stimuli or images are projected onto the retina and produce reactions in the receptor cells, the *rods* and the *cones*. They sense luminance and crominance of respective parts of the image. The output of neighboring receptor cells is collected and finally fed to ganglion cells which are located in the output layer of the retina. Due to individual synaptic connections, each ganglion cell reacts on particular spatial-temporal stimuli.

Around the optical axis is an area with a very high density of receptor cells and thus, a high spatial resolution and sensitivity, the *fovea*. Visual information received in the fovea is primarily fed towards the visual cortex in which complex recognition tasks can be performed.

A group of specific ganglion cells with large dendritic trees and therefore large receptive fields are connected to the colliculus complex, a reflex center which is generating nervous signals to initiate eye movements. These ganglion cells are particularly sensitive to spatio-temporal changes of the visual stimulus and serve for visual novelty detection and attention control. Once a visual novelty is detected, the eyes focus on it, allowing for more sophisticated recognition of visual objects which are projected onto the fovea.

The visual mapping from the retina to the colliculus is *retinotop*, i.e. retinal images are represented by respective spatial activations of neurons in the colliculus. A retinal image in this sense can be seen as the spatial-temporal derivative of the external visual stimulus.

3.2 Technical Implementation

The governing principles of biological visual information processing are used to find suitable neuromorphic abstractions and generalizations for the technical implementation. The retinal image as represented by the ganglion output signals is taken as a distinctive interface between visual sensing and low level image processing. A characteristic functionality of this interface is the processing of a spatio-temporal derivative of the image.

The temporal derivative is evaluated for our novelty detection paradigm, whereas the spatio-temporal derivative will be used in future high level visual recognition processes. For the novelty detection process, a low resolution is artificially introduced by averaging pixel values within a fixed neighborhood area. The original image matrix consists of 640×480 pixels. It is scaled down to 16×16 macropixels each of which integrating 40×30 pixels. At a certain time step (t), the grayscale value $mm_t(u,v)$ of a macropixel (u,v) is calculated as the normalized sum of the squared differences $d_t(i,j)$ of the pixel values $m_t(i,j)$ and $m_{t-1}(i,j)$ of two subsequent image matrices. The summation increases the robustness of the detection in the context of noise. Examplary subimages of two different visual output maps are shown in figure 8.

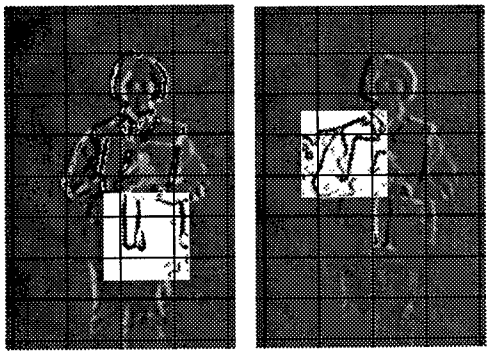


Figure 8. Fovea control according to the visual estimate, using maximum change of the macropixels.

In order to recognize the scenery, the grid of the macropixels is drawn in front of the corresponding two difference images, and the the edge detection is applied to the content of the fovea.

The visual topological maps in the colliculus complex are modelled by a set of homogeneous 16x16 matrices, resulting from the novelty detection in the video images. Together with the corresponding auditory maps, they are the basis for the auditory-visual sensor fusion for fovea position control. The relatively low resolution is postulated from the assumption that the area of the fovea shall be large enough to recognize desired objects or movements, and thus, a small mispositioning will not affect the recognition process significantly.

Instead of establishing a complex retinal model, we implemented a very simple and quick technical solution which is based on the SUSAN algorithm described in [15]. A local neighborhood of pixels, representing a receptive field, is shifted through the image, generating a new image. The number of pixels with greyscale values which do not differ too much from the one of the center pixel is accumulated, creating the new greyscale value of the center pixel. The mathematical procedure is a weighted summation. In areas of the image with a small spatial derivative within the range of the neighborhood the new pixel values reach a maximum, whereas close to corners or edges a significant decrease of the new value is realized. Edge and corner detection can thus be performed by simple thresholding. The temporal derivative is derived by processing subsequent difference images which are associated to a temporal grid of 40 ms according to the PAL video norm.

First algorithms that represent these cortical visual information processing and which will be implemented on base of the existing system are the spatial location of faces (e.g., [16], [17]) based on an evaluation of color information or on morphological operations, the recognition of faces (e.g., [18]) for supervisor identification, reading lip information (e.g., [19]) in order to improve speech recognition in noisy or distorted environment, or recognition of hand gestures (e.g., [20]) with the help of a codebook in order to give sign language commands. As an extension towards auditory-visual human-machine communication, a technical system is introduced in [12] which describes the conversion of auditory in visual speech signals by converting acoustic cues into naturally looking movements of an animated human face. All these high level processing units are meant to improve human-machine communication significantly.

4 Auditory-visual Sensor Fusion

4.1 Biological Example

As can be seen by now, the colliculus complex collects auditory and visual novelty information in form of topological representations of neuron activation. For this reason, the colliculus is assumed to be a primary center for auditory-visual sensor fusion as far as novelty detection is concerned. The fusion of this bimodal information is performed by the specific synaptic connections of the neurons in the processing layers of the colliculus which perform basically a nonlinear template matching between the auditory and visual maps.

4.2 Technical Implementation

In biological examples the auditory maps representing spatial information in the different frequency bands seem to be already fused into a general representation before the information reaches the colliculus. Our technical abstraction of the fusion process ignores the importance of this fact and therefore consists of three major steps, i) the fusion of a set of subsequent auditory maps belonging to the same interval of two subsequent visual maps, ii) the fusion of the auditory maps in different frequency ranges, and iii) the fusion of the resulting auditory with the visual map.

The idea of locally distributed topological representations of the perceptional space cue a decomposition of the localization problem by means of separate auditory and visual estimates, combined with a subsequent bimodal sensor fusion process. In technical terms this leads to particular fusion paradigms of sensory information on feature level (see [21]). Exemplary work on fusion of passive sound and vision for emulating human perception can be found in [16] or [22].

The three steps in sensor fusion are performed in the following way: At first, the auditory maps belonging to one video interval are fused by calculating averages for each position. This is necessary to take the unknown delay between audio and video recording into account. Then, the maps for the three filters are merged by calculating the average map. For the auditory-visual fusion, each position is virtually widened with the help of a Gaussian function, and the the intersection is taken as a first estimate. The finalestimate of the position is calculated by doing so for all positions of the map and averaging the result. Then, the fovea position is calculated as a quadratic prediction, using also the last two auditory-visual estimates.

References

[1] E. I. Knudsen (1982): Auditory and visual maps of space in the optic tectum of the owl, J. Neuroscience 2, pp. 1177-1194.

[2] E. I. Knudsen, S. du Lac, S. D. Esterly (1987): Computational maps in the brain, J. Neuroscience 10, pp. 41-65.

[3] B. E. Stein, M. A. Meredith: (1993): The Merging of Senses, MIT Press, Cambridge.

[4] E. I. Knudsen, M. S. Brainard (1995): Creating a unified representation of visual and auditory space in brain, Ann. Rev. Neuroscience 18, p. 19.

[5] Y. Dan, J. J. Atick, R. C. Reid (1996): Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory, J. Neuroscience 16, pp. 3351-3362.

[6] C.H. Keller, K. Hartung, T. Takahashi (1998): Head-related transfer functions of the barn owl: measurement and neural responses, J. Neuroscience 118, pp. 13-34.

[7] L. Heimer (1995): The Human Brain and Spinal Cord. Springer-Verlag, New-York.

[8] J. Blauert (1997): Spatial Hearing: The Psychophysics of Human Sound Localization. The MIT Press, Cambridge.

[9] J.B. Allen (1985): Cochlear modelling, IEEE ASSP Magazine 2, pp. 3-29.

[10] R.F. Lyon, C. Mead (1988): An analog electronic cochlea, Proc. IEEE Trans. Acoustics, Speech, and Signal Processing 36, pp. 1119-1134.

[11] L.A. Jeffress (1948): A place theory of sound localization, J. Comp. Physiol. Psych. 61, pp. 468-486.

[12] H.-H. Bothe (1997): Audio to audio-video speech conversion with the help of phonetic knowledge integration, Proc. 1997 IEEE Int. Conf. on Systems, Man, and Cybernetics, Oct. 1997, Orlando, FL, pp. 848-851.

[13] D.H. Ballard, C.M. Brown (1992): Principles of animate vision, Computer Vision Graphics and Image Processing 56:1, pp. 3-21.

[14] V. Bruce, P.R. Green, M.A. Georgeson (1996): Visual Perception, Physiology, Psychology and Ecology (3rd ed.). Psych. Press, London.

[15] S.M. Smith, J.M. Brady (1995): SUSAN, a new approach to low level image processing, Technical Report TR95SMS1c, FMRIB Oxford University and SW.

[16] S. G. Goodridge (1997): Multimedia Sensor Fusion for Intelligent Camera Control and Human-Computer Interaction, Dissertation, North Carolina State University, Raleigh, NC.

[17] H.P. Graf, T. Chen, E. Petajan, E. Cosatto (1995): Locating faces and facial parts, Int. Workshop Automatic Face and Gesture Recognition, Zurich/CH, June 1995, pp. 41-46.

[18] M. Bichsel (1995): Human face recognition: From views to models - from models to views, Int. Workshop Automatic Face- and Gesture-Recognition, Zurich/CH, June 1995, pp. 59-64.

[19] N. Li, S. Dettmer, M. Shah (1995): Lipreading using eigensequences, Int. Workshop Automatic Face- and Gesture-Recognition, Zurich/CH, June 1995, pp. 30-34.

[20] T.S. Huang (1995): Hand gesture modelling, analysis, and synthesis, Int. Workshop Automatic Face- and Gesture- Recognition, Zurich/CH, June 1995, pp. 73-79.

[21] L. Hall, J. Linas (1997): An introduction to multisensor data fusion, Proc. IEEE 85:6.

[22] U. Bub, M. Hunke, A. Waibel (1995): Knowing who to listen to in speech recognition: visually guided beamforming, Proc. 1995 IEEE Int.l Conf. on Acoustics, Speech, and Signal Processing, May 1995, Detroit, MI, pp. 848-851.

A Novel Minimal Norm Based Learning Subspace Method①

De-Shuang Huang

(Beijing Institute of System Engineering, P. O. Box 9702-19, Beijing 100101, China
Tel: 86-010-66356616, Fax: 86-010-64879213, email: bisenet@bisel.bise.ac.cn)

Abstract: *This paper proposes a novel* minimal norm based learning subspace method (*MNLSM*), *which can satisfy the requirements of being insensitive to the order of presentation of the training samples. This MNLSM is applied to recognition of simulating high resolution radar (HRR) targets (two for ships, one for chaff). Experimental results show that the performance of proposed MNLSM such as rate of correct recognition and convergence speed is satisfactory.*

Keywords: Self-Supervised Learning, Subspace, Pattern Recognition, Minimal Norm, Disordered Learning, Radar Targets.

1. Introduction

The *learning subspace method* (LSM) proposed by Kohonen in 1978[1], in essence, is an adaptive method of extracting principal components of pattern vectors from each class. This approach assumes the class labels for all input samples to be known, and uses Hebbian rule to update the basis vectors corresponding to each subspace. So it is also called the self-supervised neural network approach [2], which designs each subspace in terms of the label for each sample. However, the essential drawbacks to the LSMs are sensitive to the order of presentation of the input samples, in other words, the prior learned samples which might be recorded in the basis vectors of the corresponding subspace may be offset or forgotten by the learning of the late-coming samples, which leads to total performance decreasing [2, 3, 4, 6]. In 1982, E. Oja et. al proposed the *averaging learning subspace method* (ALSM)[3,4]which can avoid the sensitiveness to the order of presentation of the input samples. But it needs to compute three conditioned correlation matrices and their eigenvalue decompositions which leads to the convergence speed much decreasing[5]. To avoid or reduce the defects for those existing methods, this paper proposes a novel self-supervised learning subspace methods, called *minimal norm based learning subspace method* (MNLSM), which are not sensitive to the order of presentation of the input samples, and much improve the convergence speed.

This new LSM, to verify its validity, is applied to the recognition of high resolution radar (HRR) targets (two for simulating ships and one for simulating chaff). The experimental results support our claims.

2. A Novel Minimal Norm Based Learning Subspace Method

2.1 General Presentation

① This work was supported by Grants 69705001 from NSF and DSR of China

suboptimal) solutions. The training sample for each iteration needs to be selected not only "randonly" but also some specific criterion based from the training sample set. In fact, it is easily thought that for some class, the training sample with the minimal orthogonal projection length on its own subspace can be selected to design corresponding subspaces. Thus this method is called as *minimal norm based learning subspace method* (MNLSM). Fig. 1 depicts the scheme of the learning process for MNLSM.

2.3 Minimal Norm Based Learning Subspace Iterating Algorithm

Firstly, assume that, at the kth iteration, the sample with minimal norm from the$i$th subspace is selected as

$$\mathbf{x}_m^{(i)} = \arg\min\{\delta(\mathbf{x}_j^{(i)}) = (\mathbf{x}_j^{(i)T} P^{(i)} \mathbf{x}_j^{(i)})^{\frac{1}{2}} , j = 1,2,\cdots,N_i\} \quad (4)$$

where $\arg\min\{\cdot\}$ denotes the operator of selecting the training sample with minimal orthogonal projection length on its own subspace, $\delta(\mathbf{x}_j^{(i)})$ the orthogonal projection length of $\mathbf{x}_j^{(i)}$ on the ith subspace.

So, the sample with minimal norm $\mathbf{x}_m^{(i)}$ is used to learn its own subspace with the positive manner and learn other subspaces with the negative manner. According to Eq. (1), the iterating formulae for the MNLSM can be stated as follows

$$L_k^{(i)} = (I + \mu_m^{(i,i)} \mathbf{x}_m^{(i)} \mathbf{x}_m^{(i)T}) L_{k-1}^{(i)}, i = 1,2,\cdots,K \quad (5)$$

$$L_k^{(j)} = (I - \mu_m^{(j,i)} \mathbf{x}_m^{(i)} \mathbf{x}_m^{(i)T}) L_{k-1}^{(j)}, \quad j \neq i \quad (6)$$

Generally, the above learning process could be unlimitedly gone on, but after several iterations, the formed subspace might become stable. The iterating algorithm for the MNLSM is summarized as follows:

Algorithm **Minimal Norm Based Learning Subspace Iterating Algorithm**

Step 1 $k = 1$, select the dimensionality $p^{(i)}$, the termination accuracy η and learning coefficient $\mu^{(i)}(i = 1,2,\cdots,c)$, set the initial basis vectors of the c subspaces, and compute the orthogonal projection matrixes $P_i^{(i)}(i = 1,2,\cdots,c)$.

Step 2 for each pattern vector $\mathbf{x}_j^{(i)}$ of the ith class, compute its orthogonal projection length (norm) on its own subspace

$$\delta(\mathbf{x}_j^{(i)}) = (\mathbf{x}_j^{(i)T} P^{(i)} \mathbf{x}_j^{(i)})^{\frac{1}{2}} , i = 1,2,\cdots,c; j = 1,2,\cdots,N_i\}$$

Step 3 select the training sample with minimal norm from the training sample set of each class

$$\mathbf{x}_m^{(i)} = \arg\min\{\delta(\mathbf{x}_j^{(i)}) , i = 1,2,\cdots,c; j = 1,2,\cdots,N_i\}$$

Step 4 rotate its own subspace with the positive manner in terms of Eq. (5) and rotate other subspaces with the negative manner in terms of Eq. (6) using the training sample with minimal norm.

Step 5 compute the averaging orthogonal projection length, T_d, of all training samples with minimal norms $\mathbf{x}_m^{(i)}$ from the ith class on their own subspaces according to Eq. (3)

Step 6 if $T_d \geqslant \eta$, skip to **Step 8**; else, continue to **Step 7**.

Step 7 k=k+1, return to **Step 2**.

Step 8 stop.

Note that whole iteratinng process will be terminated until the termination accuracy η is satisfied.

3. Experimental Results

The simulating range profiles of radar tar-

Assume that the ith class ω_i has N_i pattern vectors $\{\mathbf{x}_k^{(i)} \in R^d,\ k = 1, 2, \cdots, N_i;\ i = 1,2, \cdots, c)$, respectively. The self-supervised learning subspace method proposed by Kohonen is stated as follows[1]

$$\begin{cases} L_k^{(i)} = (I + \mu_k^{(i)} \mathbf{x}_k^{(i)} \mathbf{x}_k^{(i)T}) L_{k-1}^{(i)} \\ L_k^{(j)} = (I - \mu_k^{(j)} \mathbf{x}_k^{(i)} \mathbf{x}_k^{(i)T}) L_{k-1}^{(j)} \\ (j \neq i = 1, 2, \cdots, c) \\ L_k^{(i)} = L[\mathbf{u}_1^{(i)}(k), \mathbf{u}_2^{(i)}(k), \cdots, \mathbf{u}_{p^{(i)}}^{(i)}(k)] \end{cases} \tag{1}$$

where $\mu_k^{(i)}$, $\mu_k^{(j)}$ are the positive learning coefficients. Generally, $|\mu_k^{(i)}| < 1/\|\mathbf{x}\|_k^2$ and $|\mu_k^{(j)}| < 1/\|\mathbf{x}\|_k^2$); T denotes matrix or vector transpose; $\mathbf{x}_k^{(i)} = [x_1^{(i)}(k), x_2^{(i)}(k), \cdots, x_d^{(i)}(k)]^T$; I is an unit matrix; $L_k^{(i)} = L[\cdot]$ indicates the ith subspace composed of $p^{(i)}$ basis vectors $\mathbf{u}_n^{(i)}(k)$ $(n = 1, 2, \cdots, p^{(i)})$ at instant k.

It should be, however, noted that the basis vectors $\mathbf{u}_n^{(i)}(k)$ $(n = 1, 2, \cdots, p^{(i)})$ should be kept orthonormal in the learning process of the LSM. Usually, the orthonormal approach available is the Gram-Schmidt one. Supposed that the converged orthogonal projection matrices corresponding to c subspaces are $P^{(i)}(i = 1,2, \cdots, c)$, respectively. For an arbitrary input vector $\mathbf{x}$, the classification rule of the self-supervised LSM for pattern recognition is that if[1]

$$\|P^{(i)}\mathbf{x}\|^2 = \mathbf{x}^T P^{(i)} \mathbf{x} = \sum_{j=1}^{p^{(i)}} (\mathbf{x}^T \mathbf{u}_j^{(i)})^2 = \max_j \|P^{(i)}\mathbf{x}\|^2 \tag{2}$$

classify $\mathbf{x}$ in class i, i.e. $\mathbf{x} \in \omega_i$.

Assume that the confidence coefficient for stopped iterating of subspaces is η $(0.5 \leqslant \eta \leqslant 1)$, if the averaging orthogonal projection of an arbitrary pattern vector $\mathbf{x}$ on the c subspaces, at the kth iterating

$$T_k = \frac{1}{c} \sum_{i=1}^{c} \mathbf{x}^{(i)T} P_k^{(i)} \mathbf{x}^{(i)} \tag{3}$$

satisfies $T_k \geqslant \eta$, c subspaces are thought to have converged to the given accuracy, the iterating will be stopped[2].

2.2 Basic Idea of Minimal Norm Based Learning Subspace Iterating Algorithm

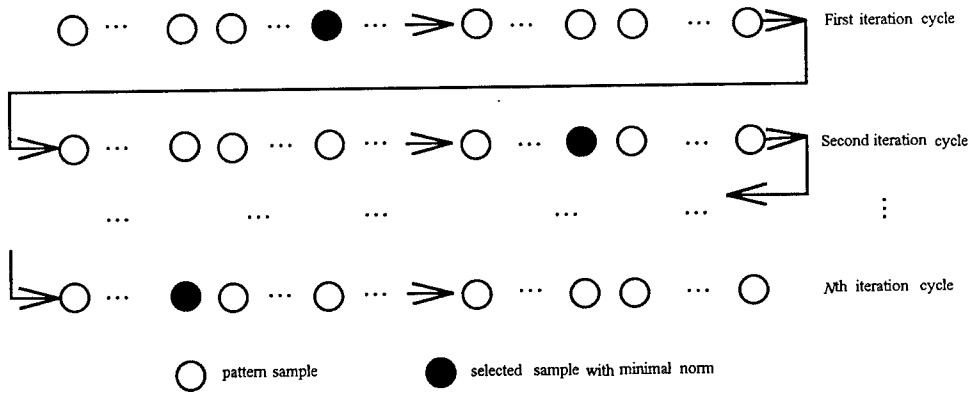


Fig 1. Scheme of the learning process for MNLSM

To avoid the effect of the order of presentation of the input samples on the classification performance, the training sample for each iteration can be selected "randonly" from training sample set. However, in doing so, it is very difficult for the learning subspace to converge to the optimal (or

gets (ship 1, ship 2 and chaff) are used as the experimental data that will be classified by the self-supervised LSM. Assume that the resolution of radar is $\triangle R = 7.5m$, the radar echoes of multiple resolution cells of targets relative to radar in the range of 3000 m to 3480 m are measured, where the azimuths of targets are changed at intervals of $\triangle\theta = 0.5°$. As a result, for each class 50 range profiles whose dimensionality is 64 are obtained. In addition, in experiment the 50 range profiles in the time-domain are transformed to those in the frequency-domain by Fourier transform. These transformed experimental data are also used to train corresponding subspaces.

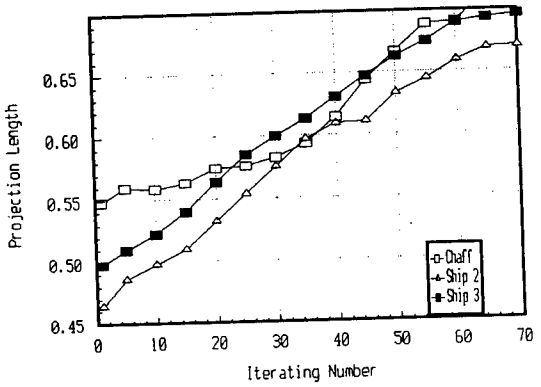


Fig 2. Dynamic learning processes of three training samples with minimal norms respectively from three classes for MNLSM classifier

Assume that the numbers of strong scatter centres of ship 1 and ship 2 are 9 and 7, respectively. In the light of the selection method of dimensionality of a subspace dicussed in [2], the dimensionalities of the subspaces corresponding to ship 1, ship 2 and chaff are selected as 10, 8 and 2, respectively. Before experiments, let all training sample vectors be normalized to unit vectors. Regardless of the properties of individual training sample vector, assume that the learning coefficient μ is selected as follows[2]

$$\mu(\mathbf{x}) = \alpha\sqrt{1 - \|\hat{\mathbf{x}}\|^2} \qquad (7)$$

where $\hat{\mathbf{x}}$ is the orthogonal projection vector of the training sample vector $\mathbf{x}$ on its own subspace; α is an adjusting coefficient. Generally, $0 < \alpha < 1$. In the practical training process $\mu_m^{(j,i)}$ is assumed to be $\mu(\mathbf{x}_m^{(i)})$, i.e., $\mu_m^{(i,i)}$.

In simulation, the obtained 50 sample vectors in the two domains for each class are divided into training sets which consist of 25 odd numbered samples and testing sets which consist of 25 even numbered samples. Assume that for each class one sample randomly selected from the 25 training samples is used to design the initial subspace. Fig. 2 shows the dynamic learning processes of three training samples with minimal norms from three classes, respectively. Table 1 gives the testing recognition results of testing samples from three classes in the two domains with the iterative number.

Table 1 The testing recognition results of testing samples from three classes in the two domains with the iterative number for MNLSM classifier

Iterative Number		10	20	30	40	50	60	70	80	90	100
Time domain	chaff	23.4%	34.2%	45.5%	56.7%	61.2%	68.7%	75.1%	81.2%	81.6%	82.1%
	ship 1	19.5%	31.5%	37.3%	45.6%	59.3%	66.4%	73.2%	76.6%	78.5%	79.2%
	ship 2	23.3%	33.7%	43.4%	49.7%	61.8%	65.2%	70.8%	73.4%	76.9%	78.4%
Frequency domain	chaff	12.5%	20.4%	26.7%	34.6%	37.5%	41.3%	48.9%	54.2%	58.5%	61.5%
	ship 1	8.6%	15.6%	21.4%	30.4%	36.6%	42.3%	51.5%	58.6%	61.3%	63.4%
	ship 2	6.7%	13.2%	21.2%	26.2%	31.5%	35.5%	40.2%	48.1%	52.6%	57.2%

In addition, assume that three kinds of self-supervised LSMs, i. e., LSM, ALSM and

MNLSM, are used to classify the simulating range profiles (time and frequency domains) of three classes. After Eq. (3) is satisfied ($\eta = 0.8$), the 25 testing samples in the two domains are used to test the rate of correct recongnition, the classification results are shown in Table 2.

Table 2 Comparisons of correct recognition rates of testing samples from three classes in the two domains using the three classification methods of LSM, ALSM and MNLSM

Classification Method		LSM	ALSM	MNLSM
Rate of Recognition (Time-Domain)	chaff	79. 8%	83. 4%	82. 8%
	ship 1	78. 9%	84. 3%	83. 7%
	ship 2	78. 3%	83. 7%	82. 4%
Rate of Recognition (Frequency-Domain)	chaff	60. 4%	64. 2%	64. 3%
	ship 1	61. 2%	63. 9%	64. 1%
	ship 2	60. 6%	64. 7%	63. 9%

From the above experiments, it can be found that the MNLSM possesses the better performance whether the convergence speed or the correct recognition rate.

4. Conclusion

Learning subspace method (LSM) for pattern recognition is one of efficient self-supervised learning neural network classifiers. This paper, based on the LSMs proposed by Kohonen, proposed a novel self-supervised LSM with higher correct classification rate and less computation time, i. e., *minimal norm based learning subspace method* (MNLSM), which is not sensitive to the order of presentation of the input samples. To verify the validities for this method, this paper discussed applying this method to recognition of simulating high resolution radar (HRR) targets. Experimental results show that the performance of proposed MNLSM such as rate of correct recognition and convergence speed is satisfactory.

Reference

[1] T. Kohonen., Self-Organization and Associative Memory. Springer-Verlag Berlin Heidelberg New York, 1989.

[2] D. S. Huang, The Study of Recognition Technique Based on One-Dimensional Image of Radar Targets. Ph. D dissertation, Xidian University, 1992.

[3] M. Kuusela, et al., "The averaging learning subspace method for spectral pattern recognition," Proc. of 6th IC on PR, Munich, pp. 134-137, 1982.

[4] E. Oja, et al., "The ALSM algorithm -an improved subspace method of classification," Pattern Recognition, PR-16, no. 4, pp. 421-427, 1983.

[5] E. Oja, Subspace Methods Of Pattern Recognition. Research Studies Press Ltd. 1983.

[6] D. L. Mensa, High Resolution Radar Imaging. Artech House, Inc., 1981.

Integration of Human Knowledge and Sensor Fusion for Machining

T.I. van Niekerk J. Huang

Port Elizabeth Technikon
Private Bagx6011 Port Elizabeth 6000
South Africa
Email: theoian@ml.petech.ac.za

Z. Katz

Rand Afrikaans University
P.O. Box 524 Auckland Park 2006
South Africa

Abstract

Machining is a complex process influenced by cutting parameters (feed, speed and depth of cut) and process conditions (tool wear, material properties, coolant and workpiece dimensions). This paper proposed a methodology of integrating human knowledge and sensor fusion for machining monitoring and control. Fuzzy logic is applied to representation of human knowledge in machining. Appropriate machining parameters are determined by using fuzzy models. Sensor fusion is made for capturing sensor features that reflect machining performance by means of DSP technology. Human knowledge and sensor data are further integrated in a neural network structure for monitoring machining quality and cutting tool state.

Key Words: machining, sensor fusion, human knowledge, neural network, fuzzy logic

1. Introduction

Machining is a complex process influenced by cutting parameters (feed, speed and depth of cut) and process conditions (tool wear, material properties, coolant and workpiece dimensions). Fuzzy logic and neural network approaches are widely applied in machining analysis and modeling. EL Baradie [1] proposed a fuzzy logic model for machining data selection. In their model, the relation of cutting speed and material hardness is built by fuzzy logic in terms of information from the machining data handbook. Fang [2] built an expert system to support fuzzy diagnosis of process states in finish-turning. Machining states are described by fuzzy features and the fuzzy feature-state matrices are established through the expert system. Ko and Cho [3] applied a neural network to cutting state monitoring in face milling. In their system, an autoregressive (AR) time series modeling is used as preprocessor for generating features from each sensor, and then a highly parallel neural network is set up for associating the preprocessor outputs with the appropriate decisions with which the cutting state is classified. Purushothaman and Srinivasa [4] studied the tool wear monitoring in turning by a back propagation algorithm. The tool condition is classified into two states: class 1 acceptable and class two unacceptable. Different patterns under different cutting conditions are studied. Using neural network for tool condition monitoring in metal cutting is reviewed by Dimla Jr, Lister and Leighton [5]. The review illustrated the extent of application of neural networks, the methods for handling sensor signals within a neural network, and the need for sensor fusion from multiple source sensor signals in tool condition monitoring.

In our system sensor fusion from multi-sensors (force, vibration, driving motor current, etc.) is realized by utilizing digital signal processing technology to perform on-line data acquisition, and determine monitoring indices (mean resultant force, power spectral densities, etc.). Neural networks are subsequently used to perform relating monitoring process states and fuzzy logic is used to determine state variables from monitoring indices.

To determine and set optimal control signals depends not only information on machining process but also on human knowledge on machining, for example, selection of machining parameters, abilities and conditions of machine tools and types of cutting tools. Sensor fusion from multi-sensors can only collect information related to machining process. This work develops a system to fusion sensory data and human knowledge for machining control, interacting with people. The system consists of four major parts, namely determination of machining parameters, sensor data acquisition and processing, machining state monitoring, and machining process control.

Determination of machining parameters and machine tool conditions is largely related to human knowledge on machining. In the process of determination, machining requirements of dimension, tolerance and surface roughness are set as quality control objective, while machine tool and cutting tool capacity and

conditions as well as tool life as constraints. Machining information is interpreted using fuzzy representation that is taken as input while machining parameters as output in a fuzzy decision making system. Human knowledge on machining is integrated into the fuzzy decision process.

The sensory data are collected and processed through DSP technology. The data processing of filtering, windowing, FFT and power spectral analysis is done on-line.

Multi-sensor data and signal features are used for building sensor fusion with neural networks. Machining parameters are further taken into account in the parallel neural networks and machining state fuzzy classification. In the system, sensor data such as force and control motor current are combined with those parameters (feed, speed and depth of cut) from the fuzzy decision system as input, and machining quality and cutting tool state respectively as output in neural network learning. In order to efficiently monitor machining process, two parallel neural networks are designed in the system with different input and output. One network is for machining quality (dimensional deviation and surface roughness) and the other for cutting tool state (tool wear state). Machining state is represented and classified according to the machining quality and tool wear state using fuzzy classification. Machining monitoring is realized in process in terms of sensor fusion and machining state specified.

2. Machining Knowledge

Machining process involves a variety of knowledge which may cover the following aspects:

- Knowledge of materials.
- Knowledge of parts to be machined.
- Knowledge of machine tools, cutting tools and fixtures.
- Knowledge of cutting liquids.
- Knowledge of manufacturing processes.

Material properties include tensile and compressive strength, hardness, endurance limit, modulus of elasticity and torsion modulus, as well as machinability, thermal properties and wear resistance etc. Machinability is a combination property of materials and directly related to machining processes. Material hardness is regarded as factor for selection of machining parameters.

Regarding a part to be machined, information of the geometry, dimension, tolerance and surface roughness as well as possible heat treatment of the part is required. The shape features of geometry with dimension are achieved by selecting different machining processes. The tolerance and surface roughness is achieved by properly planning machining process and selection of machining parameters.

Manufacturing process requires planning and arrangement of machine tools, cutting tools and fixtures. Tool wear and life is a major concern in machining. It greatly affects machining quality and manufacturing cost. Cutting liquids are necessary for most continuous machining in cutting hard materials. Cutting zone temperature and chips can be dispersed effectively by using coolants.

Although there are experimental data and machining data book available as reference, selection of machining parameters and planning machining process still need human operators' intervention using their empirical knowledge. To achieve automation of machining planning and on-line monitoring and control of machining process, it is necessary to realize integration of human knowledge with machining process by utilizing fuzzy logic and neural network or other expert technologies.

3. Fuzzy logic modeling

Many relations in machining may not be precisely represented by using traditional mathematical modeling equations. For example, the relation of material hardness and cutting speed may be generally described as that if the material hardness is high then the corresponding cutting speed is low and vice versa. The tool wear may be described as slight wear, medium wear and severe wear. Those relations can be effectively described by fuzzy logic modeling. In our system developed, two procedures are taken in modeling. The first procedure is to select appropriate machining parameters by fuzzy decision making with respect to machining requirements and conditions. The second procedure is to integrate selected parameters and sensory data into multi-layer perceptron (MLP) for cutting tool states and machining states monitoring and control. The tool and machining states are described by fuzzy representation.

Determination of machining parameters is made with a process of multi-level fuzzy decision making. Generally speaking, machining parameters selection relates to machining requirements of parts to be machined, machine tool capacity and cutting tool life. There are three steps taken into account in modeling for determination of machining parameters as follows.

(a) Select depth of cut in terms of rough machining, half finish and finish machining with respect to different materials and material hardness. Harder a material is, more force it requires in machining. Thus more power consumption is taken, obviously which is limited by machine tool capacity. We have general rules as

IF it is rough machining, THEN the depth of cut is taken deep.
IF it is half finish machining, THEN the depth of cut is taken medium.
IF it is finish machining, THEN the depth of cut is taken low. (1)

Furthermore, taking material hardness into consideration, there are rules

IF the material hardness is very hard, THEN the depth of cut is taken very low.
IF the material hardness is hard, THEN the depth of cut is taken low.
IF the material hardness is medium, THEN the depth of cut is taken medium.
IF the material hardness is soft, THEN the depth of cut is taken deep.
IF the material hardness is very soft, THEN the depth of cut is taken very deep. (2)

(b) Depth of cut and feed determine the size of cut, i.e. the volume of material to be removed in machining. The use of large feed and depth can efficiently increase the material removal rate. And if combined the large feed and depth with low cutting speed, a long tool life can be obtained. However, although a large feed and depth are beneficial to tool life and efficient material removal, the maximum size of cut is limited by a number of factors which can not be ignored for achieving a high quality machining. These factors are:

- the maximum power available from the machine tool;
- the maximum forces that the cutter can stand;
- the maximum permissible deflection of the machine tool and work consistent with the accuracy required;
- the tendency to chatter;
- the fact that surface finish grows worse as size of cut is increased. It is verified that feed has much higher effect on surface roughness than does the depth.

In recent years the trend of high speed machining is progressive. With consideration of the factors above, the machine tool power, cutter and work rigidity, and surface roughness are taken as variables for selection of feed. For rough machining, there are rules as

IF the cutter and work rigidity is very high, THEN the feed is very high.
IF the cutter and work rigidity is high, THEN the feed is high.
IF the cutter and work rigidity is medium, THEN the feed is medium.
IF the cutter and work rigidity is low, THEN the feed is low.
IF the cutter and work rigidity is very low, THEN the feed is very low. (3)

The cutter and work rigidity is roughly measured according to cutter (shank) and work (diameter) sizes. For half finish and finish machining, there are rules as

IF the surface quality is very high, THEN the feed is very low.
IF the surface quality is high, THEN the feed is low.
IF the surface quality is medium, THEN the feed is medium.
IF the surface quality is low, THEN the feed is high.
IF the surface quality is very low, THEN the feed is very high. (4)

The power consumption will be checked after the machining parameters are determined.

(c) Cutting speed is a major factor that effects on machining quality (surface finish) and tool life. Ordinarily, surface finish improves with increase of cutting speed. The change is made up to certain point when the speed arrives at some critical value due to a continuous reduction in size of the build-up edge. After the build-up edge has become insignificant in size, little further improvement on surface finish is made with increase in cutting speed. The relation of tool life and cutting speed is illustrated by Taylor empirical equation

$$VT^n = C_t \tag{5}$$

where V is cutting speed. T indicates the cutting time. Ct is a constant whose value depends on machining conditions. It is numerically equal to the cutting speed that gives a tool life of 1 min. n is exponent whose value varies to some extent with machining conditions. Generally, their relation can be stated that if the cutting speed is increased, the tool life will decrease. So it is necessary to select a cutting speed satisfying a desired tool life given by equation (5).

With desired tool life, we use surface roughness and material hardness as factors to select cutting speed. The relationship of cutting speed and surface quality may be described as the following rules

IF the surface quality is very high, THEN the cutting speed is very high.
IF the surface quality is high, THEN the cutting speed is high.
IF the surface quality is medium, THEN the cutting speed is medium.
IF the surface quality is low, THEN the cutting speed is low.
IF the surface quality is very low, THEN the cutting speed is very low. (6)

The relationship of cutting speed and material hardness may be given as

IF the material hardness is very hard, THEN the cutting speed is very low.
IF the material hardness is hard, THEN the cutting speed is low.
IF the material hardness is medium, THEN the cutting speed is medium.
IF the material hardness is soft, THEN the cutting speed is high.
IF the material hardness is very soft, THEN the cutting speed is very high. (7)

The above relations represent human experimental and empirical knowledge on machining conditions and parameters selection. To give an example, a fuzzy model for the relations of (6) and (7) may be established for describing these relations as

$$B = A \circ R \tag{8}$$

where B denotes the speed fuzzy set, A either the surface quality or the material hardness fuzzy set and R is fuzzy relation of A and B. The symbol "o" denotes fuzzy compositional operator. Fuzzy sets A and B may be represented by triangular shape membership functions (u_A, u_R) shown in Figure 1.

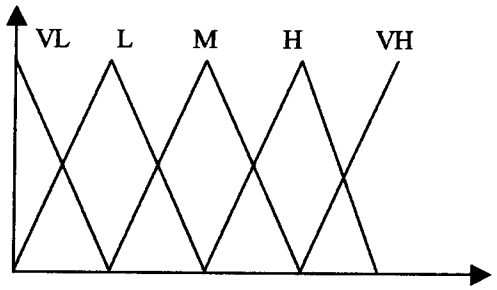


Figure 1. Fuzzy membership function

The VL, L, M, H and VH denote the linguistic states of the cutting speed, surface quality or material hardness, i.e. very low, low, medium, high and very high respectively. The fuzzy relation R can be determined from the relations described in (6) and (7). For example, the rule

IF the material hardness is very hard, THEN the cutting speed is very low

may establish a relation as

$$u_{R1} = u_A(VH) \wedge u_B(VL) \tag{9}$$

where $u_A(VH)$ denotes the grade of very hard in terms of the hardness membership function, $u_B(VL)$ the grade of very low in terms of the speed membership function, refer to Figure 1. u_{R1} describes the fuzzy relation matrix of A and B. In the same way other four rules in (7) may be expressed as

$$\begin{aligned} u_{R2} &= u_A(H) \wedge u_B(L) \\ u_{R3} &= u_A(M) \wedge u_B(M) \\ u_{R4} &= u_A(S) \wedge u_B(H) \\ u_{R5} &= u_A(VS) \wedge u_B(VH) \end{aligned} \tag{10}$$

Finally the five rules are combined together using fuzzy OR operator. It may be expressed as

$$u_R = \bigvee_{i=1}^{5} u_{Ri} \tag{11}$$

Similarly, from the rules (6) the fuzzy relation of the cutting speed and surface quality can also be established. The speed inferred from the relation (6) and (7) in terms of equation (8) may not be same. Some decision making approaches may be applied [1].

4. Knowledge integration in machining

In order to on-line monitoring machining process, a model is designed for integration of human knowledge and sensory data shown in Figure 2.

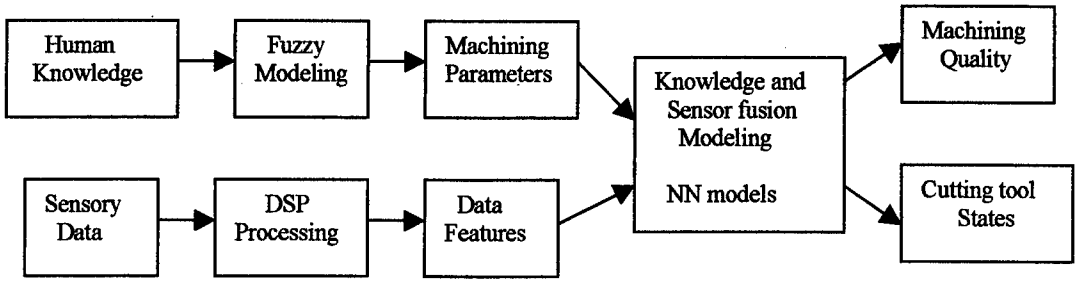


Figure 2. Modeling of integration of knowledge and sensor fusion

Human knowledge on machining is represented by fuzzy sets and integrated with fuzzy models. With fuzzy modeling human experimental and empirical knowledge and linguistic inference can be properly handled. Appropriate machining parameters selected in terms of fuzzy modeling are taken as a part of input information for machining monitoring and control. On the other hand, sensors are implemented with machining systems for capturing machining performance. Machining performance is obviously related to the machining parameters selected and also affected by other factors from machine tool, cutting tool and workpiece as well as their combination and interaction in machining process. Therefore, integrating machining parameters and sensory data may be better reflection of machining performance in monitoring.

The sensors currently implemented are a load cell for cutting forces measurement and a current transducer for spindle motor current. Vibration accelerometer will be also implemented. These sensors can reflect machining performance from different perspectives. DSP technology is used for sensory data processing in real time. It does data collection (sampling), filtering and FFT. A neural network structure is utilized for integration of the machining parameters and sensory features. This is a three layer perceptron shown in Figure 3.

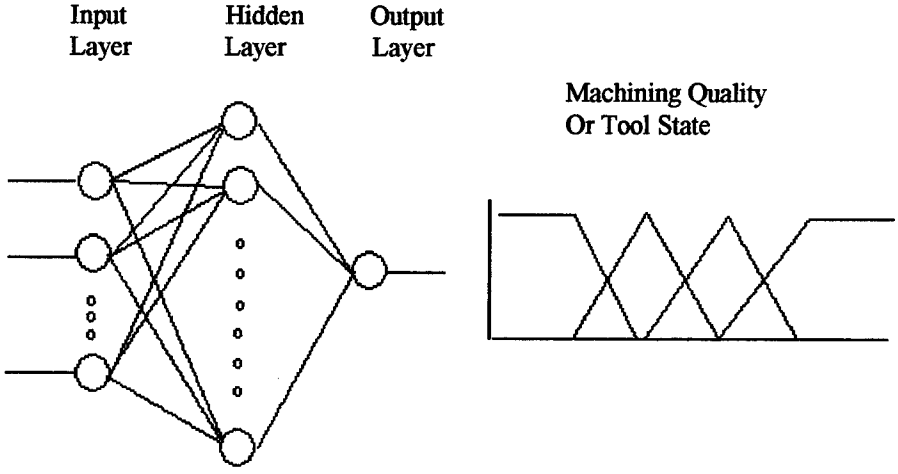


Figure 3. A NN model for integration of knowledge and sensory data

The input layer provides information of cutting speed, feed and depth of cut from fuzzy modeling and sensory data of forces Fx, Fy, Fz and spindle current Is as well as their combination (fusion) from sensor measurement and processing. The output is either machining quality (surface roughness Ra) or cutting tool wear Vb. The machining quality is classified as high, medium, acceptable and unacceptable. The tool states are classified as slight wear, medium wear, severe wear and tool breakage.

The classification of machining quality and cutting tool state is also made by using fuzzy classification.

This system developed has two parallel neural networks. One network is for monitoring machining quality and the other one for cutting tool monitoring. The two networks have similar structures in principle. The training process is carried out off-line until satisfied results are obtained. A back propagation training algorithm is applied in the neural networks. The back propagation algorithm concerns the actual and desired output vectors. The actual output from a given input vector (a weighted sum) is compared with the desired or target output. If there is no difference or the difference is within a predefined error range, no weights are changed. Otherwise, the weights are updated to reduce the difference. In the learning process a gradient search technique is used to minimized a cost function that is set as the mean square difference between the desired and actual outputs. In the BP network an input to a neuron is obtained as the weighted sum given by

$$net = \sum_{i=1}^{n} w_{ij} x_i + b_i \qquad (12)$$

where bi is bias vector acting as threshold. The output of the neuron is obtained by the activation function f (net) which normally has a sigmoid form as

$$out = f(net) = \frac{1}{1+\exp(-net)} \qquad (13)$$

The updated weights at iteration (n+1) are calculated according to the difference of the actual and desired outputs Oi and di.

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij} \qquad (14)$$

The change in weights Δw_{ij} is given by

$$\Delta w_{ij} = \alpha O_i \delta_j \qquad (15)$$

where α is a training rate coefficient, Oi is the output of neuron i in the previous layer, and

δ_j is the error coefficient related to the difference of the desired and actual outputs in the layer. The error for neurons in the layer may be given by

$$\varepsilon = \sum_i (O_i - d_i)^2 \qquad (16)$$

In order to provide faster convergence, momentum technique is implemented with the back propagation learning. Momentum acts like a low pass filter and allows a network to respond not only the local gradient, but also to recent trends in the error surface. It prevents the network from getting stuck in a shallow local minimum. With the learning feature of neural networks, system function for on-line machining monitoring can be adjusted with respect to different machining requirements.

5. Conclusion

This paper proposed a methodology of integrating human knowledge and sensor fusion for machining monitoring and control. Its validity is verified by experimental testing. Human knowledge in machining is interpreted by fuzzy representation in fuzzy inference models and appropriate machining parameters are selected by using the fuzzy models. These parameters are further integrated with sensory features in a neural network structure. Combining human knowledge and actual sensory data, machining performance is effectively evaluated. On-line monitoring of machining quality and tool wear is achieved.

6. References

[1] M. A. EL Baradie, "A fuzzy logic model for machining data selection", Int. Journal of Machine Tools and Manufacture, Vol. 37, No. 9, pp. 1353–1372, 1997.

[2] X. D. Fang, "Expert system-supported fuzzy diagnosis of finish-turning process states", Int. Journal of Machine Tools and Manufacture, Vol. 35, No. 6. Pp. 913-924, 1995.

[3] T. J. Ko and D. W. Cho, "Cutting state monitoring in milling by a neural network", Int. Journal of Machine Tools and Manufacture, Vol. 34, No. 5, pp. 659-676, 1994.

[4] S. Purushothaman and Y. G. Srinivasa, "A back-propagation algorithm applied to tool wear monitoring", Int. Journal of Machine Tools and Manufacture, Vol. 34, No. 5, pp. 625-631, 1994.

[5] D. E. Dimla Jr, P. M. Lister and N. J. Leighton, "Neural network solutions to the tool condition monitoring problem in metal cutting – a critical review of methods", Int. Journal of Machine Tools and Manufacture, Vol. 37, No. 9, pp. 1219-1241, 1997.

Receptor-effector neural-like growing network -- an efficient tool for building intelligence systems

Yashchenko Vitaly
Institute of Mathematical Machines & Systems NANU
42, Glushkova Avenue
Kyiv, 252187,
Ukraine

Abstract

The report describes a new type of neural networks - receptor-effector neural growing network (ren-GN) which contains a reconfigurable growing receptor area and a reconfigurable growing effector area. That allows to model conditions coming from the external environment, and in accordance with these conditions generate control influences to the environment. Besides this, multidimensional ren-GN allows us to perceive information in different forms (e.g., text, sound, graphics and others) and generate control signals to the corresponding executive devices. Such a network may be highly efficient for the creation of intelligence systems.

Key Words: Neural-like networks, intelligence systems, robots.

1. Introduction

The experience of building of intellectual systems and robots accumulated in different countries for many years has shown that modern universal computing facilities, in spite of their intensive improvement, are insufficiently efficient. The problem of building of principally new computing facilities for solving such problems is absolutely obvious.

In this connection, on the basis of analysis of scientific ideas that reflect regularities in the construction and operation of biological structures of brain, as well as an analysis and syntheses of knowledge, work out by different directions in Computer Science, the basis of theory of *a new class of neural-like growing networks*, not having analogue in the world practice was designed. A new technology of information handling, which unite in itself the best qualities of the technologies of information handling in semantic networks, neural networks and intellectual systems was designed [Yashchenko, 1998].

2. Neural-like growing networks

A neural-like growing network *(n-GN) is understood as a collection of determined by a definite way of interconnected neural-like elements, intended for receiving and transformation of information, moreover in the process of receiving information the network increases in size - grows.*

In the theory of neural-like networks, the main notions are notions of a structure, which reveal relationship scheme and interactions between elements of a network, as well as a notion of architecture.

Neural-like networks are presented by the following categories: a *topological (spatial)* structure - graph of relationships of elements in a network;

a *logical* structure defines principles and rules of establishing relationships, as well as logic's of network operation;

a *physical* structure - a scheme relationship of physical elements of a network (in the case of hardware realization of an neural-like network).

The *architecture of a network* is defined as principles of building of a network, which express the unity of physical and logical structures.

The *topological structure of a neural-like growing network* is defined as a coherent oriented graph (fig.2.1). The processes of passing and remembering of information in the network are considered by means of graphs in the n-GN theory.

Neural-like growing networks will formally be assign as follows: $S = (R, A, D, M, P, N)$, where $R = \{r_i\}$, $i = \overline{1,n}$ - is a finit set of receptors; $A = \{a_i\}$, $i = \overline{1,k}$, - a finit set of neural-like elements; $D = \{d_i\}$, $i = \overline{1,e}$, - a finit set of arcs that link receptors with neural-like elements, and neural-like elements between themselves; $P = \{P_i\}$, $i = \overline{1,k}$, $N = h$, where P - a threshold of excitation of a top a_i, $P = f(m_i) > P_0$ (P_0 - is minimum allowed threshold of excitation) provided that set of arcs D, which come to the top a_i, corresponds to an set of weighted factors $M = \{m_i\}$, $i = 1,w$, whereas m_i can take both positive, and negative values.

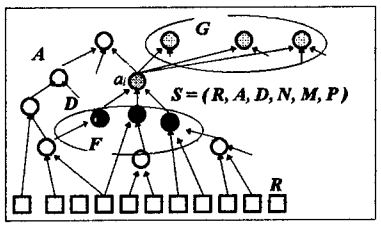


Fig.2.1.

In a network, a subset F of exitated tops from the set of tops, having direct relationship with the top a_i, and subset of excited tops of the network G, not having downwards relationships with other excited tops stands out. Symbols $\overline{\overline{F}}$ and $\overline{\overline{G}}$ mark the powers of subsets F and G, accordingly.

Logical structure n-GN is defined by the set of its building rules and operation.

Rule 1. If during the perception of information, a subset of tops F from the set of tops, having direct relationship with the top a_i, is excited, and $\overline{\overline{F}} \geq h$ the relationships of a top a_i with tops from the subset F are liquidated and a new top a_{i+1} joins the network, whose entries are connected with entries of all tops of the subset F, and the exit of a top a_{i+1} is connected with one of the inputs of a top a_i, whereas the input relationships of the top a_{i+1} are assigned weighted factors m_i, corresponding to the weighted factors of liquidated relationships of the top a_i, and top a_{i+1} is assigned the threshold of excitation P_i, equals $f(m_i)$, (function from weighted relationship factors, which fall into the top a_{i+1}). Outcoming relationship of this top is assigned a weighted factor m_i, equal $f(P_i)$. Relationships, outcoming from receptors, are assigned a weighted factor, $f(b_i)$, function from the code of sign b_i, corresponding to a given receptor (fig.2.2).

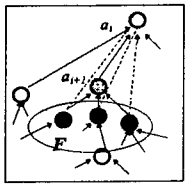


Fig.2.2.

Rule 2. If during the perception of information, a subset of tops G is excited, and $\overline{\overline{G}} \geq h$ a new associative top a_{i+1}, joins the network, which is connected by turning arcs with all tops of the subset G. Each of turning arcs is assigned a weighted factor m_i, equal $f(P_i)$ of a corresponding top from the subset G, and a new top a_{i+1} is assigned a minimum threshold of excitement P_i, equal to the function of weighted factors m_i of incoming arcs (fig.2.3).

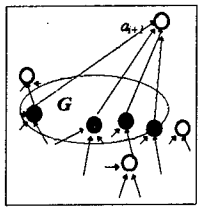


Fig.2.3.

Information in neural-like growing networks is stored as a result of its reflecting in the structure of a network. New information input into the network causes a process of building of its structure.

Neural-like growing networks are a dynamic structure, which changes depending on values and time of arrivals of information on to receptors, as well as former condition of the network. Information on objects is presented in it by set information on objects of excited tops and relationships between them. Storing the object descriptions and situations is accompanied by input in to the network of new tops and arcs when turning a group of receptors and neural-like elements became excited. The process of excitation spreads on the network, as a wave.

3.Receptor-effector neural-like growing networks

It is known that "An organism is educated by buildings sensing and motor schemes: it extracts from its experience correlation's between information, perceived by its sensor systems, and its own actions (motor activity)" [P.Lindsey, D.Norman].

Thereby, the education and interaction of biological objects with the environment is realized through acts of motion. For ensuring a possibility of modeling of processes of education and acquisition by the system of the knowledge on the external world, n-GN were developed in to receptor-effector neural-like growing network (ren-GN).

The topological structure of ren-GN is presented by a graph (fig.3.1). In ren-GN one reveals receptor R field (an analogue of sensor and receptor areas of biological objects), effector E field (an analogue of a motor area of biological objects), receptor A_r and effector A_e zones.

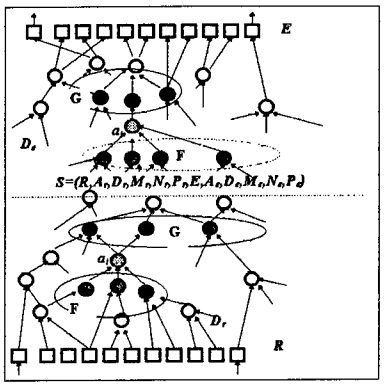


Fig.3.1

Formally, ren-GN will be assigned as follows: $S = (R, A_r, D_r, P_r N_r, E, A_e, D_e, P_e, M_e, N_e)$, $R = \{ r_i \}$, $i = \overline{1,n}$ - a finit set of receptors, $A_r = \{ a_i \}$, $i = \overline{1,k}$ - finit set of neural-like elements of a receptor zone, $D_r = \{ d_i \}$, $i = \overline{1,e}$ - a finit set of arcs of a receptor zone, $E = \{ e_i \}$, $i = \overline{1,e}$ - a finit set of effectors, $A_e = \{ a_i \}$, $i = \overline{1,k}$ - a finit set of neural-like elements of an effector zone, $D_e = \{d_i\}$, $i = \overline{1,e}$ - finit set of arcs of an effector zone, $P_r = \{P_i\}$, $P_e = \{ P_i \}$, $i = \overline{1,k}$, where P_i - a threshold of excitation of a top a_{ir}, a_{ie}, $P_i = f(m_i)$ under the condition that the set of arches D_r, D_e coming on tops a_{ir}, a_{ie}, corresponds to set of weight factors $M_r = \{m_i\}$, $M_e = \{m_i\}$, $i = \overline{1,w}$, and m_i can accept both positive, and negative values. N_r, N_e - are variable factors of connectedness of receptor and effector of zones.

In ren-GN, subsets of exited tops F_r and F_e of receptor and effector zones stand out accordingly, and subsets of exited tops G_r and G_e of receptor and effector zones of the network. Symbols $\overline{\overline{F}}$ and $\overline{\overline{G}}$ marked the powers of subsets F_r, F_e, and G_r, G_e accordingly.

The logical structure of ren-GN. As far as the composition of ren-GN includes receptor and effector zones, interacting between themselves, there is a need in the building of rules of development and operations of the network. These rules are formulated as follows.

Rule 3. If during the perception of information by a receptor field a subset F_r from the set of tops, having direct relationship with the top a_i^r, is excited, herewith $\overline{\overline{F}} \geq h$, and at generating of actions by effector zone a subset G_e is excited and $\overline{\overline{G}} \geq h$, the relationships of the top with tops from the subset F_r are liquidated, and a new top a_{i+1}^r, joins the network, which entries are connected with outputs of all tops of the subset F_r, and output of a top a_{i+1}^r is connected with one of the inputs of a top a_i^r, with assigning to incoming relationships of top a_{i+1}^r weighted factors m_i, corresponding to weighted factors of liquidated

relationships of top a_i^r, while top a_{i+1}^r is assigned a threshold of excitation Pa_{i+1}, which equals to function from the weighted factors of relationships, incoming into a_{i+1}^r top. Outcoming relationships of a top . a_{i+1}^r are assigned weighted a factor m_i which equals $f(Pa_{i+1}^r)$.

Relationships, coming from receptors, are assigned a weighted factor, which equals to the code of sign b_i, corresponding to a given receptor ave assigned to the relationships coming out from receptors. In effector zone a new assotiative top a_{i+1}^e, which connects by outcoming arcs with all tops of a subset G_e connects to the netwotrk. Each of outcoming arcs is assigned a weighted factor, m_i, equal , $f(Pa_i^e)$ of a corresponding top from the subset G_e, while to a new top , a_{i+1}^e a minimum threshold of excitement Pa_{i+1}^e , which equals to functions from weighted factors, m_i of incoming arcs, is assigned. Top a_i^r of a receptor zone is connected outcoming arc with the new top of an effector zone. New tops immediately after the introduction to the network are in the excited condition (fig.3.2).

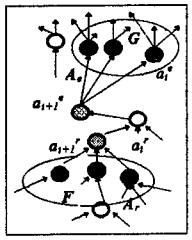


Fig.3.2

Rules 4, 5, 6 are formulated in accordance with fig.3.3, 3.4, 3.5 and in connection with restrictions on the volume of presented material which are not shown here.

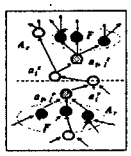


Fig.3.3

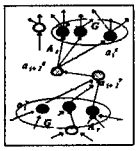


Fig.3.4

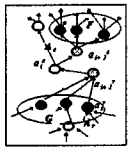


Fig.3.5

Information in receptor-effector neural-like growing networks is stored as a result of its reflection in the structure of network. Input of new information to the network causes a process of building of its structure and shaping the control influences to the external environment (i.e. to educate a network to work out control signals), in accordance with knowledge obtained by the network as a result of accumulations, analysis, categorizations and generalising information from the external world.

4. Multidimensional receptor-effector neural-like growing networks

For remembering and processing of the descriptions of images of objects or situations of a problem area, as well as generations of control influences by means of different information spatial presentations (text, sound, graphics etc.), multidimensional receptor-effector neural-like growing network (mren-GN) are entered.

Topological structure of multidimensional receptor-effector neural-like growing network is presented by the graph (fig.4.1). Formally mren-GN will be assign as follows.

$S = (R, A_r, D_r, P_r, M_r, N_r, E, A_e, D_e, P_e, M_e, N_e)$, $R \supset Rv, Rs, Rt$, $A_r \supset Av, As, At$, $D_r \supset Dv, Ds, Dt$, $P_r \supset Pv, Ps, Pt$, $M_r \supset Mv, Ms, Mt$, $N_r \supset Nv, Ns, Nt$, $E \supset Er, Ed, Ed$, $A_e \supset Ar, Ad1, Ad2$, $D_e \supset Dr, Dd1, Dd2$, $P_e \supset Pr, Pd1, Pd2$, $M_e \supset Mr, Md1, Md2$, $N_e \supset Nr, Nd1, Nd2$, here ***Rv, Rs, Rt*** - is a final subset of receptor, ***Av, As, At*** - is a final subset of neural elements, ***Dv, Ds, Dt*** - is a final subset of arches, ***Pv, Ps, Pt*** - is a final set of thresholds of excitation of neural elements of a receptor zone, belonging, for example, to visual, acoustical and tactile information spaces, ***Nr*** - is a final set of variable factors of connectivity of a receptor zone, ***Er, Ed1, Ed2*** - is a final subset of effectors, ***Ar, Ad1, Ad2*** - is a final subset of neural elements, ***Dr, Dd1, Dd2*** - is a final subset of arches of effector a zone, ***Pr, Pd1, Pd2*** - is a final set of thresholds of excitation neural elements of the effector zone, belonging, for example, to the speech information space and to the space of actions. ***Ne*** - is a final set of variable factors connectivity of a effector zone. The logical structure of mren-GN is described by rules 3-7.

Rule 7. If during arrival of different information spaces of external information on receptor fields, in receptor zone of these information spaces a subset Q_r of finite tops, belonging to this descriptions is excited, and herewith in effector zones of corresponding information spaces subset Q_e finite tops, working out a set of actions, corresponding to input information is excited, then tops of receptor areas of these information spaces, belonging to subset Q_r, are connected between themselves with bi-directional arcs. The tops of effector zones, belonging to a subset Q_e, are also connected between themselves with bi-directional arcs (fig.4.1).

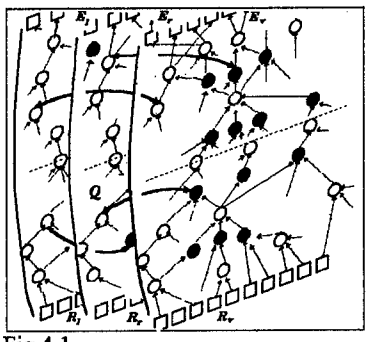


Fig.4.1

Thereby, in ren-GN and mren-GN, information about the external world, its objects, their conditions and situations which describe relations between them, as well as information on actions, caused by these conditions, is saved due to its reflecting in the structure of a network, while the arrival of new information causes a shaping of new associative tops and relationships and their redistribution between tops, which appeared earlier, herewith general parts of these descriptions and action appear, which are automatically generalised and are classified.

The main distinctions and comparative features the neural-like growing networks and common neural network are given in tabl.1

tabl.1

Neural-like growing networks	Neural networks
Neural-like element Microprocessor with memory.	**Neural element** Threshold element
It is defined some arbitrary function of enters, for example, Bayes' formula **P(H:E)=P(E:H) P(H)/(P(E:H) P(H)+ P(E:not H) P(not H))***	It is defined the weighted sum of enters, non linearly processed $x_i = F_i(\sum_{j=1}^{n} w_{ij} x_j + f_i - h_i)$
Connections and weights are set and appear equal as as many it is necessary	**Connections and weights** are defined by the architecture of the network. The number of networks is redundant. The special methods of the elimination of the connections are required.
The factor of connectivity enables to control connection neural element relations.	Absent
Reconstructed structure. Neural elements are connected by sense.	**Fixed structure.** Elements are connected everyone to everyone.
The possibility of composition and decomposition (deduction - induction). The object is defined by set of attributes and vice versa the set of attributes is defined by object.	Absent
Multilevel structure. The number of levels (layers) is arbitrary and is defined by sense.	**Usually not more three levels (layers) are used.** Using more then 3 levels has no sense.
The duration of training is from some minutes to some seconds.	**The duration of training** is from many hours to some seconds.
Effector a zone. Develops, classifies, and generalizes actions adequate to conditions formed in the receptor zone.	Absent
Appearance of false phantoms (false attractors) is absent	**Appearance of false phantoms** (false attractors) is present.
Network capacity 100%	**Network capacity** 20-30%
Parallelism of the computation is realized on the brunches of activity in all layers parallel. The efficiency of the computation is heightened (count on active part of the network).	**Parallelism of the computation** is realized on layers sequentially. The efficiency of the computation is reduced (count on the whole matrix).

*) P(H) – a priory probability of the outcome in the case of absence of additional illustrations.

P(H:E) – the probability of some hypothesis H realization in the presence of certain confirmation of illustrations E.

P(E:H), P(E:неH) – correspondingly , the probability of receiving answer "Yes", if the possible outcome is correct or incorrect .

5. Example of construction n-GN

Instance 1. The principle of constructing n-PC (for simplicity of perception) will be looked at the example of constructing the multiconnection growing network.

Formaly n-PC is described so: ***S=(R, A, D, N).***

Let be learning access, which consests of k-notions: *1. a,b,c,d; 2. b,c,d,e,g,h; 3. d,e,f; ... k. d,e,h.*

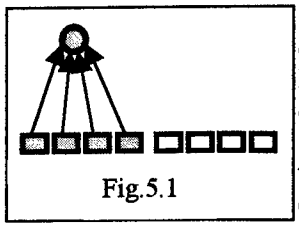


Fig.5.1

Let's set up variable coefficient of connectivity N≥5. In this case when entering the description of the first notion (a,b,c,d) on the receptor field, the receptors 1,2,3,4 are changed over to the state of excitation. The vertex *a,b,c,d* is formed and the connections between vertex and excited receptors are set up (fig.5.1.). The vertex is changed over to the state of excitation. In a definite time the excitation is taken of from receptors and a vertex.

When entering the description of the second notion (b,c,d,e,g,h) on the receptor's field, the receptors 2,3,4,5,7,8 are changed over to the state of excitation. The number of sighns, coincided with the description of the first notion (b,c,d)=3, then N=3 and in this case the second vertex b,c,d,e,g,h is formed (fig.5.2). The vertex is changed over to the state of excitation. The excitation of the vertex and receptors is taken off.

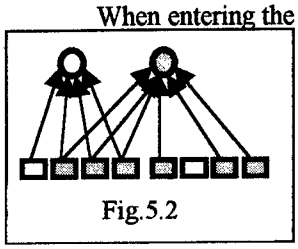


Fig.5.2

When entering the description of the third notion (d, e, f) on the receptor's field the receptors 4,5,6 are changed over to the state of excitation and N=2 then in this case the third vertex d,e,f is formed, (fig.5.3). The vertex is changed ove to the state of excitation. In a definite time the excitation is taken off from the vertex and receptors.

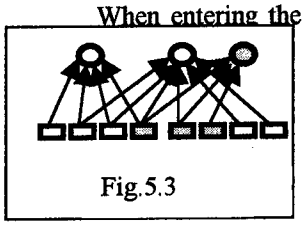


Fig.5.3

When entering the description of the k-notion (d,e,h) on the receptor's field, the receptors 4,5,8 are changed over to the state of excitation, N = 2, the k-vertex is formed (fig.5.4).

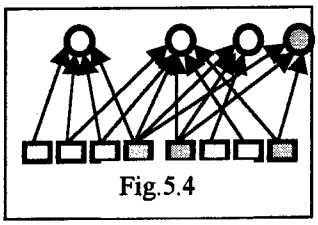


Fig.5.4

The vertex is changed over to the state of excitation. Then the excitation is taken off from the vertex and receptors.

This it is formed on-layer m-PC in which the description of k-notion is stored.

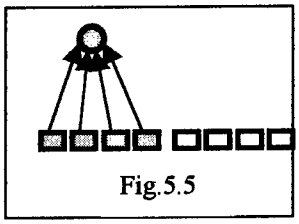


Fig.5.5

Instance 2. If the variable coefficient of the connectivity will be set N≥3. In this case when entering the description of the first notion (*a,b,c,d*) on the receptor field, the

receptors 1,2,3,4 are changed over to the state of excitation. The vertex *a,b,c,d* is formed and the connections between vertex and excited receptors are set up (fig.5.5.). The vertex is changed over to the state of excitation. In a definite time the excitation is taken of from receptors and a vertex.

When entering the description of the second notion (*b,c,d,e,g,h*) on the receptor's field *R*, the receptors 2,3,4,5,7,8 are changed over to the state of excitation. The number of signs, coincided with the description of the first notion (*b,c,d*)=3, then N=3, the vertex (*b,c,d*) and (*b,c,d,e,g,h*) are formed. The connection of the vertex *a,b,c,d* with receptors 2,3,4 are liquidated. Inputs of the vertex *b,c,d* are connected with receptors 2,3,4 and outputs of this vertex are connected with inputs of the vertex (*a,b,c,d*) and *b,c,d,e,g,h*, and these vertices are changed over to the state excitation (fig.5.6). In a definite time the excitation is taken off from the vertex (*b,c,d*) (*b,c,d,e,g,h*) and receptors.

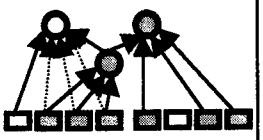


Fig.5.6

Fig.5.7

When entering the description of the third-notion on the receptor's field, the new vertex (*d,e,f*) is formed (fig.5.7).

When entering the description of k-notion on the receptor's field, the new vertex is formed (fig.5.8). In this case the separation of the common sighns, described notions is performed.

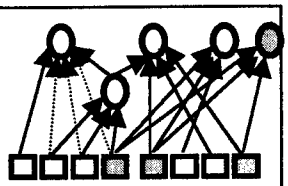


Fig.5.8

Thus the description of the notion (the vertex of the network) and sighns are stored. Besides ehat, the information, which enteres the receptor's fields of the networks, is classified and structured automicially.

When forming new vertex in n-GN, weight coefficients of connection m_i and thresholds of the excitation of the vertex P_i are considered, that is constructing n-GN is performed analogically with building m-GN, but in accordance with rules, which are described in the matereals presented before.

5.1. Prospects for using the receptor-effector neural-like growing networks for intellectual system

Using the idea of organization of n-GNs in their physical representation, we can create an intelligent multimicroprocessor system with a neuron-ensemble structure (IMSNS). The architecture of this system consists of a collection of microprocessor modules, each represented by an array of microprocessors.

The high-intelligence multimicroprocessor system with a homogeneous multidimensional array neuron-ensemble structure is a new-generation artificial intelligence system.

The main advantages of IMSNS stem from new approaches to architectural and program organization of the system using the theory of growing semantic networks with a neuron-ensemble structure, analysis of array structures with neuron-ensemble organization, new information processing technologies and two-level organization of system architecture, modularity and homogeneity of hardware and software tools, and use of array structures with multidimensional organization which dispense with physical realizations of connections between the nodes of the neuron network..

IMSNS can be manufactured as a separate PC board, an intelligent PC coprocessor, or a powerful new-generation high-intelligence PC.

The proposed conception of a multimicroprocessor system with a homogeneous array multidimensional neuron-ensemble structure makes it possible to store and classify the input information, and to execute operations in accordance with this information allowing for the frequency of occurrence of events, their probability and significance. It is intended for solving complex problems that require a large volume of information, processing of large data files, generation of knowledge bases and artificial intelligence systems. This creates the potential for substantial gains in the productivity of the user's intellectual labor.

The functions performed by the system include description of situations and concept formation; transformation of situations and extraction of new concepts; creation of associative links; associative search; action planning; instruction and self-learning.

Some of the problems solved by the system include parallel processing of computational tasks; generation of a feature space and class description; image recognition; testing and diagnosing of technical systems; creation of real-time expert systems with powerful hardware support in various domains of human activity, such as biology, medicine, military science, meteorology, geology, nuclear physics, criminology, production control, economics, environmental science, etc.

The hardware implementation of neuron networks makes it possible to achieve high information processing speeds. However, the difficulties with implementation of a large number of interconnections limit the size and correspondingly the efficiency of the modeled network.

In the proposed IMSNS these difficulties are avoided by dispensing with physical realization of interconnections. Networks thus can be created with an unlimited number of pseudoconnections. The specific features of the system make it possible to obtain the result almost instantaneously (after some initial learning).

The IMSNS architecture is characterized by a high level of decentralization and parallelism.

The multimicroprocessor system with a neuron-ensemble structure (NS) that ensures automatic effective

parallelization and classification of information streams is characterized by two-level organization. The bottom level, which contains the tools for representation and preliminary processing of input information (transformation of input information from natural language representation to machine language representation in the form of feature code combinations), has been implemented in a traditional von Neumann architecture. The top level performs multilevel parallel processing of a set of information features (detects the presence of features and connections between described objects, establishes arc weights and element thresholds, processes the information corresponding to the given feature in the nodal elements of the array), and thus generates a multidimensional neuron-ensemble structure in a three-dimensional array.

Thus, the IMSNS, while implemented in a traditional von Neumann architecture, on the whole functions according to the principles of associative artificial intelligence systems.

In classical multiprocessor computing systems, the modular organization of hardware and software tools imposes a dependence of information processing efficiency on the consistency of the physical structure of the processing resources with the logical problem architecture. In the absence of this consistency, a large part of the resources may remain idle awaiting the results of intermediate computations.

In IMSNS this difficulty is resolved by dynamic linking of the logical and physical structures in the process of loading the information (concurrently with the processing of feature codes) into the array of microprocessor elements.

The IMSNS architecture differs in an important respect from existing implementations of multiprocessor systems with overall step-by-step clocked control. In IMSNS, step-by-step control is implemented only within each microprocessor element, whereas on the whole the clock is replaced with an indexing mechanism, which fixes the termination of the transients excited by changes in input signals.

Another distinguishing property of the IMSNS architecture is the possibility of combining the data base with the knowledge base in the array of microprocessor elements. Data in the NS are represented by the set of excited nodes (microprocessors storing the physical parameters of the data), while knowledge is represented by the interconnections between the nodes and also by the weights of the interconnections and the excitation thresholds of the nodes. The basic operations of the NS, which computes the interconnection weights, compares the results with the nodal excitation thresholds, and performs other functions, do not require special software or hardware tools for data base or knowledge base management. The programs that compute the nodal activity coefficients are distributed throughout the network, simple to execute, and their structure is independent of the content of knowledge and the specifics of the application domain. The network is transformed under the control of an array of controlling microprocessors by a special recursive algorithm, which generates in the interior of the microprocessor element module a growing neuron structure that changes in response to each new information input. The multilevel structure of the system makes it possible to store knowledge about knowledge, i.e., to transform previously stored rules in accordance with new rules.

The IMSNS can be implemented using ordinary medium-scale-integration digital logic connected by line buses into an array.

To ensure execution of these operations and solution of specialized problems (e.g., pattern recognition on its own), IMSNS can be implemented using special-purpose VLSI, which should substantially reduce their size and power consumption. Moreover, the proposed conception of a multidimensional neuron-ensemble structure can be implemented using optimal neuron networks with holographic memory based on semiconductor laser arrays.

New-generation computers based on IMSNS may find wide uses in various areas and penetrate into many spheres of the world market.

At present, the development is in the stage of theoretical substantiation and experimental testing. Partial modeling has been carried out. The recursive algorithm constructing an array with a multidimensional neuron-ensemble structure has been demonstrated to function as intended.

6. Conclusions

At present, the development is in the stage of theoretical substantiation and experimental testing. Partial modeling has been carried out. The recursive algorithm constructing an array with a multidimensional neuron-ensemble structure has been demonstrated to function as intended.

References

[Yashchenko, 1998] Vitaly Yashchenko. Neural-like growing networks - new class of the neural networks. In Proceedings of the International Conference on Neural Networks and Brain Proceeding, pages 455--458, Beijing, China, Oct. 27-30' 98.

[P.Lindsey, D.Norman, 1974] P.Lindsey, D.Norman CONVERSION information beside the person (Introduction To the psychology).Under edit. A.R. Lyriy. - M.: - 1974. - c.549.

Session RC1
Emerging Applications II
Chair: Peter Wide, Orebro University, Sweden

DATA ANALYSIS IN THE GRAVITY PROBE B RELATIVITY EXPERIMENT

M.I.Heifetz
Gravity Probe B, W.W.Hansen Experimental Physics Laboratory,
Stanford University, Stanford, CA, U.S.A.

G.M.Keiser
Gravity Probe B, W.W.Hansen Experimental Physics Laboratory,
Stanford University, Stanford, CA, U.S.A.

Abstract *Gravity Probe B (GP-B) is a gravitational experiment designed to measure two predicted by General Theory of Relativity precessions of a free-falling gyroscope placed in a polar orbit about the Earth. The frame-dragging effect (drift perpendicular to the orbital plane) has never been directly measured before, while the geodetic effect (drift in the orbital plane) will be measured with an unprecedented accuracy. GP-B Data Analysis includes processing telemetry data from several physical sources placed on the GP-B spacecraft: the science gyroscope's readout system, telescope optical system, Global Positioning System (GPS), and the spacecraft's attitude control system. We discuss here only one of the numerous problems that need to be resolved through Data reduction: precise estimation of the gyroscope's relativistic drift rates. The two-step nonlinear filtering approach is presented and estimation recursive algorithms that will be used in the GP-B Data Analysis are discussed.*

Keywords: Data analysis, Multi-sensor signal processing, Nonlinear filtering.

1 Introduction

The Gravity Probe-B Relativity experiment [1] makes use of gyroscopes in Earth polar orbit to measure two effects of Einstein's General Theory of Relativity with previously unachieved accuracy - the precessions of the local inertial frame free falling about the Earth with respect to the inertial frame of the distant universe. In a polar-orbiting spacecraft, with the gyroscope spin axes lying in the plane of the orbit and perpendicular to the Earth's rotation axis, the two effects to be measured - the geodetic effect and the never before measured frame-dragging effect - are at right angles with respect to each other. The magnitudes of the two effects in a 650 km polar orbit are 6.6 arcsec/year for the geodetic effect and between 33 and 42 marcsec/year for the frame-dragging precession, depending upon the choice of guide star. Figure 1 is a schematic representation of the GP-B experimental concept.

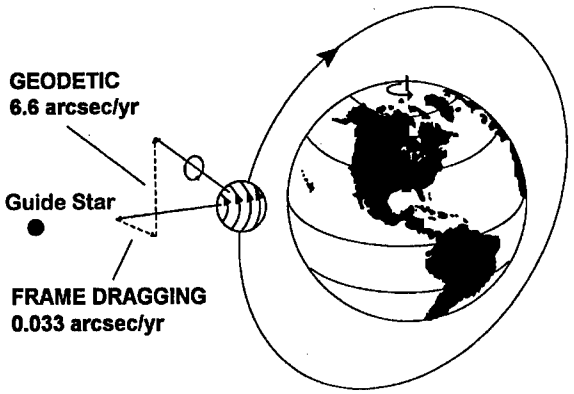


Figure 1: GP-B experimental concept

The experiment will measure these precessions with respect to the line-of-sight to a reference star whose position and proper motion with respect to the inertial frame of the distant

universe (the "fixed" stars) is determined in separate astrometric measurements. The goal is to measure the geodetic precession to better than 1 part in 10^4 and the frame-dragging precession to better than 1 percent. The fundamental objective of the relativity mission is to measure the angular rate between the local frame (free-falling about the Earth) and distant inertial space (defined by the "fixed" stars) to an accuracy better than 0.5 marcsec/year, independently in each direction, for a one year experiment.

Changes in the direction of local inertial space are detected by measuring the Science Gyroscopes (SG) spin axis direction relative to the spacecraft (S/C) in which the SG are contained with a low noise, non-interfering readout system (four gyroscopes are used for redundancy with certain systematic effects removed by spinning them in opposite directions). The spacecraft is referenced to distant inertial space, as calibrated by a guide star, by the Science Telescope (ST) fixed to the spacecraft. The SG and ST data are subtracted from each other and corrected for known effects (such as aberration of starlight and others) in a data reduction process whose output is the measured drift between the local and distant inertial spaces; i.e., general relativistic drift. Extremely low levels of acceleration on the SG's are required to keep the Newtonian drifts of the gyroscopes from overly corrupting the experiment data. The spacecraft is therefore operated drag free to minimize the effects of disturbances on the science gyroscopes and guarantee that they remain in a purely gravitational orbit (geodesic). One science gyroscope will be used as the drag free proof mass to virtually eliminate the disturbing forces on that SG. The GP-B spacecraft and its attitude control system are designed to point the telescope continuously towards the guide star (or towards its apparent position) and to minimize the body-fixed pointing error.

The GP-B spacecraft also rotates at a constant roll rate about the line of sight to the guide star. Rolling the satellite allows to move the science signal to the roll frequency, where the readout measurement noise (which has 1/f spectrum) is lower. Rolling also allows a single gyroscope pick-up loop and its readout to measure both the geodetic and frame-dragging precessions. Good orbital information of both the Earth and the S/C is required for the experiment calibration against known fundamental processes. The readout system scale factor is precisely calibrated during the experiment using the optical aberration of starlight due to the spacecraft motion around the Earth and the Sun.

The GP-B gyroscope has been designed as a near perfect inertial instrument: Newtonian precession due to the classical torques is supposed to be less than 0.3 milliarcsec/year. This means that the gyroscope's measured precession angle is assumed to be caused mainly by the relativistic effects. A detailed analysis of classical torques acting on the GP-B gyroscope is presented in [2].

In this paper we describe our approach to the GP-B Data Analysis as the set of filters that estimate the model-dependent system state vectors and calculate covariance matrices that represent statistical errors of the relativistic drift measurements due to the gyroscope and/or telescope readout noise and unmodeled disturbances.

2 GP-B Science Signal Model

The accuracy required in the GP-B experiment demands resolving numerous problems of the 'optimal' data processing of the GP-B science signals. Here we discuss only one of them: estimation of the relativistic geodetic and frame-dragging drift rates of the GP-B gyroscope from the data provided by the gyroscope's readout system. The GP-B readout system is based on the effect of magnetic field generated by a spinning superconductor(London moment) [3]. Precession of the gyroscope angular momentum results in the variations of the London moment, that is aligned with the instantaneous gyroscope spin axis. The science signal measured by the readout system

represents the London moment magnetic flux through the gyroscope pick-up loop converted to the output voltage by the SQUID magnetometer.

The simplified model of the gyroscope science signal can be represented as

$$z(t) = C_g\Big[(NS_0 + R_g t - \varepsilon_1(t))\cos(\omega_r t + \delta\phi) - (EW_0 + R_f t - \varepsilon_2(t))\sin(\omega_r t + \delta\phi)\Big] + b(t) + \nu(t), \quad (1)$$

where C_g is the readout system scale factor, NS_0 and EW_0 are North-South (in the orbital plane) and East-West (perpendicular to the orbital plane) initial misalignments, R_g and R_f are the average drift rates, ω_r is the spacecraft roll rate, $\delta\phi$ is the roll phase offset, ε_1 and ε_2 are optical aberration components, b is the readout system bias. Measurement noise $\nu(t)$ is assumed to be a white noise with zero mean and known covariance matrix R.

Optical aberration is a shift in the apparent direction towards the guide star due to the velocity of the spacecraft to the line of sight to the star. There are two categories of aberration for the GP-B spacecraft: orbital aberration caused by the satellite's orbital motion around the Earth, and annual aberration, due to the motion of the Earth around the Sun. Optical aberration signals are continuously calculated based on the information from the on-board GPS and NASA/JPL Earth ephemerides.

3 Relativistic drift rate estimation

Introducing the state vector of parameters to be estimated,

$$\mathbf{X} = \Big[R_g,\ R_f,\ C_g,\ \delta\phi,\ NS_0,\ EW_0,\ b\Big]^T, \quad (2)$$

and under reasonable assumption that some components of the state vector (2) may vary with time during the experiment, the data analysis problem is recognized as the *nonlinear* filtering problem: for the linear state vector model

$$x_{k+1} = \Phi_k x_k + \Gamma_k w_k,\ \ w \sim N(0, Q_k), \quad (3)$$

and nonlinear measurement model

$$z_k = F(x_k, t_k, \varepsilon_1(t_k), \varepsilon_2(t_k), \omega_r) + \nu_k, \quad k = 1, 2, \ldots, N \quad (4)$$

find the estimate $\hat{x}$ that minimizes the least-square cost function

$$J = \frac{1}{2}\left[\sum_{k=1}^{N}(z_k - F(x_k, t_k))^T R^{-1}(z_k - F(x_k, t_k)) + \sum_{k=1}^{N-1} w_k^T Q_k^{-1} w_k\right] \quad (5)$$

For the GP-B science signal structure (1), as it has been shown by numerous simulations, the standard nonlinear estimators, such as the extended Kalman filter (EKF) and the iterated extended Kalman filter (IEKF) [4] give, , a *biased* estimates of relativistic drifts R_g and R_f. The reason is that both EKF and IEKF linearize the measurement equation (4) and the cost function (5). To overcome that difficulty, a new nonlinear recursive two–step estimator has been developed [5]).

Instead of linearizing the cost function, it breaks the minimization procedure into two steps. A new set of states is defined for the first–step filter using nonlinear combinations of the unknowns, so that the measurement equation becomes linear with respect to the new ones. The choice of first step states is dependent on the particular problem being addressed. The first–step linear problem can be solved optimally by exploiting a linear Kalman filter. The second-step states are then calculated by treating the first step state estimates as 'new' measurements and by using an iterative Newton–Raphson searching algorithm.

By choosing the first step states as

$$y = f(x) = \begin{bmatrix} C_g(NS_0\cos\delta\phi - EW_0\sin\delta\phi) \\ -C_g(NS_0\sin\delta\phi + EW_0\cos\delta\phi) \\ C_g(R_g\cos\delta\phi - R_f\sin\delta\phi) \\ -C_g(R_g\sin\delta\phi + R_f\cos\delta\phi) \\ -C_g\cos\delta\phi \\ C_g\sin\delta\phi \\ b \end{bmatrix} \quad (6)$$

we convert the nonlinear measurement equation (1) into a linear one:

$$z = H(t)y + \nu, \tag{7}$$

where

$$H = \Big[\cos\omega_r t,\ \sin\omega_r t,\ t\cos\omega_r t,\ t\sin\omega_r t,$$

$$\varepsilon_1\cos\omega_r t + \varepsilon_2\sin\omega_r t,\ \varepsilon_1\sin\omega_r t - \varepsilon_2\cos\omega_r t,\ 1\Big] \tag{8}$$

Applying now the two–step estimator [5] we obtain the following recursive estimation procedure.

First–Step Optimization: ($\hat{y}$ and P_y are the optimal first step estimate and covariance matrix).

Measurement Update:

$$\hat{y}_k = \bar{y}_k + P_{y,k}H_k^T R_k^{-1}(z_k - H_k\bar{y}_k),$$

$$P_{y,k} = \left(M_{y,k}^{-1} + H_k^T R_k^{-1} H_k\right)^{-1} \tag{9}$$

Time Update:

$$\bar{y}_{k+1} = \hat{y}_k + f_{k+1}(\bar{x}_{k+1}) - f_k(\hat{x}_{k+1})$$

$$M_{y,k+1} = P_{y,k} + \left[\left(\frac{\partial f}{\partial x}\right) M_{x,k+1} \left(\frac{\partial f}{\partial x}\right)^T\right]\Big|_{x=\bar{x}_{k+1}};$$

$$-\left[\left(\frac{\partial f}{\partial x}\right) P_{x,k} \left(\frac{\partial f}{\partial x}\right)^T\right]\Big|_{x=\hat{x}_k} \tag{10}$$

Second–Step Optimization: ($\hat{x}$ and P_x are the optimal second step estimate and covariance matrix).

Iterative Measurement Update (i– iteration number):

$$\hat{x}_{k,i+1} = \hat{x}_{k,i} - P_{x,k,i}\, q_{k,i}^T; \qquad k = 1, 2, \ldots, N$$

$$P_{x,k,i+1}^{-1} = \left[\left(\frac{\partial f}{\partial x}\right)^T P_{y,k}^{-1} \left(\frac{\partial f}{\partial x}\right)\right]\Big|_{x=\hat{x}_{k,i}}$$

$$q_{k,i} = -(\hat{y}_k - f_k(\hat{x}_{k,i}))^T P_{y,k}^{-1} \left(\frac{\partial f_k}{\partial x_k}\right)\Big|_{x_k=\hat{x}_{k,i}} \tag{11}$$

Time Update:

$$\bar{x}_{k+1} = \Phi_k \hat{x}_k; \qquad k = 1, 2, \ldots, N-1$$

$$M_{x,k} = \Phi_k P_{x,k} {\Phi_k}^T + \Gamma_k Q_k {\Gamma_k}^T \tag{12}$$

Matrices H, Φ, Γ, Q and R, as well as nonlinear transformation $y = f(x)$, are defined above.

The two-step nonlinear estimator (2)–(12) has been used intensively for the general error analysis of the GP-B experiment. Figure 2 shows the dynamics of the estimation process and the potentially achievable accuracy of estimation. Qualitatively, the combination of the orbital (100-min period) and annual (1-year period) aberrations allows to determine the readout system scale factor C_g and roll phase offset $\delta\phi$ (science instrument dynamic calibration), which in turn allows to get the best estimate of the geodetic (R_g) and frame-dragging (R_f) relativistic drifts.

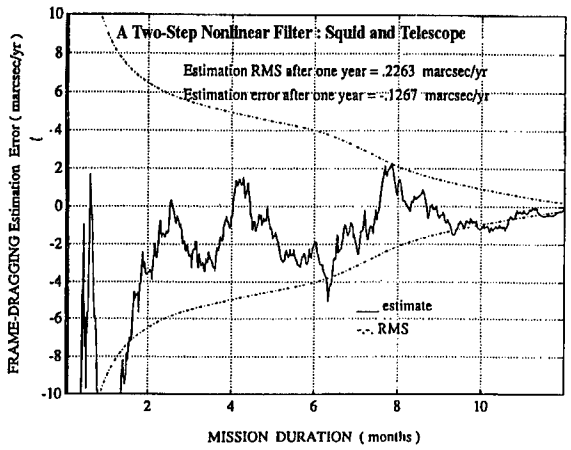


Figure 2: Relativistic drift estimation

The two-step filtering approach, described above, is also being used for various GP-B Data Analysis problems of the multi-sensor signal processing, where in order to achieve the required accuracy of the relativistic drift measurements, it is necessary to combine and optimally process data from the four GP-B science gyroscopes, science telescope's photodetectors, spacecraft attitude control system, on-board GPS, together with the auxiliary

information about the on-board environmental temperature and magnetic fields variations during the science mission. The 'bank' of filters with the different content of the state vectors, based on the above described methodology, is planned to be used in the data reduction that will start soon after the GP-B satellite's launch scheduled for October of the year 2000.

Acknowledgements

This work was supported by NASA grant NAS 8-39225 to Gravity Probe B.

References

[1] C.W.F.Everitt. *The Stanford Relativity Gyroscope Experiment: History and Overview.*, in *Near Zero,pp.587-639, W.H.Freeman and Co., New York, NY, 1980.*

[2] A.S.Silbergleit,M.I.Heifetz,G.M.Keiser. *Classical Torque Errors in Gravity Probe B Experiment*, in *Proceedings of the Third William Fairbank Meeting on the Lense-Thirring Effect, Rome-Pescara, Italy 1998.*

[3] F.London. *Superfluids Vol 1: Macroscopic Theory of Superconductivity.* Wiley, New York, 1953.

[4] T.F.Elbert, *Estimation and Control of Systems.* Van Nostrand Reinhold, New York, 1984.

[5] G.Haupt,N.Kasdin,G.Keiser,B.Parkinson. *Optimal Recursive Iterative Algorithm for Discrete Nonlinear Least-Squares Estimation.* Journal of Guidance, Control and Dynamics, **19**, 3, 643–649, 1996.

Hyperspace Data Mining with Applications to Biotech

Nianyi Chen*, Longjun Chen and Dongping Daniel Zhu*****

*Laboratory of Data Mining, Shanghai, P. R. China

**CISCO Systems, Inc., San Jose, CA, USA

***Zaptron Systems, Inc. 415 Clyde Avenue #108, Mountain View, CA 94043, Email: dan@zaptron.com

Abstract - Data mining offers an effective solution to biotech in that it requires no additional experiments, adds no new equipment, and causes no interruptions to operations. The most important role of data mining is the ability to separate data into different classes so that an accurate model can be built. This paper reports a map recognition (MREC) method and its application to biotech field. As a powerful data mining solution, MREC consists of data separation, hidden projection, back mapping, feature selection, and model building. The solution can be quantitative or qualitative depending on the pattern of the original data set. Examples using MREC are given to demonstrate its efficacy in drug manufacturing process.

Key words: data mining, drug design and processing, pattern recognition, projections, principal components.

I. Introduction to Drug Design and Production

1.1 Yield Enhancement in Drug Production

Drug manufacturing processes can be broadly classified into two methods: *synthetic* and *fermentation*. In synthetic methods, the production flow in general consists of many steps described by a production flow chart. The process usually starts from a raw material that can be converted into a number of intermediate products by a series of chemical reactions with other compounds. In drug production, yield is defined as the productivity (amount of product) per unit of raw materials, while manufacturing productivity is defined as the amount of the product in each batch of production (such as from a fermentation tank).

In synthetic drug manufacturing, the production flow charts are usually long, consisting of many steps, but the overall yield of final products is usually low. For example, if a synthetic process includes five steps and the yield of each step is 80%, the overall yield is equal to 32.07%. If we increase the yield of each step up to 95%, the overall yield will be is increased to 77.38%. In other words, the overall yield will be more than doubled when the step yield is slightly increased and the same amount of raw materials and manpower are used in each step.

Therefore, it is of great economic value to find methods that enhance the overall yield by optimizing the yield in each step in synthetic drug production. Data mining techniques have been used to process the chemical synthesis data in each step to find the "best" condition for yield enhancement. It can help increase business profit with little investment. Indeed, many chemical processes involving organic chemical reactions have been optimized with significant economic effects by data mining

Another type of drug production process is the *fermentation process* where drugs such as antibiotics are produced in a fermentation tank. It is well known that fermentation processes are, in most cases, very sensitive to a large number of influencing factors. They are usually so complicated that it is very difficult to find an optimization model to enhance the overall production yield. Data mining offers an effective solution to this problem by finding the best operation parameters to significantly enhance production yield.

1.2 Drug Design Issues

A major issue (a target) in drug design is to discover quantitative relationships between a drug's molecular structure, i.e., the way drug molecules are arranged, and its bioactivity, i.e., the effect of medicine or toxicity. It is also desirable to find relationships in molecule-molecule interactions, for example, the drug molecule-DNA molecule interaction. The bioactivity of a drug is, in most cases, experimentally measured by conducting animal tests. The molecular structure is usually described by a number of design parameters or features (factors).

Inventing a new drug by drug screening is a very lengthy, troublesome, and expensive task. People have to synthesize a huge number of new compounds, and test their bio-activities and toxicity. This is followed by biological tests and clinical tests before production. Drug design uses a series of computational methods to make "predictions" based on theoretical reasoning and modeling, in order to increase the efficiency of screening processes in exploring new drugs. Although these "prediction results" may not be 100% optimal, they can save huge amount of investment capital in drug screening, by giving design advisory or approximate models that are better than those obtained by human experiments. Data mining offers a powerful tool for finding such relationships among the various properties and parameters in drug screening.

1.3 Drug Diffusion Capability

Another very important issue in drug design is related to the diffusion capability of the molecules of a drug in human body, where the inner part of each cell is like a water solution, and the wall of each cell, made of oil-like materials, has the nature of "oil". For instance, if a drug can kill bacteria effectively but can not reach them in the human body, it is of no use.

A drug can diffuse quickly if it has suitable solubility both in "water" and in "oil". This issue is usually expressed by the octanol-water distribution of drug molecules. Sometimes we have no experimental data about this distribution, but want to use the molecular structure of a drug to predict the octanol-water distribution by a number of distribution coefficients between octanol and water. How to effectively calculate these coefficients is an important task. Thus identifying the relationship between molecular structure and solubility of a drug is yet another important task in drug research.

Oil (fat) and water can not dissolve each other, but they form two distinct layers in human body. However, if a third matter (substance) is added to this "two-layer" system, it will dissolve partially in oil and partially in water. The ratio of the concentration of this third matter in oil to that in water is called "coefficient of distribution," or "distribution coefficient". If a matter (substance) can dissolve in water but not in oil, the distribution coefficient is 0.0. If it dissolves in oil only, the coefficient will be infinity. Since human body is a mixture of water and oil (fat), the drug diffusion process in human body depends on the distribution coefficient of this drug in water and in oil. Ideally, it should not be too large or too small; otherwise its diffusion will be prevented by water and/or by oil. Since octanol is "oil-like", drug scientists use a drug's distribution coefficient in water and in octanol to correlate the diffusion ability of a drug in human body.

1.4 Organization of the Paper

In this paper we present the principle of the MREC method and its application to drug manufacturing. A number of techniques have been developed and built into the MasterMiner™ software suite. The software provides a set of effective and user-friendly tools to solve general data mining problems where various intrinsic data structures (models) manifest themselves.

Section 2 reviews various computational methods, with discussions on their advantages and limitations. Section

3 describes the proposed hyperspace data mining techniques including data separability, envelope method, feature selection and reduction, auto-boxing method. Section 4 shows a few real-world examples of using MasterMiner in biotech applications.

II. Problem Background

By nature, drug design or manufacturing is an optimization problem, and methods in pattern recognition and data mining can be used to offer effective solutions. Most pattern recognition methods are based on the computerized recognition of the multidimensional graphs (or their two-dimensional projections) of the distribution of samples from different classes in a multidimensional space. Independent variables (often called system input, features or factors) influencing the target (dependent variable or system output) are used to span a multidimensional space.

We can describe samples of different classes as points with different symbols in these spaces. Various pattern recognition methods can be used to "recognize" the patterns shown in the graph of *distribution zones* of different samples. In this way, a mathematical model can be obtained that describes the relationship (or regularity) among targets and factors. If we adjust criterion of classification, semi-quantitative models describing the regularities can be found at medium level of noise.

Unlike regression methods (linear regression, nonlinear regression, logistic regression, etc.) or the artificial neural networks (ANN) [4] that provide *quantitative* solutions, pattern recognition methods often provide *semi-quantitative* or *qualitative* solutions as well. This is of course a limitation of pattern recognition methods. However, this is not always a disadvantage, because many data sets exhibit strong noise, and a quantitative calculation would be *too precise to* represent them. Besides, practical problems in many cases are of the "yes or no" type. For example, a problem may be stated as "whether a chemical reaction will occur or not", or "whether an intermetallic compound will form or not." Pattern recognition is especially suited to offering adequate solutions to these types of problems.

As an important part of informatics, chemometrics and phamakocinetics, traditional techniques of pattern recognition, such as linear and nonlinear regression, partial least square (PLS) and artificial neural networks (ANN), have been widely applied to materials and drugs design for many years. In pattern recognition applications, the PLS method are usually used to find quantitative structure-activity relationships. However non-linearity exists among target and factors, and PLS often fails to give meaningful results. Commercial software products, based on a single computational technique, such as nonlinear or linear regression, ANN method, or PLS, are used in daily analysis at drug companies. MasterMiner™ software has a number of noticeable advantages over those based on pure ANN or pure regression (such as PCA). In particular, its data separability and the classification power on bio-active and bio-inactive compounds have been proved to be very effective in practice.

Data Mining

Data mining [1] [5] [6] is the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition technologies as well as statistical and mathematical techniques. Data mining in fact is an optimization technique, and it has found practical applications in designing and diagnosing products or process in various industries, including steel making, power generators, petro-chemical, materials design and manufacturing, drug screening and production, and operations management

[1] [8].

Data mining techniques [1] [2] [5] [6] [7] offer effective solutions to biotech, and are especially suitable for processing complex data sets in *non-linear*, *highly noisy*, and *multivariate* environment. It is used to build effective data models for control and prediction. Since complex data are common in many practical applications, traditional linear regression method is not appropriate, and advanced techniques are needed. In real-world applications, different techniques of soft computing are synergistic rather than competitive.

Feature Selection Issue

Drug design seeks to identify relationships between the structure of molecules and the bio-activity of drugs. How to select and use the design parameters to describe the molecular structure of a drug is a complicated problem. Molecular structure can be described by various parameters, including: (1) partial atomic charges in molecule studied by quantum chemical calculation; (2) intensity and format of the electrostatic field surrounding molecules studied by semi-empirical methods; (3) geometric arrangements of water molecules around the molecule; (4) dynamic parameters of drug molecules studied by molecular mechanics calculation; and (5) bio-activity parameters obtained by biological tests.

Up to now, features are extracted by human intelligence in most drug design cases. For example, some medical chemists believe that for some molecules the structure of "three oxygen atoms spanning a certain angle" favors the bio-activity of some anti-tumor drugs. This is discovered by human brain, but not by computers. Molecular structure is very complicated and it is often described by a large number of design parameters. The difficulty is how to choose the right set, or a reduced subset, of design parameters that correlate the bioactivity to the structure of drug molecules for the best result. It seems that such empirical rules can also be discovered by effective computer software that implements a few powerful criteria for feature selection and reduction.

Data Separation

The data separability criteria, implemented in the MasterMiner software, are rather useful in selecting key factors that influence the bioactivity of a drug. People often use nonlinear regression in drug design. MasterMiner software has been proved by real-world examples to be very useful in simplifying the selection of nonlinear terms in regression. It has been compared favorably against other popular software products, since it uses far less terms in mathematical models, and produces lower PRESS (prediction residue error squared sum) value (< 0.3) in data modeling.

Hansch analysis provides another standard for new drug screening. It describes the "affinity" of an organic compound toward "water" and "oil-like liquids". Since a drug usually needs to diffuse through human body before it arrives at the focus of infection, molecules of drugs must be soluble both in aqueous medium and oil-like medium. Hansch analysis uses water-octanol distribution coefficient of a drug compound to describe the ability of drug diffusion in human body.

Two critical questions need to be answered in any drug design: (1) how to find the relationship between these diffusion coefficients and bio-activity, and (2) how to find the relationship between the molecular structure and the distribution coefficient in water and octanol? These questions can be answered by data mining and modeling solutions offered by MasterMiner.

III. MREC - A Hyperspace Data Mining Method

MREC (Map RECognition by hidden projection) is a

novel approach to statistical pattern recognition and it outperforms the classical PCA (principle component analysis), Fisher and PLS methods. It is equally applicable to many nonlinear problems ranging from chemistry and materials analysis and designs to pattern recognition to general optimization. MREC methodology includes (1) data separation by a *hidden geometric transform*, (2) feature selection by *data geometric pattern* ("one-sided" or "inclusive" type), and (3) building model that reduces a complex nonlinear problem to a set of simple linear models in sub-spaces.

MREC Background

Statistical pattern recognition methods are based on computerized recognition of m-D graphs (or their 2-D projections) of sample distribution in m-D space. Independent variables (features) influencing the model are used to span an m-D space. If one can describe samples of different classes as points with different colors in the space, a mathematical model can be obtained that describes the relationship (regularity) between target functions and features. Unlike the regression methods (linear, nonlinear, logistic regression, etc.) or the neural nets that provide *quantitative* solutions, MRE can provides *semi-quantitative* and *qualitative,* as well as *quantitative* solutions. This is advantageous because real-world data exhibit strong noise, and quantitative models would be *too precise to* represent them. The PCA-based regression builds linear models without data separation (see Fig-1), whereas MREC regression first tries to separates data, and then builds more realistic models from a *reduced set* of data (see Fig-2).

Data Separation

The data separability test of MREC is designed to investigate the possibility of separating data from different populations or clusters in the hyperspace. Building a model for a non-linear problem is possible only if the data set is *separable.*

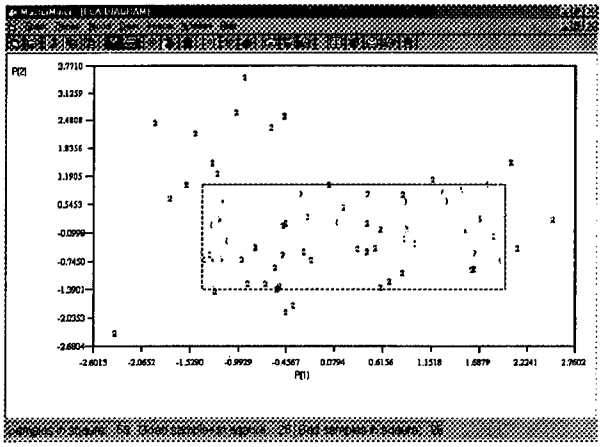


Fig-1. No data separation by PCA.

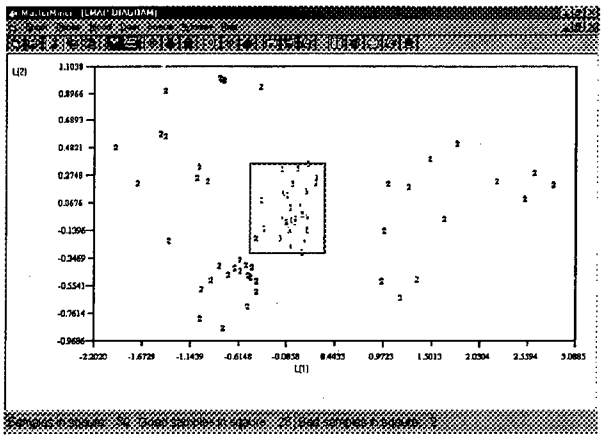


Fig-2 Good separation just after 1 projection by MREC

At each iteration, MREC chooses the "best" projection map with *maximum separation* from a series of *hidden projections*, and discards those samples outside the optimal zone (see the red box in Fig-2). After each projection, samples of class "1" (red) are automatically enclosed by a "tunnel" (two projections are shown to form an "auto-square" in Fig-3), and *a reduced data set* is formed that contains only samples within this "tunnel." Then a second MREC is performed on this reduced set to obtain the next "best" projection to further separate data into classes. After a series of such

projections, a complete (close to 100%) separation could be realized, and the resulting data set is used to build an *accurate* model. The physical meaning of MREC is explained by Fig-3, where each "auto-square" represents two "tunnels" in the original m-d space, and several such tunnels would form a *hyper-polyhedron* in space. This hyper-polyhedron, enclosing all "1" (red) but no (or few) "2" (blue) samples, defines an *optimal* zone in the m-d space. MREC has been shown to be much more powerful than various regression methods. When data separability is not good, two reasons are possible: (1) the data are too noisy, and (2) the form of the optimal zone (region of "1" samples) is complicated and cannot be described by a single *convex* hyper-polyhedron. An effective approach to the second type of inseparability is by "local view" treatment whereby one can cut a multidimensional space into several *convex* subspaces to achieve better data separability in each of the subspaces.

Back Mapping

Since MREC transforms data from the original measurement space into a number of orthogonal sub-spaces, one needs to back map the transformed data into the original feature space to derive mathematical models for practical use. Two methods, called linear and non-linear inverse mapping (LIM, NLIM) or PCBs (principal component back-mapping) [8], have been developed whereby a point in a low-dimensional principal component subspace is back-projected to the high-dimensional space of the original features. Let **X** be a training set with n samples and m features, and **Y** the sample set in the PC (principal component) space corresponding to **X** in the original feature space, with **Y = XC, and C = {C*1*, C*2*, ... C*m*}**. The columns of **C** are the eigenvectors of the covariance matrix **D,** with **D** = $\mathbf{X}^T\mathbf{X}$. The 2-d (Cu,Cv) sub-space of PCs is defined as the *main map* where samples are assumed to be completely classified. Let P represent an unknown sample point in the main map, and it is described by (ypu, ypv). In general, P is expected to be an optimal sample if all its neighbors are also optimal samples. To back transform P to the original space, i.e., to find $\mathbf{X}^*_p$, one needs determine its boundary conditions; otherwise, an uncertain solution will occur.

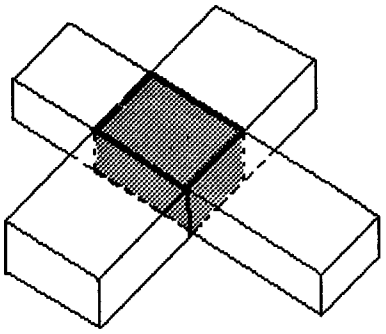


Fig-3 An "auto-square" formed by two "tunnels."

In non-linear inverse mapping (NLIM), denote dpj as the distance from sample p to all known samples in the subspace defined by the principal component coordinates, (u, v), and dpj is the same distance as in the original space. A non-linear algorithm is used to compute $\mathbf{X}^*_p$ by minimizing a cost function **E** [8].

In linear inverse mapping (LIM), besides the 2-d (Cu, Cv) subspace of PCs, there exists an (m-2)-dimensional subspace of PCs consisting of Ci (i = 1, 2, ..., m, for $i \neq u, v$), since **C** is derived from **D** (m x m). The projection of a point p, described by (ypv, ypv) in the main map, is determined by ypi (i = 1, 2, ..., m and $i \neq u, v$) in the (m-2)-dimensional subspace, and a set of *linear* equations can be obtained as $y_{pk} = \sum_{j=1}^{n} c_{jk} x_{pj}$, where k = *1, 2, ..., m*. These equations can be solved for the parameters of point p. This linear inverse mapping will always produce an exact solution. Fig-4 and Table-1 show one example of the PCB algorithm where a set of linear equations (inequalities) are obtained from (red) class "1" samples inside the auto-box and 100% data separation is achieved. These

equations constitute the mathematical model we sought.

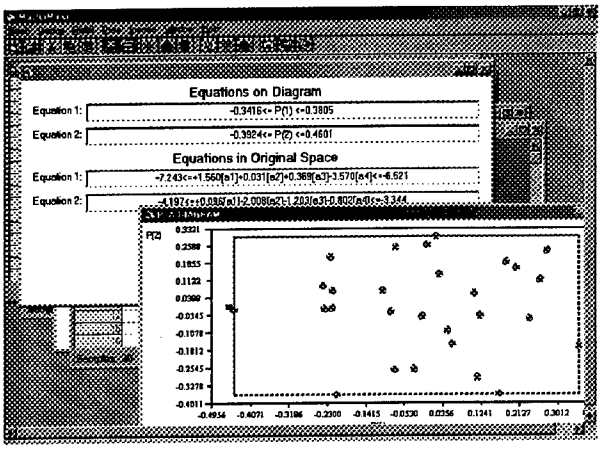


Fig-4 Modeling result of MREC with 100% separation.

Feature Reduction

The rate of data separation, **R**, is defined as **R** = (1-N2/N1), where N_1 and N_2 are the number of "1"and "2" samples inside the hyper-polyhedron respectively. If > 70%, the separability is "acceptable," otherwise it is "unsatisfactory." **R** is used as a criterion to reduce features – a feature can be removed if **R** remains the same after removing it. In practice, **R** has been used to reduce feature number by 1/3 to 1/2.

Concave Polyhedron

Since MREC only forms a *convex hyper-polyhedron*, it may not separate data when they form a *concave* polyhedron in the space. In this case, the BOX method, shown in Fig-5, offers a powerful solution whereby samples of class "2" are cut off from the polyhedron so that all samples inside are of type "1."

IV. Example - Fermentation of Glutamate

In glutamate fermentation, glucose solution is added continuously every 30 minutes into the biochemical reaction in a fermentation tank. There are four targets that need to be optimized simultaneously in the fermentation process. They are:

1) Conversion rate (from glucose to glutamic acid) – should be high
2) Productivity – should be high
3) Yield – should be high
4) Fermentation time period – should be short

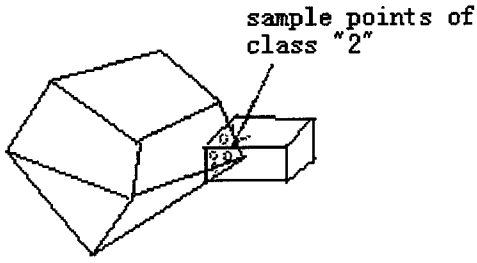


Fig-5 Principle of the Auto-Box method

Table-1 MasterMiner™ Example: modeling by PCB.
all inequalities in original space:

+7.842<=0.465[a1]+24689[a2]+0.254[a3]+3.115[a4]<=+6.658
-0.045<=+1.089[a1]+5.236[a2]-0.907[a3]-1.317[a4]<=+05528
-6.243<=+3.560[a1]+5.031[a2]+3.369[a3]-5.570[a4]<=-6.321
-9.197<=+0.996[a1]-2.087[a2]-8.203[a3]-0.902[a4]<=-4.345
-3.447<=+1.801[a1]-4.041[a2]+3.015[a3]-6.572[a4]<=-7.235
-7.661<=+4.076[a1]+6.567[a2]-0.780[a3]-7.747[a4]<=-0.786

where [a1], a[2], a[3] and a[4] are original features. The Auto-Box on the right covers all red points, showing 100% data separation.

Features that Influence Yield

A fermentation process may take 20-30 or more hours. Measurements of glucose feed would form a long time-series of data which are divided into several segments, each having tens of features. To compress information, these segments are averaged to produce a representative segment with a number of useful features. At the same time, the composition of the gas

in the biochemical reactor is detected by a gas sensor, and glucose concentration in liquid phase is measured by a glucose sensor.

Data Conditioning

Data of each batch (tank) was considered as one sample. Data from 80 production batches were used as training set. To simplify data processing, data of each fermentation period, which is usually 30 or more hours long, was divided into three segments:

1st segment - data from 0 to 12 hours
2nd segment -data from 12 to 24 hours
3rd segment - data from 24 hours to end

Some of data, including PH value, ventilation rate and temperature, were then averaged in each of the three segments to generate the averaged features for each segment. Data samples were further classified into three classes according to the range of these features:

- Class 1: the glucose-to-glutamate conversion rate is larger than 50%, and the fermentation time is less than 34 hours.
- Class 2: the glucose-to-glutamate conversion rate is less than 49.5%, and the fermentation time is larger than 34.5 hours.
- Class 3: the rest.

Feature Selection:

Features selected by MasterMiner include (1) *operation parameters*, such as glucose feed amount, tank temperature, ventilation (air flow rate), PH value of the liquid in the tank, and etc., and (2) *physical-chemical data* (as time series data) of liquors, such as glucose concentration in liquid phase from a glucose sensor, OD value (OD is an optical property of sugar, e.g., glucose), gas phase composition of chemicals in the biological reactor, and etc. Since a change in OD value indicates a change in substance concentration or composition, the OD value is a useful parameter for monitoring the process. The complete set of features for the overall fermentation process is listed below:

No.	Feature Explanation
X1	transparency of liquor
X2	glucose concentration at starting point
X3	PH of liquor
X4	PH value at the final stage of germ plantation
X5	increase (change) in OD value
X6	PH value in first segment (averaged)
X7	PH value in second segment (averaged)
X8	PH value in third segment (averaged)
X9	average ventilation rate (m^3/min) in 1st segment
X10	average ventilation (m^3/min) in 2nd segment
X11	average ventilation (m^3/min) in third segment
X12	temperature in 1st segment
X13	temperature in 2^{nd} segment
X14	temperature in 3rd segment

Findings:

By PLS (partial least square) regression, it has been found that no linear relationships existed among features or between feature and target. It therefore seems impossible to change any single feature to improve the targets. It appears that we have to define the optimal zone directly in the multi-dimensional hyperspace. By MasterMiner software, at last we found that the optimal zone is near a hyper-plane. After comparing the operation data against this hyper-plane, we have adjusted four features {x4, x6, x7, x8} to slightly lower than original values, and other parameters slightly higher. The results were rather good.

After finding the boundary of the optimal zone, MasterMiner also offers an operational advisory to the fermentation technicians by adding and testing a number of *virtual* samples to the optimal zone. A test sample generated by MasterMiner that falls inside this optimal zone is considered as an optimal sample, and it

could be expected to lead to the simultaneous optimization of the 4 targets in fermentation. Three such optimal samples, called predicted values, are listed in the following table: (the mean value is used).

Table of optimal samples by MasterMiner

X1	X2	X3	X4	X5	X6	X7
.37	-.14	.47	-.09	-.078	-.11	-.05
.84	-.13	.30	.03	-.016	.09	-.04
.56	-.01	.18	-.14	-.026	.01	-.03

X8	X9	X10	X11	X12	X13	X14
-.24	24	20	18	-.27	1.34	.59
-.20	33	30	46	.39	1.08	.64
-.12	-3	6	71	-.15	.74	.35

These predicted values were in good agreement with the result obtained by another method whereby single features were adjusted to search for the best operational direction. However, some features, such as X2, were not in agreement in these two methods.

MasterMiner offered the following advisory on this fermentation production:

- PH value should be slightly reduced
- Temperature should be slightly increased.
- Ventilation rate should be slightly increased.

Results:

In real-world fermentation operations at the client site, the optimization of the four targets, i.e., conversion rate, productivity, yield, and fermentation period, has been achieved simultaneously by using the optimal samples obtained by MasterMiner. Results from 390 production batches (tanks) were averaged to give the following results:

- Conversion rate was increased by 2.90%,
- Yield was increased by 2.56%
- Productivity was increased by 1.45%
- Glucose was save by 1.43%.
- Profit generated: US $200,000 per year

The model obtained by MasterMiner has since been used in production by the client with satisfactory results.

V. Conclusions

This paper presents basic background on drug design and manufacturing, reviews technologies in drug processing, and proposes a hyper-space data mining methodology. Data mining offers an effective solution to biotech when combined with other pattern recognition and statistical methods. The most important role of data mining is the ability to separate data into different classes so that good models can be obtained to describe the relationship between a drug's structure and its bio-activity. The proposed data mining methodology consists of data separability, hidden projection, back mapping, feature selection and reduction, and model building. When data exhibit linearity, a quantitative model can be built. When they exhibit non-linearity, a semi-quantitative or qualitative model can be obtained. Application examples have shown the efficacy of the proposed method in biotech.

References

[1] N. Chen. C. Li, R. Chen and D. Zhu, "N-factor pattern recognition method applied to process optimization and materials design," Proc. of the First International Conference on Multisensor-Multisource Information Fusion, Las Vegas, July 6-9, 1998.

[2] H. Liu Y. Chen and N. Y. Chen, "PCB method applied to material design - computer aided synthesis of a super-conductor," Journal of Chemometrics, Vol. 8, pp. 429-443, 1994.

[3] P. Geladi, B. R. : Analytica Chemica Acta, vol. 185, No. 1, 1986

[4] J. Holland, "Adaptation in nature and artificial intelligence," The University of Michigan Press, Michigan, 1975.

[5] Xiaohua Hu and Nick Cercone, "Data dining via discretization, deneralization and Rough Set Feature Selection," Knowledge and Information Systems: An International Journal, Vol. 1, No. 1.

[6] U. Fayyad, "Knowledge discovery and Data mining: towards a unifying framework," Proceedings of KDD-96, Menlo Park, CA, AAAI Press, 1996.

[7] S. Anand, "A general framework for data mining based on evidence theory," Data and Knowledge Engineering, Vol. 18, pp. 189-223, 1996.

[8] N. Chen, W. Wang and D. Zhu "Industrial diagnosis by hyper space data mining," AAAI Spring Symposium on Equipment Maintenance and Diagnosis Services, Stanford, CA, March 22-24, 1999

A Multi-Spectral Data Fusion Approach to Speaker Recognition

J. E. Higgins, R. I. Damper and C. J. Harris
Image Speech and Intelligent Systems Research Group
Department of Electronics and Computer Science
University of Southampton
Southampton SO17 1BJ, UK.

Abstract *This paper describes a multi-spectral, multi-source approach to the important problem of speaker identification. The wideband speech signal is filtered into several sub-bands and the output time trajectory of each is individually modeled by linear prediction cepstral coefficients. These individual models are then matched against reference data and the scores combined using the sum rule of information fusion, before using a k-nearest-neighbor rule to decide the identified speaker. Multi-spectral processing is shown to deliver performance improvements over wideband recognition. The optimal number of filters is found to be 16. These results are interpreted in light of the hypothesis that the multi-spectral approach solves the bias/variance dilemma of commonly manifest in systems that are trained on example data.*

Keywords: speaker recognition, feature-set construction, multi-spectral fusion

1 Introduction

Automatic speaker recognition (ASR), whereby a computer attempts to recognize an individual from their voice, is an important, emerging technology with many potential applications in commerce and business, security, surveillance etc. This paper is concerned with the application of modern data engineering techniques to the problem of ASR. The main idea here is the use of a multi-spectral approach, in which the wideband acoustic signal is pre-processed by a bank of bandpass filters to give a set of time-varying outputs – so-called *sub-band* signals. Because these time trajectories vary slowly relative to the wideband signal, the problem of representing them by some data model should be simplified. A major goal for this paper is to test if this is so, and if so, to determine the optimal number of sub-bands. Since we now have several time trajectories to consider rather than just one, the question arises of how to (re)combine or fuse the information in each, to reach an overall decision about speaker identity.

The remainder of the paper is organized as follows. Section 2 provides a motivation for research into recognition. Section 3 introduces the multi-spectral aspect of the recognition system and includes fuller discussion on the possible benefits to an identification system. In section 4, the component parts of the baseline multi-spectral system which provides the foundation for this research are described in turn. Finally, section 6 concludes with discussion of the issues raised by multi-spectral recognition and some possible avenues of future work.

2 Speaker Recognition

Recognition can be divided into speaker *verification* and speaker *identification* tasks, each of which may in turn be *text-independent* or *text-dependent* [1, 2]. In verification, there is an *ab initio* claim about speaker identity, and the aim is to determine if a given utterance was produced by the claimed speaker. This is done by testing the model of the claimed speaker against the utterance, comparing the score to a threshold, and deciding on the basis of this comparison whether or not to accept

the claimant. In identification, there is no *ab initio* identity claim, and the system must typically decide who the person is, or that the person is unknown.

In text-independent recognition, there are no limits on the vocabulary employed by speakers. This is in contrast to text-dependent recognition, where the presented utterance must be from a set of predetermined words or phrases. As text-dependent recognition only models the speaker for a limited set of phonemes in a fixed context, it generally achieves higher recognition rates than text-independent recognition, which must model a speaker for a variety of phonemes and contexts.

Speaker recognition is an example of biometric personal identification [3]. Biometric techniques based on intrinsic characteristics (such as voice, finger prints, retinal patterns) have an advantage over artifacts for identification (keys, cards, passwords) because biometric attributes cannot be lost or forgotten. Biometric techniques are generally believed to offer a reliable method of identification, since all people are physically different to some degree. Automatic speaker identification and verification are often considered to be the most natural and economical methods for avoiding unauthorized access to physical locations or computer systems [1]. Thanks to the low cost of microphones and the universal telephone network, the only cost for a speaker recognition system may be the software.

In this paper, we are primarily interested in text-dependent identification. Success depends on extracting and modeling the speaker-dependent characteristics of the speech signal which can effectively distinguish one talker from another.

Figure 1 shows the structure of a typical, simple identification system. In general, identification consists of five steps:

- digital speech data acquisition
- feature extraction
- pattern matching
- identification decision
- enrollment to generate reference models of each speaker

Initially, the acoustic sound pressure wave from an unknown speaker is transformed into an analog signal by a microphone or telephone handset. The analog signal is then passed through an anti-aliasing filter before being sampled to form a digital signal by an analog-to-digital converter.

In feature extraction, each frame of speech (typically spanning 10–30 ms of the speech waveform) is mapped into a multidimensional feature space creating a sequence of feature vectors $\mathbf{x}_i$. This sequence is compared to existing speaker models, created during the enrollment step, by pattern matching, resulting in a match score z_i for each of the speaker models. The match score gives an indication of the similarity between the sequence of vectors and the models of all the known speakers. The last step consists of a decision as to speaker identity. Before use, speakers must *enroll* on the system. During enrollment, speaker models are created for all authorized users and stored for later reference during identification.

3 Multi-Spectral Processing

In a seminal and influential paper, Allen [4] popularized the earlier notion of Harvey Fletcher that the decoding of speech signals by humans is based on decisions in narrow frequency bands that are processed independently of each other. Decisions from these frequency bands are combined such that the global error rate is equal to the product of the band-limited error rates within the independent frequency channels. This means that if any frequency band yields a zero (or low) error rate then the resulting global error rate would also be zero (or very low), regardless of the error rates of the remaining bands. While this has come to be known as the Fletcher-Allen principle, Allen himself refers to it as "the Stewart-Fletcher multiindependent channel model" (p. 572). He further characterizes the approach as "across-time" rather than the more usual "across-frequency" processing (p. 575) typified by template matching in automatic speech recognition. In this paper, we will refer to this as *multi-spectral processing*.

The positive benefits of this new approach to speech recognition are starting to be investigated

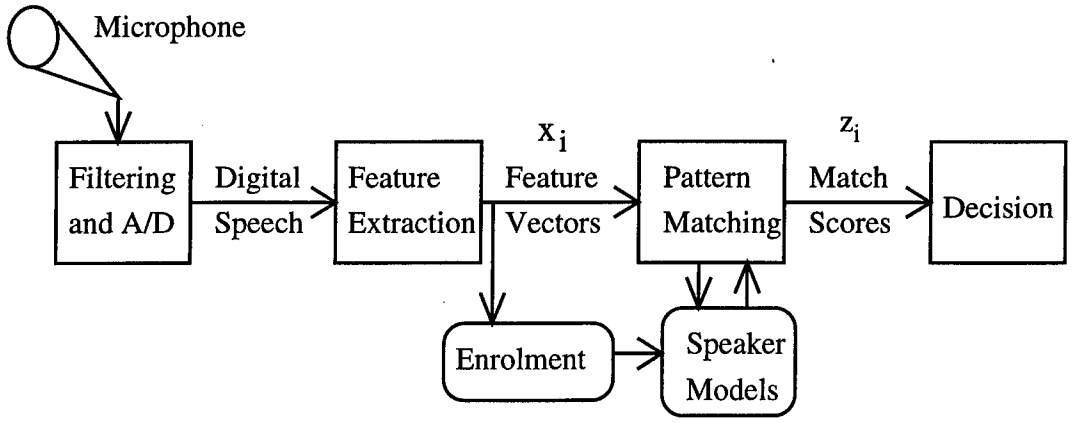


Figure 1: Block diagram of a typical speaker-identification system.

and reported [5, 6, 7, 8]. There are several cogent reasons why it might also profitably be applied to *speaker* recognition:

- The deleterious effect of narrow-band noise may be reduced. If noise only affects some frequency bands, then the remaining (clean) bands should carry sufficient information to allow the correct decision still to be reached. This follows from the (idealized) Fletcher-Allen principle according to which only one error-free band is required for correct recognition.

- Certain bands may contain more speaker-specific information than others. Weighting these to emphasize their contribution to the overall score should lead to better recognition rates. In fact, some bands might be better for some speakers than others, so that speaker-specific weighting during the information fusion – or (re)combination – stage may be possible. Note, however, that this assumes a form of fusion in which weighting can be sensibly done. (If, for instance, combination is by multiplication of scores, then weighting has no effect.)

- Successful recognition critically depends on building good speaker models from the training data. Data modeling, however, is subject to the well-known *bias/variance dilemma* [9]. According to this, models with too many adjustable parameters (relative to the amount of training data) will tend to overfit the data, exhibiting high variance, and so will generalize poorly. On the other hand, models with too few parameters will be over-regularized, or biased, and so will be incapable of fitting the inherent variability of the data. Multi-spectral processing offers a practical solution to the bias/variance dilemma by replacing a large, unconstrained data modeling problem by several smaller (and hence more constrained) problems. Empirical support for this notion in the specific context of speaker recognition comes from the work of Reynolds [10], who writes: "giving too much spectral resolution will degrade performance by modeling spurious spectral events or introducing too many parameters to be trained" (p. 642).

There are, however, several practical issues to be resolved before these advantages might be realized:

- The number, width and location of the frequency bands must be optimized. Sub-bands designed for speech recognition may not be suitable for speaker recognition: it may be that the frequency division should best be

done on a speaker-specific basis for speaker recognition.

- Some knowledge is required of which bands contain the most speaker-dependent information. The scores from these bands might then be emphasized to improve recognition.
- The features to be used for recognition must be decided. Again, features designed for speech recognition may not be suitable for speaker recognition [2]. It is also possible that features which are appropriate for wideband speaker recognition are less so for multi-spectral processing.

To date, relatively few workers have studied this problem. In the conference literature, [11], [12], [13] and [14] have all presented empirical results which confirm that worthwhile performance advantages can be gained from multi-spectral processing in speaker recognition. Taken together, however, these prior works do not cover anything like the full range of implementation options, so that many of the aforementioned questions remain open. Further, there is still only a rudimentary understanding of multi-spectral processing – and precisely how it delivers performance improvements – from a theoretical perspective.

4 Identification System

This section describes the different components that make up the identification system.

4.1 Database

The text-dependent Millar database from British Telecom was specifically designed and recorded for text-dependent speaker recognition studies. It consists of 43 male and 14 female native English speakers saying the digits *one* to *nine*, *zero*, *nought* and *oh* 25 times each. Recordings were made in five sessions spaced over three months, to capture the variation in speakers' voices over time which is one of the most important aspects of speaker recognition [15]. The speech was recorded digitally in a quiet environment using a high-quality microphone, and a sampling rate of 20 kHz with 16 bit resolution. The database was also made available at an 8 kHz sampling rate. In this version, the speech has been band-passed to telephone quality and then downsampled. Only this latter version was used.

For the experiments, 12 male speakers were used saying the word *seven*. The first two sessions (i.e. 10 repetitions of *seven*) were used as references and the remaining three sessions (15 repetitions) were used for testing.

4.2 Sub-Band Processing

The wideband signal was split into various numbers of sub-bands. Filters were simple second-order Butterworth, spaced on the psychophysical mel scale [16], covering the frequency range up to 3,600 Hz. There are many possible features that can be extracted from a speech signal, e.g. fundamental frequency, formant frequencies, and linear predictor (LP) coefficients. For recognition purposes, it is important to use a feature set that maximally discriminates between speakers. In this research, the feature set is based on cepstral coefficients. Cepstral analysis is motivated by, and was designed for, problems centered on *voiced* speech [17] but also works well for unvoiced sounds. Cepstral coefficients have been used extensively as the features in speaker recognition [18, 19]. This is because a simple recursive relation (see below) can be used to transform the LP coefficients into cepstral coefficients.

The time trajectories in each sub-band were modeled using an analysis frame of 20 ms, Hamming windowed and overlapping by 50%, and 12th order linear prediction [20]. These were then used to create cepstral coefficients via the recursion described by Atal [21]. That is, the LP cepstrum (or pseudo-cepstrum) is used, rather than the original (power or complex) cepstrum which would be obtained from Fourier analysis.

4.3 Pattern Matching

A popular method of pattern matching in speaker recognition systems uses 'templates'. The input signal is represented as a series of feature vectors that characterize the speech of a particular

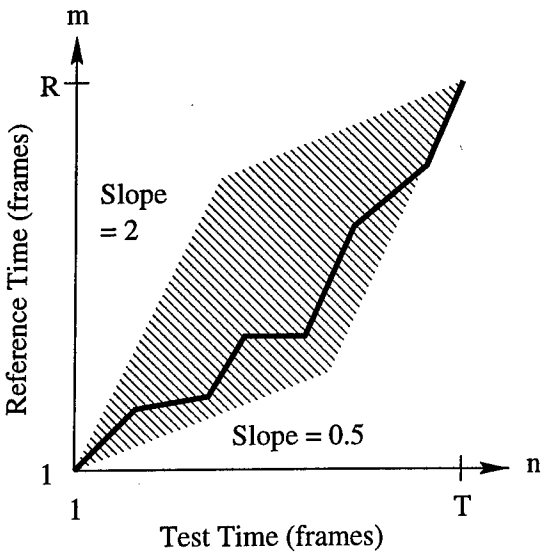


Figure 2: Typical DTW plot, illustrating the optimal warp path mapping the test time axis n into the reference time axis m.

speaker [22]. This time-ordered set of features constitutes the template. Even if the same speaker utters the same word on different occasions, the duration changes each time with nonlinear expansion and contraction. Therefore, any template matching algorithm needs to be able to cope with this: we use the popular technique of dynamic time warping (DTW) because of its ability to handle nonlinear time scale variations. It combines alignment and distance computation through a dynamic programming procedure [23]. It is normal to use the Euclidean distance measure when working with cepstral coefficients. Figure 2 depicts the DTW procedure schematically.

4.4 Fusion

Kittler, Hatef, Duin, and Matas [24] recently developed a common theoretical framework for combining classifiers which use distinct pattern representations. They outlined a number of possible combination schemes such a product, sum, min, max, and majority vote rules, and compared their performance empirically using two different pattern recognition problems. Kittler et al. found that the sum rule outperformed the other classifier combination schemes. This surprised them, because the statistical assumptions underlying this rule are stronger than, say, those for the product rule and it is clear that these assumptions do not hold well.

To explain this empirical finding, they investigated the sensitivity of various schemes to estimation errors. Their analysis showed that the sum rule is the most resilient to estimation errors, so almost certainly explaining its superior performance. Accordingly, the sum rule is used, at least initially, for combination purposes in this research while recognizing that this is one area which could benefit from further research by investigating other rules and methods of combination.

4.5 Decision Rule

There are 15 test utterances per speaker, each of which is matched to the 10 reference utterances for all 12 speakers – a total of 120 comparisons. These are then ranked (closest matches first) and the k-nearest-neighbor rule applied with $k = 5$. That is, the speaker maximally represented among the top five ranking matches is declared to be the identified person.

5 Results

To investigate the benefits of multi-spectral processing, as well as answering the question of the optimal number of sub-bands, we have collected identification results as the number of filters varies from 2 to 24. For comparison, recognition was performed using the wideband (unfiltered) speech signal also. Figure 3 displays the results.

It is clear that a multi-spectral recognition system can perform better than one using just the wideband signal. Using the wideband spectrum, the system achieved 85% recognition rate. By contrast, the best-performing multi-spectral system, using 16 mel-spaced sub-bands, produced a recognition rate of 96%. This is a very considerable improvement.

Using a small number of filters (< 6)), performance was generally worse than the wideband system. The reason for this is currently unknown, but we conjecture that too much spectral energy is removed by the filterbank, i.e. the regions of overlap between adjacent filters are too wide. Conversely,

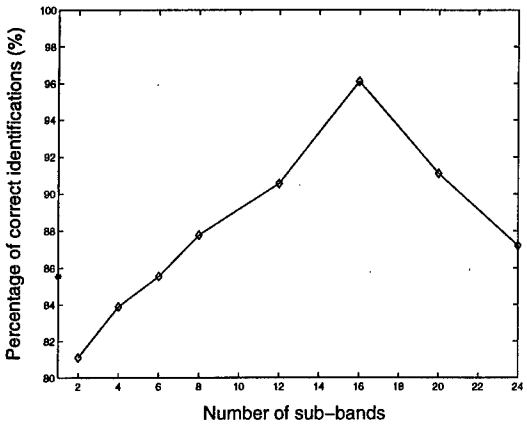


Figure 3: Percentage of correct identifications for different numbers of mel-spaced sub-bands (* indicates wideband).

it is possible to have too many filters. Performance reduces when there at 20 filters or more. We attribute this to attempting to fit too many parameters in the data models describing each speaker.

From the perspective of time-frequency duality, it seems intuitively reasonable that there should be some such trade-off. With a small number of filters, we will be attempting to fit the time trajectories too closely, having only a few parameters to do so. With a large number of filters, we will be attempting to fit the frequency distribution too closely but with more parameters than can be reliably estimated from the data. There is an interesting convergence with Allen's comment [4]: "It has been reported ... that 10 bands is too few, and 30 bands gives no improvement in accuracy over 20" (p. 572).

6 Discussion and Conclusions

The results highlight the advantage of a using multi-spectral approach to speaker recognition. We believe that the approach offers a practical solution to the bias/variance dilemma manifest in trainable systems, and so leads to improved data modeling. The problem of fitting parameters to training data is constrained by requiring them to be more or less uniformly deployed across frequency. Although multi-spectral processing increases performance, there is a limit to how many sub-bands can be used before performance starts to decrease. Here, it seems that 16 is the optimal number. This finding is interpreted in data-modeling terms as reflecting an attempt to fit too many parameters for the available training data. By contrast, the wideband approach (or use of a small number of filters) attempts data modeling with too few, unconstrained parameters.

The traditional approach to identification has been to base the development of recognition systems on *a priori* knowledge. The prior knowledge has been applied to such things as choosing the type and number of feature parameters and determining the pattern matching method to use. Current speaker identification systems produce reasonable results but still lack the necessary performance if they are to be used routinely by the general public. Furui has listed 16 open questions about speaker recognition which need to be addressed if performance is to be improved. One of these concerns the selection of feature parameters: commonly cepstral (or delta cepstral) coefficients. These are employed principally (or only) because they are familiar from their use in speech recognition. Hence, they may not optimally discriminate between different speakers. From this perspective, there seems much to be gained from automatic (data-driven) selection of features – and other architectural parameters.

Future work will look at possible ways of implementing a data-driven strategy for number and placement of the filters, and for automatically determining the type and number of features to be used in each sub-band. We will also explore other combination schemes and will extend the work to speaker verification. Finally, we propose a direct test of our hypothesis of improved data modeling, by varying the number of parameters fitted in the different filtering scenarios.

7 Acknowledgements

The authors would like to thank Robert Finan for kindly providing some of his code and the UK Engineering and Physical Sciences Research Council for their support.

References

[1] J.P. Campbell. Speaker recognition: A tutorial. *Proceedings of the IEEE*, 85(9):1436–1462, Sep 1997.

[2] S. Furui. Recent advances in speaker recognition. *Pattern Recognition Letters*, 18(9):859–872, September 1997.

[3] G.R. Doddington. Speaker recognition – identifying people by their voices. *Proceedings of the IEEE*, 73(11):1651–1664, Nov 1985.

[4] J. B. Allen. How do humans process and recognize speech? *IEEE Transactions on Speech and Audio Processing*, 2(4):567–577, October 1994.

[5] H. Bourlard and S. Dupont. A new ASR approach based on independent processing and recombination of partial frequency bands. In *Proceedings of International Conference on Spoken Language Processing (ICSLP '96)*, pages 426–429, Philadelphia, PA, 1996.

[6] H. Hermansky, S. Tibrewala, and M. Pavel. Towards ASR on partially corrupted speech. In *Proceedings of 4th International Conference on Spoken Language Processing, ICSLP 96*, volume 1, pages 462–465, Philadelphia, PA, 1996.

[7] S. Tibrewala and H. Hermansky. Subband based recognition of noisy speech. In *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, volume II, pages 1255–1258, 1997.

[8] H. Hermansky and S. Sharma. TRAPS – Classifiers of temporal patterns. In *Proceedings of 5th International Conference on Spoken Language Processing, ICSLP 98*, Sydney, Australia, 1998. Paper 615 on CD-ROM.

[9] S. Geman, E. Bienenstock, and R. Doursat. Neural networks and the bias/variance dilemma. *Neural Computation*, 4:1–58, 1992.

[10] D. A. Reynolds. Experimental evaluation of features for robust speaker identification. *IEEE Transactions on Speech and Audio Processing*, 2(4):639–643, October 1994.

[11] L. Besacier and J. Bonastre. Subband approach for automatic speaker recognition: Optimal division of the frequency domain. In *Proceedings of 1st International Conference on Audio- and Visual-Based Person Authentication (AVBPA)*, Crans-Montana, Switzerland, 1997.

[12] R. Auckenthaler and J. S. Mason. Equalizing sub-band error rates in speaker recognition. In *European Speech Communication Association (ESCA) Conference, Eurospeech '97*, pages 2303–2306, Rhodes, Greece, 1997.

[13] P. Sivakumaran, A. M. Ariyaeeinia, J. A. Hewitt, and J. A. Malcolm. An effective sub-band based approach for robust speaker verification. *Proceedings of the Institute of Acoustics*, 20(6):69–72, 1998.

[14] P. Sivakumaran, A. M. Ariyaeeinia, and J. A. Hewitt. Sub-band speaker verification using dynamic recombination weights. In *Proceedings of 5th International Conference on Spoken Language Processing, ICSLP 98*, Sydney, Australia, 1998. Paper 1055 on CD-ROM.

[15] R.A. Finan, A.T. Sapeluk, and R.I. Damper. Imposter cohort selection for score normalisation in speaker verification. *Pattern Recognition Letters*, 18(9):881–888, September 1997.

[16] E. Zwicker and E. Terhardt. Analytical expressions for critical-band rate and critical bandwidth as a function of frequency. *Journal of the Acoustical Society of America*, 68(5):1523–1525, 1980.

[17] J.R. Deller, J.G. Proakis, and J.H.L. Hansen. *Discrete-Time Processing of Speech Signals*. Macmillan, New York, 1993.

[18] S. Furui. Cepstral analysis technique for automatic speaker verification. *IEEE Transac-*

tions on Acoustics, Speech and Signal Processing, ASSP–29(2):254–272, April 1981.

[19] D.A. Reynolds and R.C. Rose. Robust text-independent speaker identification using Gaussian mixture speaker models. *IEEE Transactions on Speech and Audio Processing*, 3(1):72–83, Jan 1995.

[20] J. D. Markel and A. H. Gray. *Linear Prediction of Speech.* Springer-Verlag, Berlin, Germany, 1976.

[21] B.S. Atal. Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification. *Journal of the Acoustic Society of America*, 55(6):1304–1312, June 1974.

[22] J.M. Naik. Speaker verification: A tutorial. *IEEE Communications Magazine*, pages 42–48, January 1990.

[23] D. O'Shaughnessy. *Speech Communication: Human and Machine.* Addison-Wesley, 1987.

[24] J. Kittler, M. Hatef, R. Duin, and J. Matas. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):226–239, March 1998.

The artificial sensor head: A new approach in assessment of human based quality.

Peter Wide
Ivan Kalaykov
Applied Autonomous Sensor Systems laboratory
Department of Technology and Science
Örebro University
Sweden
Email: peter.wide@ton.hoe.se
ivan.kalaykov@ton.oru.se

Fredrik Winquist
The Swedish Sensor Center
Linköping University
Sweden
Email: frewi@ifm.liu.se

Abstract

The design of a new electronic sensor head using artificial senses is described. The system involves a chewing process that mimics the human behavior. Before entering the test sample in the artificial mouth, the sensor system uses a video camera to identify the test object. The artificial sensor mouth is then measuring the crushing and chewing process of the samples, mixing it with saliva like liquid. In parallel it measures the aroma with an electronic nose, detect the chewing resistance and listens to the crushing sound. Further, the taste of the mixed solution from the sample is measured with an electronic tongue sensor.

To the amount of information received, we apply feature extraction analysis and a fuzzy clustering to assess the quality. By combining data from different artificial sensor systems into a single set of meaningful features, we achieve information that is of greater benefit than the aggregate of its contributing sensors. The combination of sensor data by fuzzy clusters has the aim to perform inferences that may be impossible from the single artificial sensors.

I. Introduction

The combining of data into more meaningful information refers to an essential technology in the problem of the information treatment to improve the quality of the sensing data. Data fusion uses various data sources to provide a better understanding of the phenomenon taken in consideration. The information proceeds usually from two types of sensor models [1], consolidating data from the same type of information [2], and in the second case, usually named multi-sensor data fusion, merging information from different and often complementary sensors to create a environmental based sensor model [3].
We have focused in this approach on a sensor model using combination of data information from five different sensor systems measuring the quality of a food product, and more specific an integration of multiple sensing data in human quality applications.
A number of single artificial sensors have been described in different human based quality related applications, electronic nose [4, 5], electronic tongue [6] and in the chewing process [7, 8]. Further, a combination of the information from artificial smell and taste sensor systems into a merged opinion has been reported [9].

II. Human analysis of quality estimation

The operator in an industrial food process, for example potato chips plant, continuously analyzes the dynamical process properties, e.g. temperature, humidity, time, sound, etc. as well as the specific product quality like color, size, taste, smell, along the process line. In the laboratory, tests are regularly

made to measure parameters as concentration of salt, color, water content and percentage of fat, a visual inspection is provided as well.

III. The artificial approach

There is a change in the attitude within measurement technology towards the way of and how to collect process information. Instead of measuring single parameters, in many cases it has become more desirable to get information of attributes such as quality, condition or state of a process. Due to different available techniques of extracting human like features from a huge information flow mimicking the human perception, there is a growing interest in the concept of artificial senses.

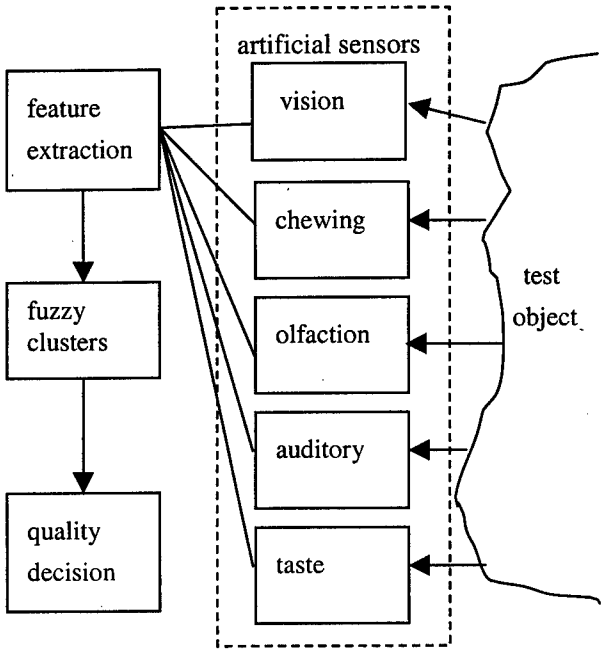


Figure 1. The model of the artificial sensor head.

Although the combination of artificial senses most likely increases the performance of the measurement, articles in this area are lacking. In [7] and [8] an electronic mouth is described. In [9] and [10] original sensor fusion methods based on human opinions about smell and taste and measurement data from artificial nose and tongue sensors is presented.

In this paper we propose to combine artificial sensors into an electronic sensor head approach containing a number of sensor systems that measure essential properties of the tested object, as shown in figure 1.

A. The electronic head

A special artificial mouth with hearing and vision capabilities, i.e. an artificial sensor head, is designed and tested in the laboratory. Stationary robot arm feeds the mouth with test samples after the vision cameras has recorded the object. In the mouth, with a temperature of 37 °C kept inside, a crushing process takes place that is similar to human chewing. In parallel the crushing sound and chewing resistance are recorded and the developed aroma pumped from the mouth to the measuring electronic nose. The chewed pieces of the sample object are further mixed with saliva like fluid and the electronic head spits the rest into a cell where the electronic tongue is measuring the taste. After this moment the cell containing the sample test is cleaned up and the system is ready to measure a new sample. The result is presented for visual acceptance on the monitor indicated by the mode of a happy or sad human face.

The electronic head system is controlled by a PLC (Programmable Logic Control) pneumatic system and interacting with the measurement PC operating under LabView software.

B. The artificial electronic nose

The sensor array consists of a number of selective semiconductor metal oxide (Taguchi) type sensors, obtained from Figaro Engineering Inc., Japan. The measurement interface was built at the laboratory. Gas samples are pumped from the mouth cell by a membrane pump at a flow rate of approximately 500 ml/min and injected into the sensor chamber, where the sensors are placed in a row. The injection of gas samples is performed at given time intervals by the opening a valve. Thus, samples are injected during

the chewing process.

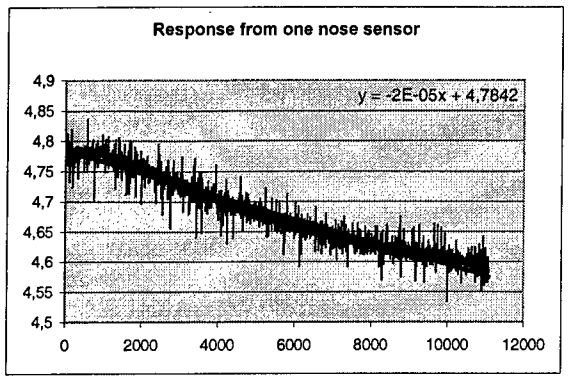


Figure 2. A calculated feature from one of the nose sensors.

Further, the sensor system collects and preprocesses the data from the sensor array. Each of the four-sensor data measurement used in this approach contains 4 variables to be further analyzed where one parameter is the derivative shown in figure 2.

C. The artificial tongue

The principle for measurement was based on pulse voltammetry carried out in a standard six-electrode configuration. In this method, current transients due to onset of a voltage pulse are measured, giving information concerning both amount and type of charged molecules and of redox active species. The electronic tongue, consisting of a six working electrode system also contains an auxiliary electrode, and a reference electrode. The six working electrodes are composed of gold, iridium, palladium, platinum, rhenium and rhodium. The whole configuration is placed in a 150-ml measurement cell. The electrical current transient responses are measured by a potentiostat connected to the measurement PC via an A/D converter.

The recorded voltammograms are based on large amplitude pulse voltammetry (LAPV). A measurement sequence starts by applying a potential of 800 mV during 0.5 sec. The voltage is then set to 0 at the instant, when the applied potential is decreased by 100 mV, and the cycle starts again. A measurement sequence covers 11 cycles, which results in a final pulse value of -200 mV, see figure 3.

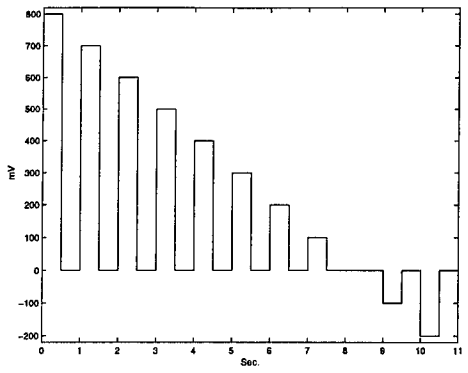


Figure 3. A series of pulses applied to a tongue electrode during a measurement sequence.

A typical recording of a full measurement over all electrodes is shown in figure 4. The sample rate is set to 20 Hz and only the amplitudes which has shown to contain sufficient information, namely from the first, second and last samples in each 0.5-second interval, are used in this experiment. Each electrode measurement is characterized by 66 samples; hence, a total tongue measurement comprises 396 samples.

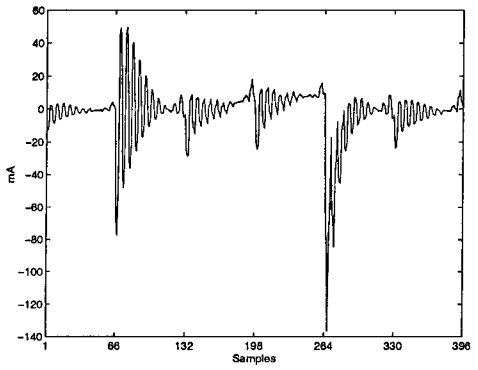


Figure 4. A typical sequence of samples in the complete tongue measurement.

D. The chewing resistance

A signal from a force sensor connected to the pneumatic driving system of the mouth is also measured. The chewing process behaves similar to the human; i.e. an initial crushing is applied to the test object before the final chewing starts. The shape of this signal reflects the deformation process.

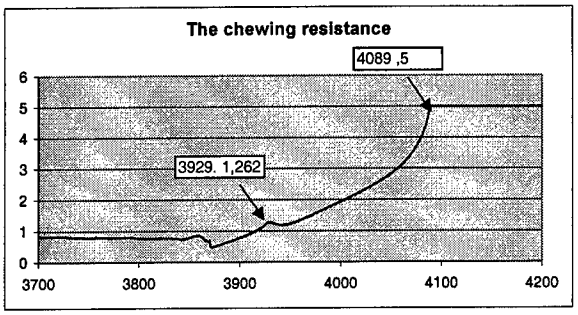


Figure 5. An example curve describing the chewing resistance.

E. The vision system

A color vision camera is used to indicate visual properties of the samples. The picture also directs the computer system to start the measurement procedure by opening the mouth. Information about color, shape and size of the sample object is measured.

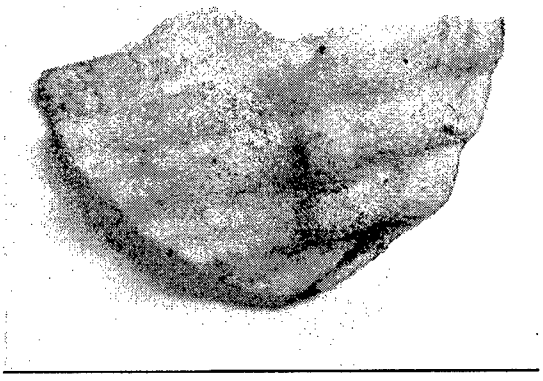


Figure 6. An image from the vision system.

F. The sound system

A microphone is embedded in the mouth construction to measure the sound from the chewing process, then a standard frequency analysis is provided on the records. Differences in the spectral power density (shift of the maximum, level of the horizontal asymptote), the amplitude spectrum (change in the parameters of the envelope curve) and the complex spectrum drawn on the complex plane (varying size and shape of the spot) show they can be used as characterizing parameters of the quality. Illustrative diagrams are presented in figure 7.

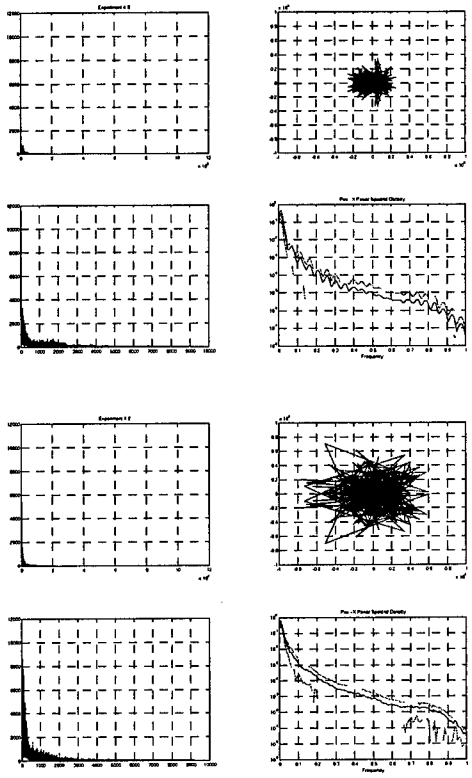


Figure 7. Crashing sound frequency domain patterns of two samples with different properties

IV. Sensor fusion and pattern recognition

This section describes an industrial problem and a proposed solution using sensor fusion. The main reasons for this experiment is twofold, first to test the

artificial sensor head in a long time test and to investigate its usefulness as an industrial on-line sensor system. The aim is also to investigate if it is possible to improve the results by combining the different sensor systems in a real world application.

The sensor fusion process in this approach is defined as a pattern recognition method, which gathers opinions from a number of task specific classifiers. Each one of the classifiers is specialized for one perception specific related property of the artificial sensors: smell, taste, vision, hearing and mouth feeling. The fusion method is then combining the features, into a single, more reliable one.

The industrial related problem considered here is taken from the food process industry: classification of different qualities of potato chips including classification of the aging processes at room temperature. The recognition task for a given sample of potato chips is to identify the type of chips and classify its quality within four different grades.

A. Feature extraction

Each sensor measurement contains different amount of information. Data reduction must be performed to form an efficient classifier. Generally, this task may be troublesome due to the problem of modeling the physical process that generates the measurements. Therefore, our approach is to compute some features of the sensor signal, and by fuzzy cluster analysis, determine its information content. For example, in case of tongue data by using the score- and loading-plots in principal component analysis (PCA) it emerged that the *range* in each of the *first* two cycles and the *last* cycle at each electrode should contain sufficient information. The range is a relative measure and should be robust with respect to bias in different measurement setups. The complete tongue pattern vectors, $\mathbf{x}_{taste}$, from a complete tongue measurement then consists of 18 elements.

The nose measurements accommodate a manageable amount of data, which has proven to contain relevant information [11]. The nose data is obtained from 4 sensors, each one with a unique gas sensitivity pattern. Thus, the pattern vector from the artificial nose, $\mathbf{x}_{smell}$, consists of 16 features.

Features from the other sensors are constructed in a similar way giving a unique vector from each artificial sensor pattern. The final pattern vector can then be formed as

$$\mathbf{x} = [\mathbf{x}^T_{smell}, \mathbf{x}^T_{taste}, \mathbf{x}^T_{vision}, \mathbf{x}^T_{chewing}, \mathbf{x}^T_{sound}]^T$$

B. Pattern Recognition

We propose here a system that is closely related to the human way of estimating quality parameters. It is based on training of a fuzzy classifier and then using it in estimating how the sample object taken from the conveyor belt fits to the already established classes of production quality. For that purpose we make lots of experiments with potato chips, the quality of which is grouped in 8 classes, depending on the 3 levels salt content, existence/absence of spices and freshness. Large amount of measurements of the full pattern vector then is stored, preprocessed and used for training of the fuzzy classifier. Other set of measurements is used for test and verification of the quality of the classifier. Two fuzzy classification algorithms, namely fuzzy c-means and Gustafson-Kessel [12, 13], are applied. The experimental data gave approximately the same results, so any of them can be used.

V. Conclusions and further work

An artificial human related sensor system evolved from human perception measurement is proposed, with emphasis on issues of complex quality determination and focusing on food measurement based on the human ability to quality estimation. The paper presents basic background in feature extraction

and fusion of artificial sensor systems. In this concept of an artificial perception head we extract feature from the following sensors:

- chewing resistance
- electronic nose
- electronic tongue
- vision system
- auditory system

Applying multivariate analysis methods can show that the sensor units evaluate the properties of experimental samples in different way. However, by combining types of sensors and features from the different sensor data it is possible to reduce the amount of data to be processed in the classification phase. To achieve that, the system has also to be learned to estimate the discriminating abilities of each sensor with respect to the quality assessment of particular product. Then, a proper combination of sensor data can contribute to performing inferences that may not be possible from single sensors. This aspect has to be further developed in more comprehensive and self-contained system, able to include other human based capabilities and enhanced fusion techniques.

ACKNOWLEDGMENT

This work has been done at the AASS laboratory and at the Swedish Sensor Center at the universities of Örebro and Linköping respectively. We thank OLW, Pär Westergren and Camilia Modig for contributing with industrial knowledge and test samples. We also thank Joakim Arnell and Andreas Fransson for contributing to the development of the demonstrator hardware and software. This work is part of research programs sponsored by KK-foundation in Sweden.

References

[1] López-Orozco, J.A., de la Cruz, J.M., Sanz, J. and Flores, J., "Multisensor fusion of environment measures using Bayesian Networks", *Proc. of the International Conference on Multisource-Multisensor Information Fusion, Las Vegas,* USA, 1999, pp 487-493.

[2] Winquist, F., Holmin, S., Krantz-Rülcker, C., Wide, P. and Lundström, I., "A hybrid electronic tongue", submitted to *Analytica Chimica Acta*

[3] Wide, P., Saffiotti, A., and Bothe, H., "Environmental Exploration: An Autonomous Sensory Systems Approach", invited article to *IEEE I&M Magazine*, September 1999.

[4] Gardner, J.W., and Barlett, P. N., "A brief history of electronic noses, *Sensors and Actuators B* 18-19, 1994, pp. 211-220.

[5] Wide, P., Winquist, F. and Driankov, D., "An air-quality sensor system with fuzzy classification.", *International Journal of Measurement Science Technology*, vol. 8, 1997, pp 138-146.

[6] Winquist, F., Wide, P. and Lundström, I., "An electronic tongue based on voltammetry", *Analytica Chimica Acta*, 357, 1997, pp. 21-31.

[7] Andersson, Y., Drake, B., Granquist, A., Halldin, L.,Johansson, B.,Pangborn, R. M. and Åkesson, C., "Fracture force, hardness and brittleness in crisp bread with generalized regression analysis approach to instrument –sensory comparisons", *Journal of Texture Studies*, vol 4. 1973, pp 119-144.

[8] Winquist, F., Wide, P., Eklöv, T., Hjort, C. and Lundström, I., "Chripsbread quality evaluation based on fusion of information from the sensor analogies to the human olfactory, auditory and tactile senses". Submitted to the *International Journal of Food Process Engineering.*

[9] P. Wide, F. Winquist, P. Bergsten, and E.M. Petriu, "The human-based multisensor fusion method for artificial nose and tongue sensor data, *IEEE Transaction on Instrumentation and Measurement*, vol. 47, No. 5, October 1998. pp 1072-0177

[10] Winquist, F., Wide, P. and Lundström. I., " The combination of an electronic tongue and an electronic nose", accepted paper, *International Journal - Sensor and Actuators B.*

[11] Lundström, I., Winquist, F.Hörnsten, E. G. and Sundgren, H., "From hydrogen sensors to olfactory images-twenty years with catalytic field-effect devices," *Sensors and Actuators B*, vol. 13, 1993, pp. 16-23.

[12] Gustafson, D. and W.Kessel, "Fuzzy clustering with a fuzzy covariance matrix", *Proc. IEEE CDC*, San Diego, 1979, pp.761-766.

[13] Roger Jang, J.-S. and N. Gulley, "*Fuzzy Logic Toolbox*", Mathworks, 1995.

Improving Resolution of Seismic Sections Based on Method of Information Fusion with Well-log Data[1]

Ke Zhang, Xuegong Zhang, Yanda Li
Dept. of Automation, Tsinghua University, Beijing 100084, P.R. China
email: zk@simba.au.tsinghua.edu.cn

Abstract - *A method of extrapolating the high frequency information from the well-log into the cross-well seismic sections was proposed in this paper. Employing the Sort-Time Fourier Transform (STFT) and an average coherence coefficient that was defined in this paper, the local correlation between two adjacent seismic traces was calculated in the frequency domain. Then depending on the proposed transfer function we transfer the high frequency information of a well-log into its nearby traces, and then from one high resolution trace into its adjacent traces one by one to get the desired high resolution seismic section.*

Keywords: information fusion, short-time Fourier transform, high-resolution seismic signal processing

1. Introduction

In today's petroleum industry, one important source of information for oil/gas exploration and production is seismic data, and improving the vertical resolution of seismic data is an important task in seismic signal processing. With the development of modern signal processing technology, the resolution of seismic sections has already been improved greatly. Yet the requirement from the industry is far from fully satisfied. As the information in the seismic records is limited, it is very difficult to further enhance their resolution depending solely on signal processing techniques. Fusing other information about the strata in seismic data processing is an effective way to solve this problem.

Usually in oil exploration bore-holes or wells will be drilled at locations of particular interests in a survey. In this case logging sondes will be placed in the bore-hole and pulled upwards, measuring rock velocity and density of the subsurface rocks as well as other geophysical parameters. Since a reflection coefficient is the difference in acoustic impedance of two layers, over their sum, and acoustic impedance is given by velocity times density, it is thus possible to construct a series that is close to the true reflective coefficient sequence from well logs. It has the highest possible vertical resolution on the spot of well. In other words, it has the highest frequency extent. How to use its high frequency information to enhance the resolution of seismic sections across the well is a newly problem in the study of seismic signal processing.

In general, it is reasonable in geophysics to assume that there exists good local correlation between most adjacent seismic traces. And, through some processing there can also exist good local correlation between the well-log and its nearby traces. This observation is the basis of improving resolution of seismic sections using information of well-logs. Based on this observation Luo and Li proposed an initial idea of extrapolating high frequency information from well-logs into seismic section using the short-time Fourier transform[1]. But there exist many aspects that need to be further improved so that the method can be practicable.

In this paper, we proposed a new method for enhancing the resolution of cross-well seismic sections by fusing high-resolution well-log information into the processing. We first extract the high frequency components of the well-logs and extrapolate them to the near-well seismic trace, to get a new higher resolution trace. Then

[1] This work is supported by NSFC, the National Science Foundation of China.

the high frequency information of this new trace is transferred to its adjacent traces. This procedure is executed trace by trace so that a whole new seismic section of higher resolution can be finally derived. The key of this method is to extrapolate the high frequency information from a well-log to a near-well trace, and from one trace to its adjacent trace. In order to estimate the local correlation between two adjacent traces exactly, we calculate the short-time Fourier transform (STFT) of a high-resolution trace and its adjacent low-resolution trace respectively. Then the high frequency component of STFT amplitude of the adjacent trace is extended or is modified by some transfer function, which is designed based on the local correlation of the two traces in frequency domain. From the modified STFT amplitude of this adjacent trace, a new seismic trace whose resolution has been enhanced can be constructed with inverse STFT. By this processing, the high frequency information contained in well-log data is effectively fused into the seismic data, and at the same time, the original information of seismic data is well preserved at the same time.

Experimental results on both synthetic seismic data and field seismic data proved the effectiveness of our new method.

2. The Model

In seismic signal processing, a well-known discrete convolution model is

$$y_t^k = r_t^k * w_t + n_t^k = \sum_l r_{t-l}^k w_t + n_t^k, \quad k=1,2,\ldots, \qquad (1)$$

where y_t is the seismic signal, r_t is the primary reflective coefficient sequence of strata, w_t is the seismic wavelet, and n_t is additive noise. The subscript t is two-way reflection time, and superscript k is the number of seismic trace. In general situation, the series $\{y_t\}$ may be analyzed as (second-order) stationary and furthermore, from detailed investigations in [2], may be taken as having a zero mean and having a square summable auto-covariance sequence so that their spectra exist in the usual mean square sense.

In equation (1) we assumed that the seismic wavelet of each trace is same in one section. This assumption is appropriate for actual post-stack sections[3]. The problem of enhancing the resolution of seismic section is to compress the width of seismic wavelet, or to extend the width of frequency band of sequence $\{y_t\}$. For convenience, we ignore the item of noise in (1) the following discussion.

If we denote the series $\{r_t^0\}$ as the reflective coefficient that is taken from the data of well-logs, we can use a wavelet $\{w_t^0\}$, whose frequency band can be intentionally selected, to make a synthetic seismic trace. We denote it as $\{v_t\}$:

$$v_t = r_t^0 * w_t^0 = \sum_n r_{t-n}^0 w_t^0 . \qquad (2)$$

If we select a higher frequency wavelet in above equation, we can get a higher resolution synthetic seismic trace, contrarily, we can also get a lower resolution synthetic seismic trace. We denote the higher frequency wavelet and lower frequency wavelet of synthetic seismic data as $\{w_t^{0H}\}$ and $\{w_t^{0L}\}$ respectively, and the synthetic seismic trace corresponding with them is denoted as v_t^H and v_t^L.

3. Estimation of Local Correlation between Adjacent Traces

Let the near trace of $\{v_t\}$ be $\{y_t^1\}$. In spite that $\{y_t^1\}$ and $\{v_t\}$ are coming from different physical methods, they all represent information of strata at the same local position. Because the change of strata is relatively slow in the horizontal direction, $\{y_t^1\}$ and $\{v_t\}$ can be considered to be highly correlated. Actually, any two adjacent traces are highly correlated in most local ranges of traces.

Let $s_{vy}^0(\tau)$ be cross-correlation function between well synthetic trace $\{v_t\}$ and its near trace $\{y_t^1\}$. It is defined as

$$R_{vy}^0(\tau) = E[v_t\, y_{t+\tau}^0] \approx \frac{1}{N}\sum_{k=0}^{N-1-\tau} v_t\, y_{t+\tau}^1, \qquad (3)$$

where N is the length of series. For computing a

local correlation between $\{v_t\}$ and $\{y_t^1\}$, we modify above equation with a window function $\{p_t\}$. Then,

$$R_{vy}^0(l,\tau)\approx\frac{1}{N_p}\sum_{k=0}^{N_p-1-\tau}v_{l,t}y_{l,t+\tau}^1, \quad (4)$$

where $v_{l,t}=v_t p_{t-l}$,

$y_{l,t}^1=y_t^1 p_{t-l}$,

and l is time, N_p is the width of the window function, which has a short-time duration.

Imitating equation (2), the auto-correlation of $\{v_t\}$ and $\{y_t^1\}$ are

$$R_v^0(l,\tau)\approx\frac{1}{N_p}\sum_{k=0}^{N_p-1-\tau}v_{l,t}v_{l,t+\tau}^0, \quad (5)$$

$$R_y^0(l,\tau)\approx\frac{1}{N_p}\sum_{k=0}^{N_p-1-\tau}y_{l,t}^1y_{l,t+\tau}^1. \quad (6)$$

Their Fourier transforms $R_{vy}^0(l,\tau)$, $R_v^0(l,\tau)$ and $R_y^0(l,\tau)$ give the cross-spectral density function between $\{v_t\}$ and $\{y_t^1\}$, and auto-spectral density functions of $\{v_t\}$, $\{y_t^1\}$, which are denoted as $S_{vy}^0(l,f)$, $S_v^0(l,f)$, $S_y^0(l,f)$ respectively. From above spectral density functions, the coherence function $\gamma_{vy}^2(l,f)$, which measures the linear correlation between the components of $\{v_t\}$ and $\{y_t^1\}$ at frequency f in a local range, can be defined as

$$\gamma_{vy}^2(l,f)=\frac{|S_{xy}^0(l,f)|^2}{S_x^0(l,f)S_y^0(l,f)}. \quad (7)$$

The coherence function can be seen being normalized, and it is sort of a correlation coefficient in the frequency domain[4]. If $v_{l,t}=y_{l,t}^1$ (maximum correlation between $\{v_{l,t}\}$ and $\{y_{l,t}^1\}$), then

$$\gamma_{vy}^2(l,f)=\frac{|S_{vy}^0(l,f)|^2}{S_v^0(l,f)S_y^0(l,f)}=1. \quad (8)$$

On the other extreme, if v_t and $y^1{}_t$ are un-correlated, then $S_{vy}^0(l,f)=0$ and $\gamma_{vy}^2(l,f)=0$. We define an average coherence coefficient $\bar{\gamma}_{vy}^2(l)$ as

$$\bar{\gamma}_{vy}^2(l)=\frac{1}{N_f}\sum_{f=0}^{N_f-1}\gamma_{vy}^2(l,f), \quad (9)$$

where N_f is number of samples in frequency domain. The function $\bar{\gamma}_{vy}^2(l)$ measure the average value of local correlation between the components of $\{v_t\}$ and $\{y_t^1\}$ in frequency domain. It is only the function of time l. Alike the coherence function $\gamma_{vy}^2(l,f)$, if $v_{l,t}=y_{l,t}$ then $\bar{\gamma}_{vy}^2(l)=1$, and if $\{v_{l,t}\}$ and $\{y_{l,t}^1\}$ are un-correlated then $\bar{\gamma}_{vy}^2(l)=0$.

4. Transfer Function and the Information Transfer

The basic method of enhancing the resolution of seismic section in this paper is to realize the extrapolation of high frequency information from well-log synthetic trace to its nearby traces, and then from one higher resolution seismic trace to its adjacent lower resolution seismic trace one by one. This means that the processing of extrapolation should be in the frequency domain. On the other hand, since strata vary along the horizontal direction, the values of correlation in the different local segments of any two adjacent traces are different. From this property the extrapolation should be of the time-varying. STFT is a useful tool for analyzing and processing the problem of time-frequency. We employed it in our method.

Denote the STFT of series $\{v_t\}$ is $SF_v(l,f)$, then

$$SF_v(l,f)=\sum_{k=-\infty}^{\infty}v(k)p(l-k)e^{-j2\pi fk}. \quad (10)$$

The above equation can be rewritten in form of the amplitude spectrum $|SF_v(l,f)|$ and corresponding phase spectra $\Phi_v(l,f)$,

$$SF_v(l,f)=|\sum_{k=-\infty}^{\infty}v(k)p(l-k)e^{-j2\pi fk}|e^{j\Phi_v(l,f)}. \quad (11)$$

Replacing v with y in equation (11) we can get $|SF^0{}_y(l,f)|$ and $\Phi^0{}_y(l,f)$.

The transfer function is used to transfer the local frequency information from a higher

resolution trace to a lower resolution trace corresponding to the local correlation. In order to reserve the phase information of original section, which is important for farther processing, we only transfer the amplitude information. Therefore, the transfer function is defined as

$$H_{vy}(l,f)=1+\frac{\bar{\gamma}^2_{vy}(l)[|SF_v(l,f)|-|SF^0_y(l,f)|]}{|SF^0_y(l,f)|} \quad (12)$$

where $H_{vy}(l, f)$ relates to the average coherence coefficient $\bar{\gamma}^2_{vy}(l)$. Let $\{\hat{y}^1_t\}$ be the series whose resolution has been enhanced through the processing of high frequency information transfer from series $\{v_t\}$ to series $\{y^1_t\}$. The amplitude spectra of its *STFT* are denoted as $|SF_{\hat{y}}(l,f)|$. Then

$$|SF_{\hat{y}}(l,f)|=|SF^0_y(l,f)|\,H_{vy}(l,f) \quad (13)$$

Using (12), this can be written

$$|SF_{\hat{y}}(l,f)|=|SF^0_y(l,f)|+\bar{\gamma}^2_{vy}(l)\,[|SF_v(l,f)|-|SF^0_y(l,f)|], \quad (14)$$

which will form the basis of our extrapolating scheme. From (14), the process of frequency information transfer relates to the local coherence coefficient $\bar{\gamma}^2_{vy}(l)$. In extreme situations, if $\bar{\gamma}^2_{vy}(l)=0$, which means that $\{v_t(l)\}$ are un-correlation with $\{y^1_t(l)\}$, then $|SF_{\hat{y}}(l,f)|=|SF^0_y(l,f)|$, i.e. the frequency information isn't transferred completely; and if $\bar{\gamma}^2_{vy}(l)=1$, i.e. $\{v_t(l)\}$ have the maximum correlation with $\{y^1_t(l)\}$, then $|y^1_t|=|SF_v(l,f)|$, in which situation the frequency information is transferred completely.

5. The Procedure of Fusing Well-log Information into Seismic Traces

In section 3, we showed the correlation measurement of two adjacent traces in frequency domain. In section 4, we gave a definition of transfer function and put forward the method of transferring the high frequency information from a synthetic trace to its nearby traces. Based on those, the whole procedure of fusing well-log information into seismic trace is presented.

First, using the reflective coefficient sequence, which is from a well-log data, and two known wavelets, one frequency is higher and another is lower, we generate two synthetic seismic traces with the convolution model (2), i.e. $\{v^H_t\}$ and $\{v^L_t\}$, which have higher and lower resolution respectively. It is worth noting that the lower resolution synthetic seismic trace $\{v^L_t\}$ is only used to match its near trace $\{y^1_t\}$, so its wavelet should be as same as $\{y^1_t\}$. In the best way, this wavelet should be extracted from $\{y^1_t\}$.

Secondly, to calculate the local cross-correlation $R_{vy}(l,\tau)$ between the lower frequency synthetic seismic trace $\{v^L_{l,t}\}$ and its nearby trace $\{y^1_{l,t}\}$ with the equation (4). Because the seismic section has the slope line, the local cross-correlation between two adjacent traces may be not in the same time point. Therefore, we should find the maximum of the cross-correlation between $\{v^L_{l,t}\}$ and $\{y^1_{l,t}\}$ in a small time range $[-M, M]$. In this situation, the equation (4) is modified as

$$R^M_{vy}(l,\tau)\approx \max_{m\in[-M,+M]}\{\frac{1}{N_p}\sum_{k=0}^{N_p-1-\tau} v^L_{l,t}y^1_{l+m,t+\tau}\}, \quad (15)$$

where parameter M is selected according to the degree of slope of the line in seismic section. Beside the $R^M_{vy}(l,\tau)$, we should calculate the local auto-correlation function of $\{v^L_{l,t}\}$ and $\{y^1_{l+m_0,t}\}$, denote $R_v(l,\tau)$ and $R^0_y(l+m_0,\tau)$, with equation (5) and (6), where $m_0\in[-M, M]$ is the time point that make $R^M_{vy}(l,\tau)$ maximum in (15). After getting $R^M_{vy}(l,\tau)$, $R_v(l,\tau)$ and $R^0_y(l+m_0,\tau)$, Fourier transform them to get power spectrums, $S^M_{vy}(l,f)$, $S^0_v(l,f)$, $S^0_y(l+m_0,f)$, then using equation (9) to calculate the average coherence coefficient $\bar{\gamma}^2_{vy}(l)$.

Finally, transfer frequency information from the synthetic seismic trace $\{v^H_t\}$ to its nearby trace $\{y^1_t\}$ so that a new higher resolution trace will be got. We denote it by $\{\hat{y}^1_t\}$. For this purpose, we compute the STFT of $\{v^H_t\}$ and

$\{y_t^1\}$, get the amplitude spectrum $|SF_y(l,f)|$ and $|SF_y^0(l+m_0, f)|$, and the phase spectrum $\Phi_y(l, f)$ from (10). Then, from equation (14) the amplitude spectrum of STFT of $\{\hat{y}_t^1\}$, i.e. $|SF_{\hat{y}}(l,f)|$ is given. Using the following equation,

$$\hat{y}_t^1(l)=\frac{1}{2\pi p(0)}\sum_{f=0}^{2\pi}|SF_{\hat{y}}(l,f)|e^{j(2\pi ft+\Phi_y(l,f))}, \quad (16)$$

where l, $t=0,1,...,N_y$ we can reconstruct the seismic trace $\{\hat{y}_t^1\}$.

Replacing the $\{v_t^H\}$ with $\{\hat{y}_t^1\}$, and replacing $\{y_t^1\}$ with $\{y_t^2\}$, we repeat above procedure, the higher resolution trace $\{\hat{y}_t^2\}$ can be given. Imitating this the information of well-log can be transfer into the whole seismic section.

6. Experiments

In order to check the performance of the method that was proposed above, we first transfer the information from a well-log into its nearby trace using the simulated signals. Fig.1(a) is a simulated sequence of reflective coefficient that was taken from well-log data. Fig.1(b) is a synthetic seismic trace which is the convolution of a higher frequency wavelet and the sequence of Fig.1(a). Fig.1(c) is the trace near the well which was generated with reflective coefficient sequence that is showed in Fig.1(f). Using the method presented above, we transferred the higher frequency information from Fig.1(b) into Fig.1(c), the result or output trace is showed in Fig.1(d). Comparing Fig.1(c) with Fig.1(d), the resolution of original trace has been enhanced evidently. Fig.1(e) is the real high resolution trace corresponding with Fig.1(c), which is generated from convolution between a high frequency wavelet and correspondent real reflective coefficient sequence Fig.1(f). Comparing Fig.1(d) and Fig.1(e), we can find that result of the procedure of information transfer is quite similar with its real situation.

Another result of the experiment with a real-world seismic data is shown in Fig.2. Fig.2(a) is an original seismic section which are from certain area in China. Fig.2(b) is the processed seismic section using above method. In fig2(b) we inserted two high frequency synthetic seismic traces that were repeated to form several same traces, which are the source of the high resolution of seismic section. From fig.2(b) we can see that the resolution is much higher than the original seismic section, while the basic structures are well reserved.

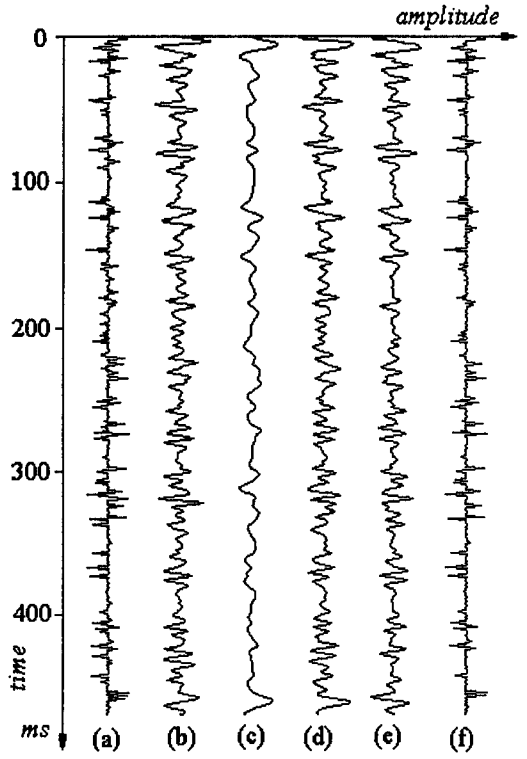


Figure1. Frequency information transfer with simulated signals of well-log and its nearby trace. (a) reflective coefficient taken from well-log data; (b) the higher frequency synthetic seismic trace; (c) the nearby trace; (d) the trace whose resolution has been enhanced; (e) the real high resolution trace corresponding with (c); (f) the real reflective coefficient corresponding with (c).

7. Conclusion

In this paper we defined the local average coherence coefficient, and proposed a new transfer function. Based on them a new method of fusing well-log information to enhance the

resolution of seismic section was given. Experiment examples showed its good effect.

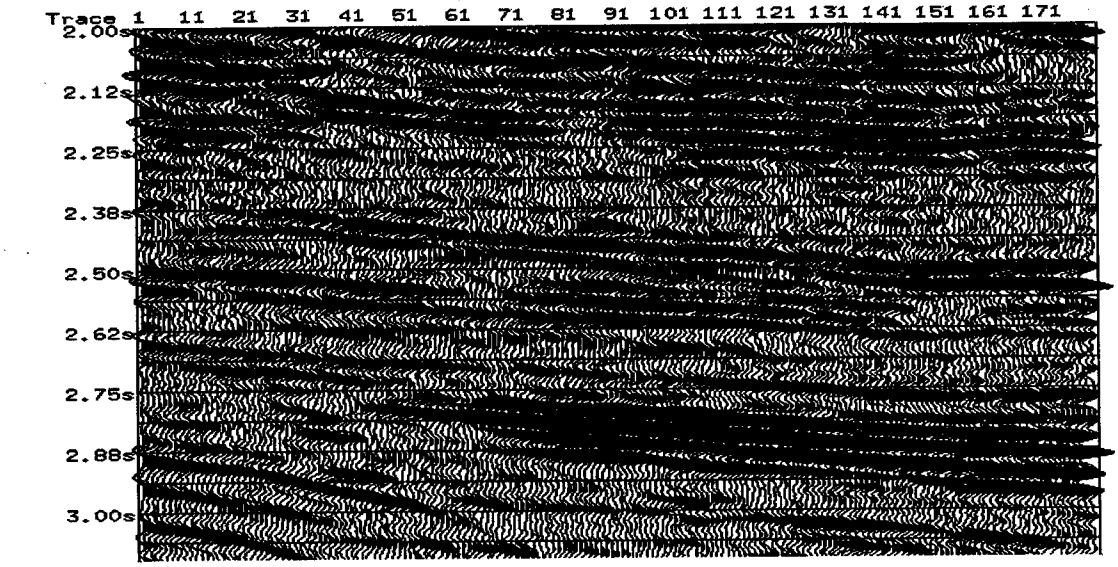
References

[1] H. Luo, A study on improving resolution of cross-well seismic section (in Chinese), Ph.D. Dissertation, Tsinghua University, 1994

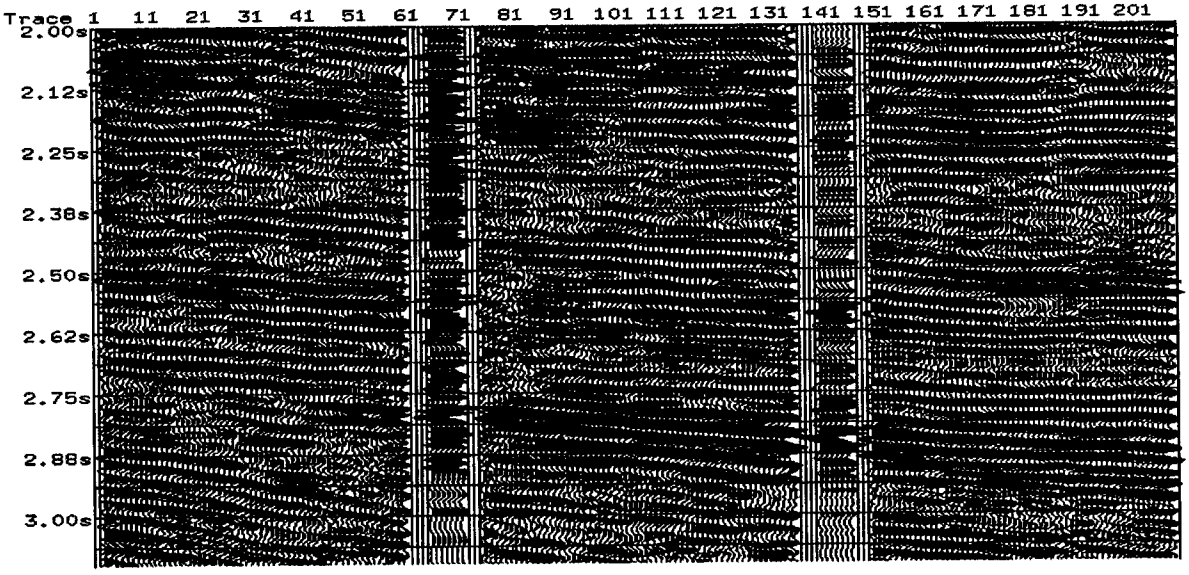
[2] A.T. Walden and J.W. Hosken, An investigation of the spectral properties of primary reflection coefficients, Geophys. Prosp., vol.33, pp. 400-435, 1985

[3] A.T. Walden and J.W. Hosken, The nature of the non-Gaussianity of primary reflection coefficients and its significance for deconvolution, Geophys. Prosp., vol.34, pp. 1038-1066, 1986

[4] R.G. Brown, Introduction to random signal analysis and Kalman filtering, John Wiley & Sons, Inc., 1983



(a)



(b)

Figure 2. Experiment with real-world seismic section (part).

(a) original seismic section; (b) processed seismic section with synthetic traces at well locations.

A New Structure of ESKD—Generalized Diagnosis Type Expert System Based on Knowledge Discovery*

Bing-ru Yang and Hai-hong Sun
Information Engineering School
Beijing University of Science and Technology, Beijing, China, 100083
Email: bingru.yang@bj.col.com.cn

Abstract - A new structure of ESKD (generalized diagnosis type expert system based on knowledge discovery system KD (D&K)) is first presented on the basis of KD (D&K)—synthesized knowledge discovery system based on double-base (database and knowledge base) cooperating mechanism. With all new features, ESKD may form a new research direction and provide a great probability for solving the wealth of knowledge in the knowledge base. This paper mainly advances the general structural frame of ESKD, describes some sub-systems among ESKD and emphasizes on dynamic knowledge base based on double-base cooperating mechanism. According to the result of demonstrative experiment, the structure of ESKD is effective and feasible.

Key Words: Knowledge Discovery, General Structure, Expert System, Dynamic Knowledge Base, Double-Base Cooperating Mechanism.

1 Introduction

Since 1965 when the first expert system DENDRAL, was developed by F. A. Feigenbaum to deduce the structure of molecule from chemical data, the expert system has developed rapidly and been used in many domains to produce great ecconomic and social benefit. But the further development of expert system met some difficulties such as poor knowledge, monotonous inference and poor ability for self-study. These caused the second bottleneck—insufficient knowledge during the research of expert system recently.

On the other hand, abundant knowledge in the knowledge base of diagnosis type expert system is crucial factor and difficult during the software developing. Presently, in the world the new method is "knowledge module method", namely the software of lower layer and quick as possible as is used to transit. Its worst quality is imperfect, and it needs prophase accumulation of longer time, then the knowledge base is made. Essentially, this adopt to a "blenching" method on the above crucial problem (whether the knowledge is abundant). From the "extended" and "radical" angle, can the crucial problem be resolved?

In accordance with above question, the article presents a generalized diagnosis type (i.e. a type of problem of seeking generalized cause-and-effect in wide field, including in fault diagnosis, intellect call center, credit card cheat and so on) expert system based on knowledge discovery (ESKD). Its theory is synthesized knowledge discovery system based on double-base (database and knowledge base) cooperating mechanism presented by us. It produces a very abundant dynamic knowledge base and corresponding integrated inference mechanism under many knowledge resources, kinds of knowledge fusion, multi-abstract levels and different knowledge layers. Therefore it especially fits for complicated big system and provides a valid path to produce the kernel technology on the structure of expert system. This system primarily improves the practical function of traditional expert system.

2 Generalized diagnosis system based on knowledge discovery (ESKD) (see fig.1)

* Project supported by Chinese National Natural Science Fund (69835001)

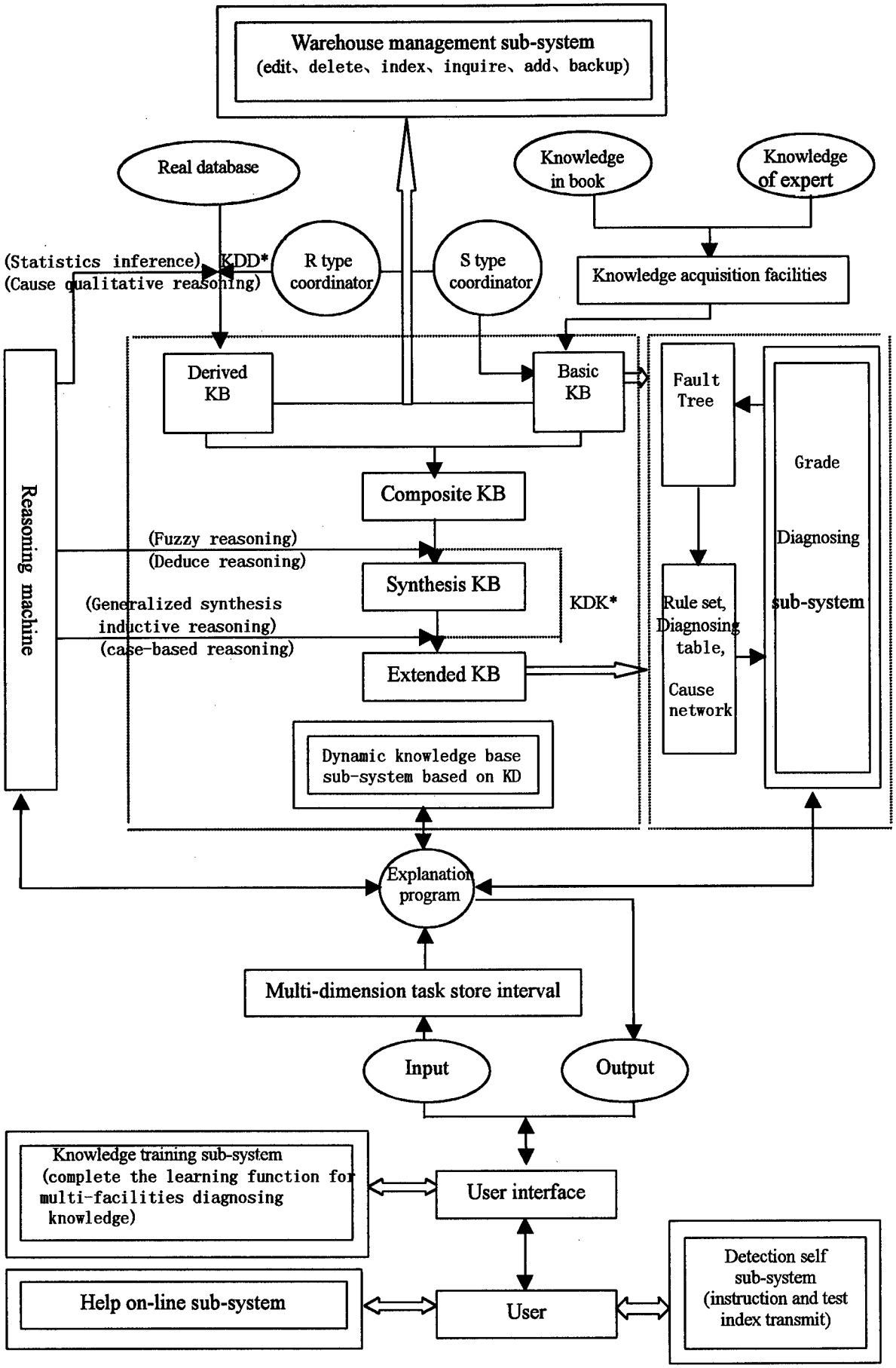


Fig.1 FESKD overall structure figure

It mainly includes the following modules:

2.1 Dynamic knowledge base sub-system based on knowledge base:

Essentially, it is a knowledge discovery system based on double-base (database and knowledge base) cooperating mechanism. As a result of knowledge discovery under different knowledge levels, the basic knowledge base directly constructed by expert experience and book knowledge is expanded continuously. Utilizing KDD*, each kind of inference mechanisms and KDK* under double bases cooperating mechanism, knowledge base sub-system which has the character of dynamic expansion is developed to manage Fuzzy uncertainty, random uncertainty and qualitative information. The cause-and-effect rules, discovered by this system, are used to modify the pristine fault tree, decision tree and example in knowledge base to fit for the solution of complex generalized diagnosis problem.

2.2 Knowledge training sub-system:

The system can not only be trained by professionals but also gain data directly by examples. It can gain the fault diagnosis knowledge of different type sets to adapt to different users.

2.3 Grade diagnose and decision sub-system:

Fault tree is used to put the whole facility to a set of tests to be determined whether there is a fault or not. These modules will be tested one by one if there is. When one module is found fault, rule base (cause network and fault diagnosing table) is used to test and
diagnostic internal module until the faulty point is found. Using correct resemble mechanism and knowledge in knowledge base, the system tests the facility and diagnoses whether the facility is normal or not. If the facility is not normal, the system will find the cause of the fault and provide solution according to decision tree.

2.4 Base management sub-system

It mainly manages real database, basic knowledge base and derived knowledge base. It can edit, delete, index, inquire, add and backup. It establishes good interface in Windows style. Users can realize expediently the operation of knowledge base and database.

2.5 Detection self sub-system

To avoid false diagnosis caused by the fault of testing hardware itself, the expert system will check the testing hardware itself by close-loop before operation.

2.6 Help on-line sub-system

This sub-system will help users use the system more effectively and get the help of corresponding information at any time.

3. Dynamic knowledge base system with double-base cooperating mechanism

3.1 General frame of dynamic knowledge base system:

Dynamic knowledge base experiences the promotion process of basis—derivation—integration—expansion. The process only finishes the first stage of discovery, i.e. the first abstraction level. The expansion knowledge base in the first abstract level is regarded as basic knowledge base in the second abstract level. The second abstract level will be finished in a process similar to each step of discovery in the first abstract level. Things like that continue. When cognition is developed and time and space environment are changed in different stages. Knowledge will constantly be enriched and promoted and cognition will deepen. So knowledge base is dynamic and in sequence (see fig.2).

The parallel structure can be applied to complicated system if the system is divided into several independent sub-systems (shown in fig.3).

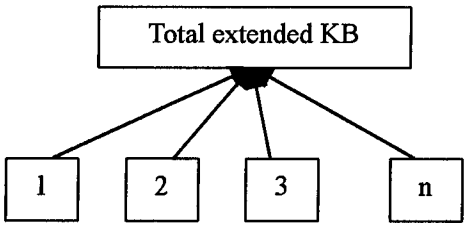


Fig.3 the parallel structure

Where 1,2,3,⋯,n are the developing processes of each independent sub-system. Each of them seems to be an abstract level in sequence structure, then they all integrate into the total extended knowledge base.

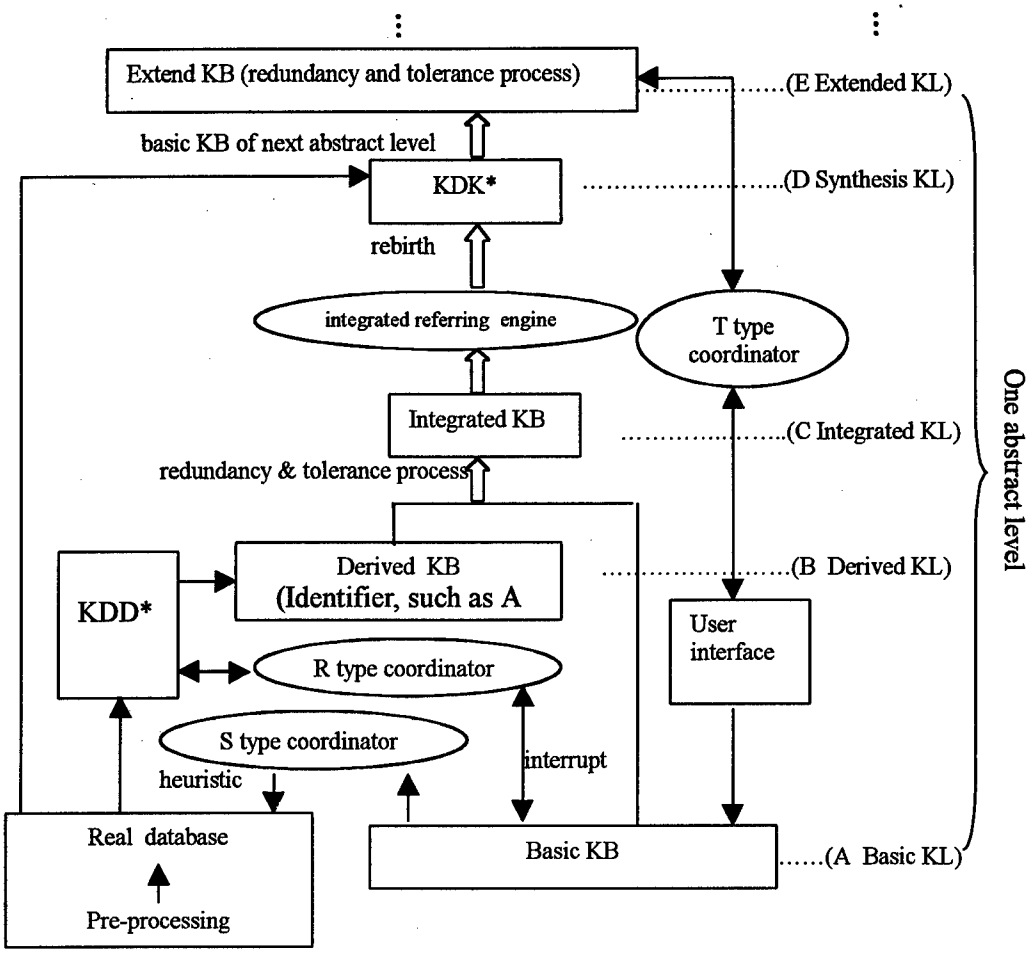


Fig.2 simplified structure in sequence

3.2 Double-base cooperating mechanism:

3.2.1 Basis Theory:

The dynamic knowledge base is established on double bases (database and knowledge base) cooperating mechanism. The large (basic) knowledge base is divided into several correlative sub-knowledge bases according to each domain; Meanwhile, the real database is divided into correlative sub-databases according to each domain. Thus the layers between knowledge nodes in knowledge base and data sub-class (structure) make a one to one mapping. The basis theory which is proposed by us is pan-homotopy conception and the following structure mapping theorem [1] [2].

Theorem (Structure Mapping Theorem):

Aiming at X, in the sub-database corresponding to sub-knowledge nodes, <E, F> of knowledge nodes and <F, D> of data sub-class (structure) are identical pan-homotopic type spaces.

This theorem presents the mapping of layers between knowledge nodes in the sub-knowledge base and data sub-class in corresponding sub-database (see fig.4).

On the basis of the research above, we can see that in the knowledge discovery system mathematical structure of database and knowledge base can be essentially come down to pan-homotopy category. Namely database is pan-homotopy category combined with data sub-type (structure) set and "mining path", which is called data mining category; and knowledge base is pan-homotopy category combined with knowledge nodes set and "reasoning arc", which is called knowledge reasoning category. Additionally more result about the isomorphy and restricting mechanism of knowledge reasoning category $C_R(E)$ in <E, F >and data mining category $C_D(F)$ in <F, D > are got, and "directional searching" and "directional mining process" are resolved

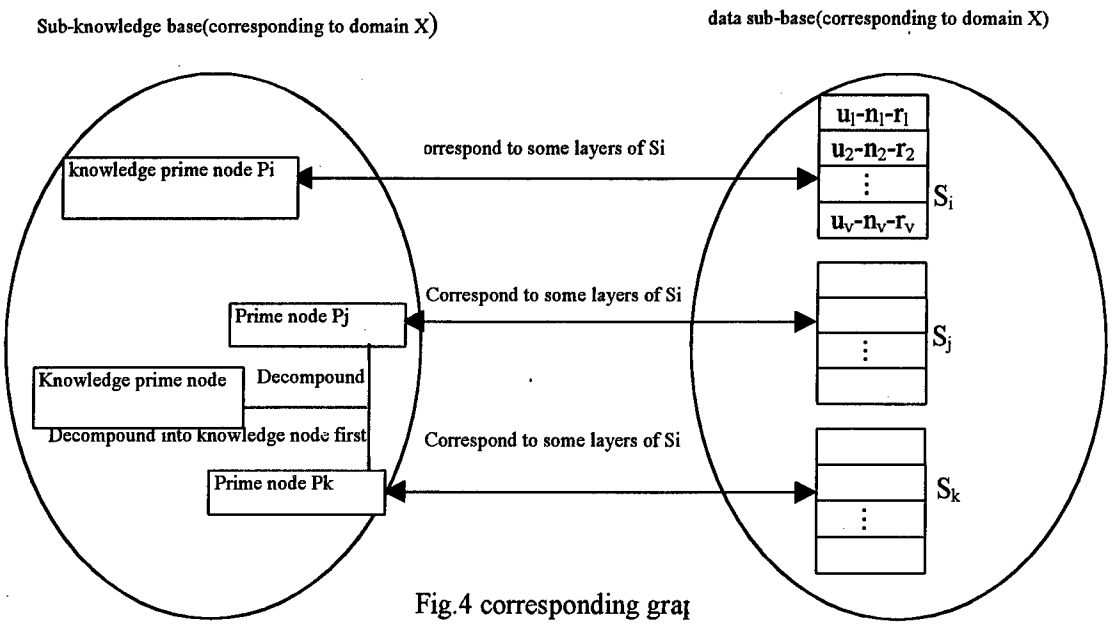


Fig.4 corresponding gra

3.2.2 The technological realization:

The technological realization of double-base cooperating mechanism is to construct R type and S type coordinator. The main function of R type coordinator is to "interrupt" the process of KDD and search whether there are repetition and controversy of the resulting rule in knowledge base after the rule (knowledge) is resulted from large amount of data in real database through setting focus. If there is repetition, the resulting rule is cancelled and returned to the beginning position in KDD. If there is controversy, the rule in knowledge base is thought right in general conditions, and the resulting rule is cancelled or set into knowledge base after the means of expanding premise is used to eliminate controversy. If there is neither repetition nor controversy, the process of KDD continues. Namely the result is set into knowledge base after evaluation. The main function of S type coordinator is to search irrelevant state of knowledge nodes in knowledge base under the principle of property on which knowledge base is established. Knowledge shortage is found. Data-class corresponding to real database uses heuristics and is activated to produce "directional mining process", The priority of "directional mining" is sorted according to relevant strength.

3.3.3 Significance

1) Besides discovering knowledge according to the factitious "interest", we proposed the new path of automatically enlightening directional mining according to "knowledge shortage" in basic knowledge base.

2) The mechanism greatly decreases "the evaluating quantity" after discovering hypothesis rule. 3) According to the above mechanism of "structure mapping", the searching space is greatly reduced and the mining efficiency is improved. 4) The mechanism rather easily resolves the redundancy and consistency problem in knowledge after new and old knowledge synthesized. 5) During the KDD process and the wide relation with basic knowledge base, KDD regarded as a open system improves and optimizes its structure, process and running mechanism.

4 Conclusions

Comparing the structure of generalized diagnosis type expert system based on knowledge discovery (ESKD) with traditional fault diagnosis structure of expert system, we will find the following performance characteristics and creative idea.

4.1 Abundance: traditional knowledge base system only uses reasoning machine to extend

knowledge in basic knowledge base. However, the dynamic knowledge base of ESKD experiences a series of promotion process of basic—derived—integrated—synthesized—exte nded. Both the quantity and quality of the knowledge reserve are quite abundant. Its management system is perfect and is high intellective to discover deep layer knowledge and estimate knowledge;

4.2 Strong reasoning (include Fuzzy, deduce, induce, qualitative, reasoning based on cases, statistics reasoning and so on) and interpretation ability;

4.3 Independence: the system use structure system analysis. The whole expert system is divided into independent sub-systems that can perform different performance. Each sub-system can both work cooperatively and be used independently by different users;

4.4 Practicality: during the process of diagnosis, when the test watch is placed where the blueprint has marked, this system can auto-send out the order and show kinds of target of hardware testing in this place. Meanwhile, it accepts the testing result and gives it to processor in turn, then the diagnosis result is showed. This system can also tell the operator whether the set is normal. The operator needn't do any other thing. Therefore, it is practical, convenient and popularizing;

4.5 Self-learning and Self-adaptability: self-learning is improved by coordinator, learning by cases and knowledge training. New knowledge is acquired and set into dynamic knowledge base; at the same time dynamic knowledge base and database based on knowledge discovery extend in time and space. New knowledge is regenerated to fit to the changing environment following the increase if abstract level. This makes the system rather self-adaptive.

4.6 Cocurrency: In accordance with generalized diagnosis problem, ESKD adapts to quite wide field. Meanwhile, ESKD support the cline/sever structure and different database system.

Under the operating platform of Internet and Windows95/98, we have finished the development of demonstrating program of the two important modules in ESKD with VC++5.0 and Oracle. The result is satisfactory. The following gives the example of turbine engine vibrating to show the validity of ESKD:

Suppose the following rules are known, it forms a branch of fault tree:

The oil viscosity decreases → the oil surface break → burning the bearing bush → the engine strong vibrates

In addition, there is a group of rules:

Dirt deposits in water beside oil-cooling equipment → oil cooling equipment fault → oil temperature is high

If we only know the above rules, the result possibly is that new bearing bush is used to change the old one when diagnosing the reason of strong vibrating of engines. But it isn't the essential reason. The equipment will still burn the bearing bush. The fault generally is caused by low level which cause the higher level system fault, namely the propaganda of fault is a process of propaganda from low level to the high.

The essential reason must be found in the diagnostic process so that the problem can be solved. The essential reason of the above problem is that dirt deposits in water beside oil-cooling equipment, The reason isn't found is absent relation between these two group of rules, namely the relation between the temperature of oil high and the oil adhesion degree decreasing.

If KDD* is used, we can find such a rule from database:

The temperature of oil is high → The oil adhesion degree decreases.

Then these rules can be connected into a cause link:

Dirt deposits in water beside oil-cooling equipment → oil cooling equipment fault → the temperature of oil is high → The oil viscosity decreases → the oil surface broken → burning the bearing bush → the engine strong vibrates.

At this time the essential reason, dirt deposits in water beside oil-cooling equipment can be found so the problem can be solved completely.

References

[1] Bingru Yang, KD (D&K) and Double-Base Cooperating Mechanism, Journal of System Engineering and Electronics, Vol.10, No.1, 1999.

[2] Bingru Yang, Double-Base Cooperating Mechanism in KDD, International Symposium on Computer, 149-152 (1998).

[3] Bingru Yang, FIM and CASE for Evaluation of Hazavd level Based on Fuzzy language Field, Fuzzy Sets and System, North Holland, Vol.95, No.2, 83-89 (1997).

[4] Wei-Min Shen, Discovery Regularities from Knowledge Bases, International Journal of Intelligent System, Vol.7, 623-635 (1992)

[5] U. Fayyad, G. Piatetsky—Shapiro, P. Smyth, Knowledge Discovery and Data Mining: Towards a Unifying Frame—work, Proc.of KDD-96, Menlo park, CA:AAAI Press,82-88.

[6] S. S. Anand, D.A.Bell, J.G. Hughs, EDM: A General Framework for Data Mining Based on Evidence Theory, Data &Knowledge Eng., 18,189-223 (1996).

[7] R.H. Lchiang, et. al., A Framework for the Design and Evaluation of Reverse Engineering Methods for Relational Databases, Data &Knowledge Engineering,21,57-77 (1997).

[8] H.-L.ong and H.-Y.Lee, A New Visualization Technique for Knowledge Discovery in OLAP, Proceedings of the First Pacific-Asia Conference on Knowledge Discovery and Data Mining,279-286 (1997).

[9] C. Li and G.Biswas, Unsupervised Clustering With Mixed Numeric and Nominal Data-A New Similarity Based Agglomerative System, Proceedings of the First Pacific-Asia Conference on Knowledge Discovery and Data Mining,35-48(1997)

Session RC2
Classification II
Chair: M. Hinman
Air Force Research Laboratory, USA

Model-Based Sensors Validation through Bayesian Conditioning and Dempster's Rule of Combination

Aldo Franco Dragoni, Maurizio Panti, Marco Pandolfi
University of Ancona, Istituto di Informatica
via Brecce Bianche, 60131, Ancona, Italy
tel +39-71-2204390
fax +39-71-2204474
email dragon@inform.unian.it

Abstract

Sensor data fusion/interpretation and the identification of failures are extremely important activities for the safety of complex, expensive or dangerous systems. If the system's description model is correct, conflicts among the data may only be attributed to temporary deterioration or permanent breakage of one or more sensors. Testing the sensors is sometimes impossible or too much expensive, as in unmanned space missions or nuclear power plants. This paper introduces and discusses three simple ideas:

1. classical "Model-Based Diagnosis" can be extended straightforwardly to encompass the sensors' models into the system's description in order to diagnose even their own faults
2. from the "log-file" of the diagnosed minimal conflicts among the sensors, one can draw interesting conclusion regarding their relative reliability (e.g., through Bayesian Conditioning)
3. the estimated reliability of the sensors is useful when assessing (e.g., through Dempster's Rule of Combination) the actual state of the monitored physical system, even in cases of conflicting data.

These ideas lead to the conception of a distributed monitoring system able to attach a statistically-evaluated relative degree of reliability to each sensor. This is especially useful for devices situated in dangerous zones or areas difficult or impossible to reach. This system is able to detect multiple faults of sensors and components.

1. Introduction

To control complex processes, (power plants, automated vehicles, aircraft) a large number of sensors are normally used. Sensor values directly affect the controllers' actions or the operators' decisions. Failures in generating adequate control actions as consequences of invalid sensor values often lead to total system shutdowns or hazards creating significant economic losses and sometimes even endangering the system's and human's safety. Hence, the reliability and the performance of the system are largely dependent on the validity and accuracy of the various sensors that are used. Errors can exist in the sensor readings due to the imperfect nature of the sensors, to permanent or temporary breakages and as a consequence of noises added to the readings, due to numerous known and unknown causes. Faults could be abrupt and/or incipient (slowly developing such as bias or drift in calibration). Sensor readings inconsistent with the normal model of the system, could be caused both by sensors faults or by breakages of the components of the monitored/controlled system, and it is very important to distinguish between the two. To improve the operational reliability it is necessary to validate the measured sensor data and isolate any faulty sensor; this is the task of Fault Detection and Isolation (FDI).

In Model-Based Diagnosis, collected data are confronted with a theoretical model of the monitored process/phenomenon in order to specify its current state (in case of a control system) or to validate the theory (in case of a scientific experiment). Discrepancies between theoretical models and sensor data can be imputed either to the sensors or to the theory (or to both of them). We may distinguish between three basic cases:

1. at least one sensor did not adequately report the quantity it should have measured
2. the theoretical model is not (completely) applicable to the actual monitored system because:
   a) the (scientific) theory has to be refined (*objective* interpretation)
   b) the physical system is not working as it "should" (*teleonomic* interpretation)

Case 1, often referred to as *Sensor Data Validation* (SDV), gained much interest in the last few years [1,2,3,4]. As illustrated in [5], methods can be distinguished into three categories:

SDV1. *data*-based: they rely on statistical models obtained from observed data

SDV2. *model*-based: they rely on an analytical model of the monitored system

SDV3. *knowledge*-based: they rely on human expertise

Case 2 has been deeply studied in Artificial Intelligence, both as a *knowledge revision* problem (BR for short, see [6] for an overview) and as a *model-based diagnostic* problem (MBD for short, see [7] for a survey). It seems evident to us that BR and MBD are dual problems. In the last decade, MBD moved from its theoretical foundation [8][9] to some practical applications (see for instance [10]). In MBD, diagnoses are found from discrepancies between observation and prediction. The intermediate step is the exhaustive generation of the "conflict sets" for the tuple (*SD*,*COMPS*,*OBS*), in which *S*ystem *D*escription and *OBS*ervations are sets of first-order sentences, *COMP*onent*S* is a finite set of constants each one representing a component of the system [11]. A diagnosis is a subset of *COMPS* that covers all the conflict sets.

A main problem with MBD is that each of its three fundamental steps, prediction, conflict recognition and candidate generation, exhibits a combinatorial explosion for large devices [12]. However, the worst problem with MBD is related to the case 2a before, i.e., the fact that it is at least difficult to find out a correct model for the system to diagnose. This paper does not deal with these problems: both of them will be cravenly avoided by imposing the relative simplicity of the apparatus to be controlled or diagnosed. Instead, this paper introduces, discusses and reports experimental results about the following three issues:

1. the problem of recognizing *sensors'* faults can be approached entirely within the framework of MBD (section 2)
2. from the diagnostics of the sensors' faults one can formulate interesting conclusions regarding the various sensors' relative reliability (section 3)
3. from the estimated reliability of the sensors one can hypothesize the actual state of the monitored physical system even in cases of not-redundant and conflicting data (section 4).

Normally, sensors come labeled with many important qualifications (accuracy, average life-time, ...) which are necessary to estimate their *a priori* current reliability. By "reliability" of a sensor we mean the "*probability* that the sensor is providing the correct measure," whatever the term "correct" may signify. However, the *actual* current reliability of a sensor may be smaller than the "a priori" one due to unpredictable and/or unknown events that might have affected it from its assembly to its current employment in the monitoring system. Of course, any sensor's current conditions can be appraised through appropriate testing devices. But, apart from the academic problem of infinite regression (which devices will test the testing devices, and so on, ...), a concrete question is that "testing" has its own costs. For instance, in the monitoring apparatus of an automatic production line, some optical sensors might have been altered after a temporary fault of the conditioning device that cleans the air from the pollution particles produced by the power generator. Since testing the sensors implies stopping the manufacturing process, other evidence about their possible deterioration would be appreciated. Furthermore, systematic controls and calibration of sensors "can lead to material degradation due to repetitive manipulations" [25]. Issue 2, above, suggests that such an evidence may come from the theoretical model of the monitored process/phenomenon and from the *global* datum provided by the *distributed* monitoring apparatus.

The *group* of the sensors acts as a testing device for each one of its own members. Of course, this evaluation depends on the average reliability of all the sensors in the group (hence including the corrupted ones) and on the completeness of the monitored entity's model. In any case, these estimates will not be comparable (nor for quality neither for typology) with the evaluations made by specifically designed testing devices. Their point is that they do not interfere in any way with the manufacturing process, thus they have no expenses at all (apart from the fixed costs of a CPU, some data-acquisition boards and a mass storage device). A key idea with this distributed auto-estimate is that of "minimal conflicts". Intuitively, if it has been detected a minimal conflict between the sensors A and B (by confronting the collected data with the theoretical model) and, subsequently, another minimal incompatibility is found involving B and C, then one may suppose more probable the deterioration of B than those of both, A and C. Dealing with probabilities, we do not want to reinvent the wheel since Bayesian Conditioning [13] (section 3) seems an appropriate tool to accomplish the task. Basically, the new reliability of a sensor S will be computed as the probability that S gave the correct value provided that it has been involved in some minimal conflicts. The greater the cardinality of these minimal conflicts, the higher the chance that S is working properly. The worst case is when S in involved in a singleton minimal conflict (i.e., it went, by itself, out of the range compatible with the theoretical model) so that its new reliability is 0. We will estimate statistically the current reliability of each sensor (over all its working life) w.r.t. the other ones.

There are cases in which the cost of testing a sensor is infinite, i.e., the examination is impossible or not

convenient. Let us think about the sensor equipment of unmanned satellite stations or about real-time domains in which you receive impossible (or utterly improbable) global data and have no time to test the sensors. These cases fall into the classic discipline of *decision support under uncertainty*. In these circumstances, the estimated *current* ranking of reliability plays an important role since, although very rough, it provides a more justified and up to date (hence more adequate) estimate than the "a priori" one. To accomplish this task, the fundamental tool we adopted in our method is Dempster's Rule of Combination in the special form in which Shafer and Srivastava apply it to the "auditing" domain [14] (section 4).

Section 5 describes the phenomenon of the "overexposure", which may penalize some sensors; in section 6 we introduce models whit fault behaviors. Section 7 compares our approaches with other related works and section 8 reports some tentative conclusions that might be drawn from our experiments, pointing the attention to the biggest obstacle we were faced with: the relative *overexposure* of some sensors.

2. Diagnosing sensor faults

Although these ideas come from an independent line of research [15, 16], diagnosing sensor faults can be done as well within the framework of MBD [9] by extending the system's description (e.g., figure 1-A) to encompass the sensors' models (e.g., figure 1-B).

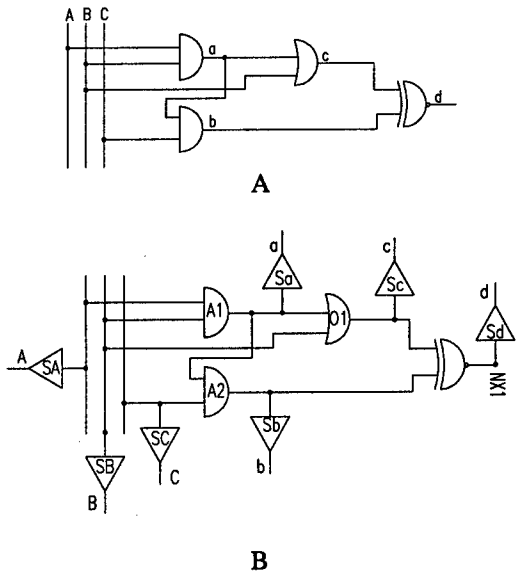


Figure 1. Extending the notion of system to encompass the sensors models

The system's description will be extended congruously (in bold below):

COMPS :

{ A1, A2, O1, NX1, S_A, S_B, S_C, S_a, S_b, S_c, S_d }

SD :

ANDG(x) ∧ ¬AB(x) ⇒ out(x) = and(in1(x), n2(x))

NXORG(x) ∧ ¬AB(x) ⇒ out(x) = or(in1(x),in2(x))

ORG(x) ∧ ¬AB(x) ⇒ out(x) = or(in1(x),in2(x))

SENS(x) ∧ ¬AB(x) ⇒ out(x) = in(x)

ANDG(A1), ANDG(A2), NXORG(NX1), ORG(O1)

SENS(S_A), SENS(S_B), SENS(S_C), SENS(S_a), SENS(S_b), SENS(S_c), SENS(S_d)

out(A1) = in1(O1), out(A1) = in1(A2),

out(A2) = in2(NX1), out(O1) = in1(NX1)

in2(A1) = in2(O1), in(S_A) = IN1(A1),

in(S_B) = IN2(A1), in(S_C) = IN2(A2),

in(S_a) = OUT(A1), in(S_b) = OUT(A2),

in(S_c) = OUT(O1), in(S_d) = OUT(NX1)

in1(A1) = 0 ∨ in1(A1) = 1, in2(A1) = 0 ∨

in2(A1) = 1, in2(A2) = 0 ∨ in2(A2) = 1

OBS : a finite set of first order ground sentences

The system components COMPS is a finite set of constants each one representing a component of the system, sensors included. The system description SD describes how the system components *normally* behave by appealing to the distinguished predicate AB whose intended meaning is "abnormal". Thus, the first sentence states that a normal (i.e. not ABnormal) and gate's output is the Boolean and function of its two inputs.

Recalling from [9], a *minimal conflict set* for (*SD*,*COMPS*,*OBS*) is a subset $\{x_1,\ldots,x_k\}$ of *COMPS* such that $SD \cup OBS \cup \{\neg AB(x_1),\ldots,\neg AB(x_k)\}$ is inconsistent and such that the same holds for no proper *sub*set of $\{x_1,\ldots,x_k\}$. Any *minimal hitting set* on the collection of all the minimal conflict sets will be a *diagnosis* for (*SD*,*COMPS*,*OBS*).

The strength of this framework is its ability to diagnose the contemporary faults of components and sensors (thus treating both the cases 1 and 2a before). However, often sensors observe physical systems in which there is no notion of component at all (e.g., distributed seismic monitoring systems [17,18]). In these cases, *COMPS* contains only the sensors, *SD* reduces to a mathematical model (maybe very complex) of the observed phenomenon and *OBS* is a simple array of numerical and/or boolean data. As an example, let us consider a metallic bar, heated at an

extremity and monitored by some thermometers as depicted in figure 2.

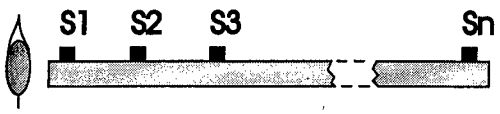


Figure 2. Diagnosing faults of pure sensor systems

Even ignoring the bar's heat transfer equation, we may yet model the system with simple constraints (in bold face below for the case of only three thermometers):

COMPS: $\{S_1, S_2, S_3\}$

SD : $\mathrm{SENSOR}(x) \wedge \neg \mathrm{AB}(x) \Rightarrow \mathrm{out}(x) = \mathrm{in}(x)$, $\mathrm{SENS}(S_1)$, $\mathrm{SENS}(S_2)$, $\mathrm{SENS}(S_3)$

$out(S_1) \geq out(S_2)$, $out(S_2) \geq out(S_3)$

OBS : a triple of numerical data

For instance, from $OBS=\{\mathrm{out}(S_1)=153°C, \mathrm{out}(S_2)=175°C, \mathrm{out}(S_3)=168°C\}$ we draw the minimal conflict sets $\{\{S_1, S_2\}, \{S_1, S_3\}\}$ and the diagnoses $\{\{S_1\}, \{S_2, S_3\}\}$.

The strongest point with the adoption of MBD in SDV relies in the notion of *good* (as we called it for the obvious duality with de Kleer's *nogood*, called "minimal conflict set" by Reiter), that is a subset $\{x_1,\ldots,x_k\}$ of *COMPS* such that $SD \cup OBS \cup \{\neg \mathrm{AB}(x_1),\ldots,\neg \mathrm{AB}(x_k)\}$ is *consistent* and such that the same holds for no proper *super*set of $\{x_1,\ldots,x_k\}$. Each good is the complement of a diagnosis w.r.t. *COMPS*, i.e., a maximally consistent set of sensors. Goods play an important role when trying to hypothesize the system's status in presence of conflicting data. In fact, because of the duality between goods and diagnoses, choosing a most probable diagnosis means choosing a most probable good, i.e., a most probable (and complete) reconstruction of the system's status.

A problem with MBD applied to SDV is that, independently of the accuracy of *SD*, the theory $SD \cup OBS \cup \{\neg \mathrm{AB}(x) \mid x \in COMPS\}$ may be *consistent* even in cases of sensor faults. These *hidden faults* may occur, for instance, in cases of contemporary breakage of more than one sensor such that the global output is still a possible (although wrong) one.

3. Estimating the sensors' actual reliability

Whereas hidden faults constitute a problem, successful recognition of *minimal* conflicts offers an invaluable opportunity to estimate, statistically, the actual current sensors' reliability from the "a priori" one. The most obvious way to do this is through Bayesian Conditioning, since we defined "sensor's reliability" as the *probability* that the sensor is returning the correct value. Let us denote with R_i and NR_i, respectively, the "a priori" and the "a posteriori" reliability of the sensor S_i, and let us denote with S the set *COMPS* restricted to the sensors. Under the assumption that the deterioration of each sensor is an independent event (!?!), the hypothesis that only those belonging to $\Phi \subseteq S$ are working properly has the combined "a priori" probability $R(\Phi) = \prod_{S_i \in \Phi} R_i \cdot \prod_{S_i \notin \Phi} (1 - R_i)$. It holds that $\sum_{\Phi \in 2^S} R(\Phi)=1$. Of course, after the recognition of a minimal conflict ϕ, $NR(\Phi)=0$ for each $\Phi \supseteq \phi$, and any other Φ is subjected to Bayesian Conditioning so that $\sum_{\Phi \in 2^S \wedge \phi \not\subset \Phi} NR(\Phi)=1$. The "a posteriori" reliability of S_i is defined as $NR_i = \sum_{S_i \in \Phi} NR(\Phi)$. If S_i is involved in *minimal* conflicts, then $NR_i \leq R_i$, otherwise $NR_i = R_i$.

Estimating the current reliability CR_i of a sensor S_i from R_i and from the history of the recognized minimal conflicts is a (debatable) statistical matter. In the experiments below, we took for CR_i the average of all the NR_i calculated during the life of the distributed monitoring system. As we'll see, such a CR_i provides an interesting relative ordering of reliability. The overall distributed sensor system acts as a testing device for each of its constituent member. Note that NR_i is calculated only on the reception of conflicting data. Another important question is that of the length of the temporal window, i.e., how far we go back in the past to record conflicting data; intuitively, the wider the window the higher the inertia of the mechanism in registering the sensors' deterioration.

4. Choosing the preferred good

In Shafer's and Srivastava's multi-source version of the belief function framework [14], the sources' degrees of reliability are "translated" into belief-function values on the given pieces of information. In our method we follow them by taking the estimated reliability CR_i as primary *evidence* in favor of the datum s_i furnished by S_i. Let Ω denote the set of all the possible configurations of the monitored system,

and let $[s_i] \subseteq \Omega$ denote only those compatible with s_i. The key assumption is that *a reliable sensor cannot give false information, while an unreliable sensor can give correct data*; the hypothesis that "S_i is reliable" is compatible only with $[s_i]$, while the hypothesis that "S_i is *un*reliable" is compatible with the entire Ω. Each S_i gives an evidence for Ω and generates the following *basic probability assignment* (*bpa*) m_i over the elements X of 2^{Ω}:

$$m_i(X) = \begin{cases} CR_i & \text{if } X = [s_i] \\ 1 - CR_i & \text{if } X = \Omega \\ 0 & \text{otherwise} \end{cases}$$

All these *bpa*s will be then combined through Dempster's Rule of Combination (DRC):

$$m(X) = m_1(X) \otimes \ldots \otimes m_n(X) = \frac{\sum_{X_1 \cap \ldots \cap Xn = B} m_1(X_1) \cdot \ldots \cdot m_n(X_n)}{\sum_{X_1 \cap \ldots \cap Xn \neq \emptyset} m_1(X_1) \cdot \ldots \cdot m_n(X_n)}$$

From the combined *bpa* m, the credibility of a set of data (hence of a good) s is given by

$$Bel(s) = \sum_{X \subseteq [s]} m(X)$$

A major problem with the belief function formalism is the computational complexity of DRC; the straightforward application of the rule is exponential in the cardinality of Ω and in the number of sensors. However, much effort has been spent in reducing its complexity. Such methods range from "efficient implementations" [19] to "qualitative approaches" [20] through "approximate techniques" [21].

5. The sensors' overexposure

Normally, conflict sets contain sensors that are correctly providing the value they should. By "exposure" of a sensor we mean "its probability to be unjustly involved in a conflict". Unfortunately, some sensors are *more* exposed than the others, and such exposure depends not only on the model but also on the real (unknown) capacity of the sensors. Let us clarify the point with a simple example. Consider the monitoring system depicted in figure 2 with only three thermometers, S1, S2 and S3. Each thermometer has its own degree of capacity (which is unknow); suppose that S2 has a degree of capacity 99, while S1 and S3 have a degree of 90. The sensor S2 is more exposed than the others for two reasons:

1 being it more accurate, probably it will be innocent in case of conflict with other sensors

2 it appears in two relations in the System Description, while the others appear in only one relation each. S2 is involved both in the errors of S1 and of S3, while S1 and S3 are only involved in errors of S2. We say that S2 is "overexposed". In order to reduce the component of overexposure which is due to the model, we can try to explicitate all the relations which are only implicit in the model. For instance, in the example above, we can espicitate the relation out(S1) > out(S3). Doing so, we do not render more complete the System Desription, however, in general, it is possible to add a new relation derived from other experiences and/or knowledge that makes the model more complete.

Unfortunately, the overexposure due to differences in the real capacity of the sensors remains unknown. However, if it is possible to observe the system in a preliminary test, and calculate the exposure of each sensor, we can try to correct such overexposure. Having calculated the deviation of the exposure for each sensor, then those exceeding a threshold (defined from σ), are the overexposed sensors.

6. Faulty modes

Till now we have considered models in which every sensor might be just normal or faulty. In general, models could contemplate modes of independent faulty behaviors. Let us think of an example with two faulty modes for some sensors. Then every sensor can be in one of the 3 possible states: Normal (N), Fault1(F1) and Fault2 (F2). The "a priori" probability for each state of each sensor must be given (rispectly R_i, F_{i1} and F_{i2}) and they must sum up to 1. In the Bayesian Conditioning, one must distribute the probability to the set of possible states (3^n, n = number of the sensors), in accordance with the formula:

$$R(\Phi) = \prod_{S_i \in \Phi} R_i \cdot \prod_{S_i \notin \Phi} F_{ij} \quad (j = 2,3)$$

Hence the "a posteriori" reliability is calculated and used to estimate the actual current sensors' reliability and to choose the preferred good.

7. Related work

The system presented in this paper belongs to both the SVD1 and SVD2 category of sensor data validation. The model-based methods presuppose the existence of an analytic model of the system and are based on a common methodology: generating and analyzing of the signals sensitive to the fault *residues* [26]. The main approaches to generate the residues can be classified in approaches based on *analytic redundancy* [25] (e.g. *parity space*) and approaches based on the *observator* (e.g. *Kalman filter*). In Dorr et al.[23-25] faults are detected by comparing each residual with thresholds defined with respect to the sensor measurement accurancy. Residuals are generated by comparing each measurement with an estimate value. They can be obtained by simple redundancy (at least two sensors that measure the same physical variable) or analytical redundancy. The model-based methods are very efficient when there is a linear and well-known model of the system. The problem has been also deeply studied in Artificial Intelligence. In [2] Lee presents a technique based on the analytic redundancy that needs an accurate knowledge of the process. Also in [22], Washio proposes a method to find the sensors' faults based on the model of the monitored process. Detection of faults is based on consistency checking between observations and optimal constraints, called "minimal over-constraints", consisting of first principles. If some inconsistencies are detected, the model-based diagnosis is applied to derive the candidate of faulty components. The individuation of the sensors' faults and the diagnosis of the components are performed contextually in the same framework. In particular, the method allows diagnosing no linear components, sensors and components, and the width of the fault (table 1).

Even if the methods overcome the non-linearity problem, all the model-based methods are sensitive to errors of the modeling. So, when it is not possible have an accurate model of the system, an alternative way is using a qualitative description based on the human experience: knowledge-based methods. Typical examples of this category are the classical expert systems (formed by a knowledge base and an inferential engine) and fuzzy expert systems.

In [27] is given a survey on the state of the art of model-based diagnosis employing artificial intelligence approaches. Emphasis is placed on neural network techniques and the use of fuzzy models for residual generation and fuzzy logic for residual evaluation. The different strategies for diagnosing the faults in continuous systems with qualitative models can roughly be divided into two groups:

1 fault model based

2 normal model based

It seems obvious that a fault diagnosis concept using a qualitative knowledge-based models can be organized in a configuration similar to that of the analytical model-based.

Table 1. Compared work

Character	Our	Dorr et al.	Washio
Method of diagnosis	Through minimal conflicts	Through residuals analysis	Through minimal conflicts
Mult.faul. of comp.	●		●
Mult.faul. of sensors	●	●	●
Non-lin. Models	●	●	●
Highly non-lin.	●		●
Dynamic behaviors	●		●
Human expertise	●		
Faults amplitud.		●	●
Estimated reliability	●		
Complexity	Exp.in the card. of Ω	Iterative method	Exp.in the numb.of undet.quant in SD

In this paper we presented a method based on the knowledge of the system that is not necessarily expressed by a mathematical model. The method needs any kind of knowledge to extract the *minimal conflicts*. This allows using both equation and constraints of real situations like "if the temperature is bigger than 100° C, then the alarm has to start"; the model-based methods cannot manage these constraints.

In the SVD, it is supposed that the components are not corrupted. In our approach this constraint can be overcome extending properly the method.

The historical analysis of the data allows exploiting information formerly draw out for solve the future conflicting situations. The systems proposed in [22] and [16] don't give indications concerning how to solve the conflicts and how to choose one of the possible diagnoses.

8. Conclusions

Elaboration of data coming from multiple sensors is critical when conflicts emerge among them. This paper introduces and discusses three issues:

1. the problem of recognizing *sensors'* faults can be approached entirely within the framework of model-based diagnosis;
2. from the history of the sensors' faults it is possible to formulate interesting conclusions regarding the various sensors' relative reliability, by means of Bayesian Conditioning;
3. from the estimated reliability of the sensors it is possible to hypothesize the actual state of the monitored physical system even in cases of not-redundant and conflicting data, by means of Dempster's Rule of Combination.

We have proposed a monitoring system which is used to detect faults and to diagnose their location (Fault Detection and Isolation). Mathematical models can be not enough to perform efficient FDI. Our method can cope with both, analytical and heuristic expert knowledge, from complex dynamic equations to simple numerical constraints, rules and facts; in this manner we increase the robustness of the diagnostic process. The hardest problem with this method is what we called "overexposure effect", which depends on the model and on the real capacity of the sensors.

9. References

[1] Alag S. and A. M. Agogino, "A Methodology for Intelligent Sensor Validation, Fusion and Sensor Fault Detection for Dynamic Systems" #95-0301-P, 1995.

[2] S. C. Lee, "Sensor Value Validation Based on Systematic Exploration of the Sensor Redundancy for Fault Diagnosis KBS", IEEE Transactions on Systems, Man, and Cybernetics, Vol 24, n° 4, April 1994.

[3] S. Iksch, W. Horn, C. Popow, F. Paky, "Context-sensitive data validation and data abstraction for knowledge-based monitoring", in ECAI 94. 11th European Conference on Artificial Intelligence, 1994, pp. 48-52.

[4] D. Himmelblau and M. Bhalodia, "On-Line Sensor Validation of Single Sensor Using Artificial Neural Networks", in Proceedings of the 1995 American Control Conference, vol. 1, 1995, pp 766-770.

[5] Stephane Canu,"Sensor Data Validation: a Review", in EM2S Esprit European Project.

[6] Dragoni A.F., Belief Revision: from theory to practice, "The Knowledge Engineering Review", vol 12, n° 2, Cambridge University Press, 1997.

[7] de Kleer, J., Hamsher, W. C., Console, L., (Eds.): Readings in Model-Based Diagnosis, Morgan Kaufmann, 1992.

[8] de Kleer, B., C., Williams, Diagnosing Multiple Faults, Artificial Intelligence 32, pp. 97-130, 1987.

[9] R. Reiter, A Theory of Diagnosis from First Principles, Artificial Intelligence 32, pp. 57-95, 1987.

[10] Beschta, A., Dressler, O., Freitag, H., Montag, M., Stru§, P., A model-based approach to fault localisation in power transmission networks, Intelligent Systems Engineering, 2(1), IEE Publisher, pp. 3-14, 1993.

[11] de Kleer, J., Mackworth, A. K., Reiter, R., Characterizing Diagnoses, Proceedings of 1990 Conference of the American Association for Artificial Intelligence, pp 324-330, 1990.

[12] Johan de Kleer, Focusing on Probable Diagnoses, Proceedings of 1991 Conference of the American Association for Artificial Intelligence, pp 842-848, 1991.

[13] T. Bayes, "An essay towards solving a problem in the doctrine of chances." Philosophical Transactions Royal Soc'y London 53:370-418, 1763. Reprinted in various publications, e.g. Biometrika 45:293-315.

[14] Shafer G. and Srivastava R., The Bayesian and Belief-Function Formalisms a General Perpsective for Auditing, in G. Shafer and J. Pearl (eds.), Readings in Uncertain Reasoning, Morgan Kaufmann Publishers, 1990.

[15] A.F. Dragoni, P. Giorgini, P. Puliti, "Finding Inconsistencies as a Basis for Evaluating the Reliability of Different Information Sources", in Binaghi, E., Brivio, P.A. and Rampini, A. (Eds.), Soft Computing in Remote Sensing Data Analysis, Series in Remote Sensing, vol 1, World Scientific Publishing, 1996.

[16] A.F. Dragoni, P. Giorgini, A. Corsi, "Sistemi di Monitoraggio Distribuiti Automonitoranti", ANIPLA 41st national Conference, Turin 5-7 november 1997.

[17] Mason, C., An Intelligent Assistant for Nuclear Test Ban Treaty Verification, IEEE Expert, vol 10, no 6, 1995.

[18] Cindy L. Mason and Rowland R. Johnson, datms: A Framework for Distributed Assumption Based Reasoning, in L. Gasser and M. N. Huhns eds., Distributed Artificial Intelligence 2, Pitman/Morgan Kaufmann, London, pp 293-318, 1989.

[19] Kennes, R., Computational Aspects of the Möbius Transform of a Graph, IEEE Transactions in Systems, Man and Cybernetics, 22, pp 201-223, 1992.

[20] Parson, S., Some qualitative approaches to applying the Demster-Shafer theory, Information and Decision Technologies, 19, pp 321-337, 1994.

[21] Moral, S. and Wilson, N., Importance Sampling Monte-Carlo Algorithms for the Calculation of Dempster-Shafer Belief, Proceeding of IPMU'96, Granada, 1996.

[22] Washio, Sakuma, Kitamura., A new approch to quantitative and credible diagnosis for multiple faults of components and sensors, Artificial Intelligence Vol 91 pp.103-130 .

[23] Dorr R., Mourot F., Kratz F., Loisy F., An application of gross errors measurements detection in nuclear power plants, IFAC Safeprocess '94, June 1994, Espoo, Finland.

[24] Dorr R., Kratz F., Ragot J., Loisy F., Germain J.L., Sensor validation in a nuclear power plant using direct and analytical redundancy, International Conference on Industrial Automation, June 1995, Nancy France.

[25] Dorr R., Kratz F., Ragot J., Loisy F., Loisy F., and Germain J., Detection, Isolation, and Identification of Sensor Faults in Nuclear Power Plants, IEEE Transaction on Control Systems Technology, Vol. 5, no. 1, January 1997.

[26] Patton R.J., Franck P.M., and Clark R.N., Fault Diagnosis in Dynamic Systems: Theory and Application. Englewood Cliffs, NJ: Prentice-Hall, Series in systems and Control Engineering, 1989.

[27] P.M. Frank, B. Köppen-Seliger, Recent Trends in Fault Diagnosis. Automazione e Strumentazione, April 1997, pp. 93-103.

Merge and Split Hypothesis for Data Fusion in the Evidential Reasoning Approach

E-h ZAHZAH, L MASCARILLA
Laboratoire d'Informatique et d'Imagerie Industrielle
Universit de La Rochelle,Avenue Marillac
17042 La Rochelle Cedex 1 (FRANCE)
e-mail: {ezahzah, lmascari }@univ-lr.fr

Abstract *Actual systems uses data providing from several sources, this data may be for different reasons heterogeneous, uncertain, imprecise, and unreliable. A system decision must manage with all of this qualitative and quantitative information. In some cases, at the finale decision step, a system must choose between a coherent result even if it imprecise or a precise result even if it is incoherent. It is obvious that a ideal system must be both precise an coherent. but when data we manage with does not allow it, one must make a choice. In this context, we propose in this paper a method to combine data providing from several sources, using the Dempster-Shafer rule. Before taking any decision, the system gives and additional information concerning the sources involved (reliable or not), and it gives also an information related the conflict produced by combining such sources.*

Keywords: Dempster-Shafer Theory , multiple hypothesis, uncertainty and ignorance management.

1 Introduction

The Dempster-Shafer Theory (DST) offers an interesting tool to combine data providing from heterogeneous sources more or less reliable by managing imprecision and uncertainty.This is particularly important when dealing with multi-modality imaging (satellite image), where the fusion of information increases the global knowledge about the phenomenon. If θ represents the universe of discourse, also called frame of discernment the DST enables to assign mass to 2^θ rather to solely θ hypothesis as in the Bayesian approach. The DST has been used in many applications as in Pattern recognition and Image Analysis, but without its all powerful. When using with singletons i.e an object belongs to a unique class even in uncertain situations, the DST falls in the Bayesian approach, which is considered as a particular case. In this context, some authors attempt to use double hypothesis but their method remains *ad hoc*. In order to better use the DST with multiple hypothesis, we propose in this paper a method to estimate the mass of ignorance from any belief measurements, and merge the close hypothesis to finally use the orthogonal Dempster-Shafer rule to split as most as possible, the most credible hypothesis (unique or multiple), with a low degree of conflict, which enables to take a final and less risky decision.

2 Fundamentals of DST

2.1 Some functions in DST

The DST uses a frame of discernment which is a set interpreted as a set of mutually exclusive propositions. The propositions of interest are assumed to be expressed as subsets of the frame θ which is assumed to be a finite set. A mass function over θ also known as a basic probability is a function $m : 2^\theta \rightarrow [0, 1]$ such

that $m(\emptyset) = 0$, and $\sum_{A\in\theta} m(A) = 1$ The set of focal elements $\Im_m$ of a mass function m is defined to be a set of subsets of theta for which m is non zero, i.e, $\{A \subseteq \theta : m(A) \neq 0\}$, the core C_m of a mass function m is defined to be a union of its focal elements, that is $\cup_{A\in\Im_m} m$ can be viewed as being a mass function over C_m, which sometimes advantageous computationally.

1. *The belief function* $Bel : 2^\theta \to [0,1]$ is given by $Bel(A) = \sum_{B\subseteq A} m(B)$, if there exists a mass function m over θ with, for all $A \in 2^\theta$, $Bel(A) = \sum_{B\subseteq A} m(B)$ Bel is said to be the belief function associated with m. To every mass function over θ there corresponds a unique belief function; conversely for every belief function over θ there corresponds a unique mass function. To recover mass function m from its associated belief function Bel on can uses the following equation

$$For\ A \subseteq \Theta, m(A) = \sum_{B\subseteq A} (-1)^{|A-B|} Bel(B)$$

2. *The Plausibility function* $Pl : 2^\theta \to [0,1]$ associated with mass function m is defined by the equations: for all $A \in 2^\theta$, $Pl(A) = \sum_{B\cap A\neq\emptyset} m(B)$ There is a simple relationship between the belief function Bel and the plausibility function Pl associated with a particular mass function m: for $A \subseteq \theta, Pl(A) = 1 - Bel(\bar{A})$, and $Bel(A) = 1 - Pl(\bar{A})$. The problem of computing values of plausibility is equivalent to the problem computation values of belief. A mass function is viewed as a piece of ambiguous evidence that may mean A, for any $A \in \Im_m$; we consider that with probability $m(A)$, it means A. $Bel(A)$ can be thought of as the probability that the ambiguous evidence implies A, and $Pl(A)$ as the probability that the evidence is consistent with A.

3. *The commonality function* $Q : 2^\theta \to [0,1]$ associated with mass function m is defined by the equations : for all $A \in 2^\theta$, $Q(A) = \sum_{B\supseteq A} m(B)$. It doesn't usually have a simple interpretation but it allows a simple statement od Dempster's rule(ref). A commonality function determines and is determined by a mass function: if Q_i is the commonality function associated with mass function m_i for $i = 1, 2$ the $Q_1 = Q_2$ if and only if $m_1 = m_2$.

2.2 Dempster's Rule of Combination

suppose we have a number of mass functions, each representing a separate piece of information. The combined effect of these, given the appropriate independence assumptions, is calculated using Dempster's rule of combination. Let m_1 and m_2 be mass function over theta. Their combination using Dempster'rule $m_1 \oplus m_2$, is defined by, for $\emptyset \neq A \subseteq \Theta$,

$$m_1 \oplus m_2(A) = k \sum_{B\cap C=A} m_1(B)m_2(C)$$

where K is a normalization constant chosen so that $m_1 \oplus m_2(A)$ is a mass function, and so is given by

$$k^{-1} = \sum_{B\cap C\neq\emptyset} m_1(B)m_2(C)$$

This combination is only defined when$k$ is non zero; this happens only when if the cores of m_1 and m_2 is non-empty.

The operation $\oplus$ is associative and commutative. The combination $\oplus_{i=1\ldots k} m_i$ of mass functions $m_1, m_2, ..., m_k$ over theta is well-defined if the intersection of all the cores is non-empty, that is $\cap_{i=1\ldots k} C_m \neq \emptyset$. In this case their combination $\oplus_{i=1\ldots k} m_i$ can be shown to be given by, for $\emptyset = A \subseteq \Theta$

$$\oplus_{i=1}^{k} m(A) = K_{1,\ldots k} \sum_{B_1\cap\ldots\cap B_k=A} m_1(B_1)\ldots m_k(B_k)$$

$$K_{1,\ldots k}{}^{-1} = \sum_{B_1\cap\ldots\cap B_k\neq\emptyset} m_1(B_1)\ldots m_k(B_k)$$

The normalization constant $K_{1,\ldots k}$ can be viewed as a measure of the conflict between

the evidences. Let $\oplus_{i=1}^{k} Bel_{m_i}$ and $\oplus_{i=1}^{k} Q_{m_i}$ be the belief and the commnonality function respectively corresponding to the combined mass function $\oplus_{i=1}^{k} m_i$, they satisfy, for $\emptyset \neq A \subseteq \Theta$

$$\oplus_{i=1}^{k} Bel_{m_i}(A) = K_{1,..,k} \sum_{\emptyset \neq B_1 \cap .. \cap B_k \subseteq A} m_1(B_1)..m_k(B_k)$$

and

$$\oplus_{i=1}^{k} Q_{m_i}(A) = K_{1,..,k} Q_{m_1}(A)..Q_{m_k}(A)$$

The last result can be rewritten as

$$\oplus_{i=1}^{k} Q_{m_i}(A) = K_{1,..,k} \prod_{i=1}^{k} m_k(A)$$

showing for commonalities, combination is just normalized product.

3 Mass Ignorance Estimation

In practical cases, and mainly in classification or pattern recognition problems, we manage with a finite set of classes, let's say $\{\Omega = \omega_1, \omega_2, .., \omega_n\}$. The mass corresponding to each class is given by a classifier as probabilistic, possibilitic or any other measures. In a bayesian process for instance, all the unity mass is assigned to classes, and the information concerning this assignment (ignorance) is not represented. To overcome this situation, we propose to add to the initial set Ω, an additional class θ which represents the ignorance which can also indicates the reliability of the source. The mass assignment to this special class is calculated based on the following idea:

If the entire mass value is assigned to ignorance, then the source is considered non informative and totally unreliable, and conversely, if its mass value is zero, then the source is considered as completely reliable.

In our case, the mass of ignorance $m(\theta)$ is estimated from the initial mass measures $m(\omega_i)$. We assume that $m(\theta)$ is maximal if the confusion degree among the given hypothesis is high, this situation occurs when all the mass degrees are identical and in the other hand the mass of ignorance is minimal if the separability between one hypothesis and the rest is maximal.(no confusion occurs). This can be expressed by:

$$m(\theta) = 1 - sep \; where$$

$$sep = m(\omega_{max}) - \frac{\sum_{i=1, i \neq max}^{n} m(\omega_i)}{(n-1)}$$

ω_{max} corresponds to the class with the maximum mass value.
After adding this special class to Ω, the mass of the new set is normalized according to DST axioms

$$\Omega_{new} = \{\Omega_{old}, \theta\}$$

$$\sum_{x \subseteq \Omega_{new}} m(x) = 1$$

4 Merge and Split hypothesis Process

4.1 Merging Hypothesis

At this point of process, the mass are assigned to a single hypothesis (class), with an additional information concerning the reliability of theses assignments. To uses the DST with its powerful in the next stage, we are not considering the 2^{Θ} hypothesis which are very complex to estimate and to represent. The hypothesis which are close are merged in what we call a multiple-hypothesis, and hope that information providing by others sources will separate these hypothesis if these latest are concordant. The merge process is performed by assigning a part of the the mass let's say $1 \leq q < 0$ to the merged hypothesis and the rest of the mass to ignorance.

1. $\Omega_{old} = \{\omega_i\} \; i = 1, ..., \theta$

2. $\Omega_{new} = \{\Omega_{old} \bigcup \omega_i\} \; if$
   $|m(\Omega_{old}) - m(\omega_i)| \leq \epsilon \; i = 1, ..., \theta$

3. Ω_{new} represents the new set of merged class, or multiple hypothesis. Ω_{old} is the previous set of classes with single or multiple hypothesis.

The update of the mass assignment of the merged class and ignorance are realized according to the following formulas:

$$m(\Omega_{ij}) = m(\Omega_i) + q.m(\omega_j)$$

$$m(\theta) = m(\theta) + (1 - q).m(\omega_j)$$

At the end of this process we get $|\Omega_{new}| \leq |\Omega|$ i.e the cardinalty of the new set of merged hypothesis is less or equal to the cardinality of the first set Ω (universe of discourse).

4.2 Split Process with the DS combination rule

The $Dempster - Shafer$ rule is a scheme of inference to aggregate bodies of evidence provided by multiple information sources. If m_1 and m_2 are the mass functions over θ the frame of discernment or the universe of discourse,their combination using Dempster'rule $m_1 \oplus m_2$, is defined by, for $\emptyset \neq A \subseteq \Theta$,

$$m_1 \oplus m_2(A) = k \sum_{B \cap C = A} m_1(B) m_2(C)$$

where K is a normalization constant chosen so that $m_1 \oplus m_2(A)$ is a mass function, and so is given by

$$k^{-1} = \sum_{B \cap C \neq \emptyset} m_1(B) m_2(C)$$

k also represents the degree of conflict between the combined sources.
A,B,C are assumed here to be both single or mutiple-hypothesis. The $DS - rule$ is used to combine all information providing by different sources according to all the hypothesis (single or multiple), with an additional information concerning the reliability of the source. (figure 1) After combining this data, the decision is taken over a single or multiple hypothesis. If enough agreeing information are providing by the sources, the decision is taken over a single hypothesis otherwise, the hypothesis remain multiple. However, the conflict degree using the $DS - rule$ with multiple hypothesis will be less or equal than the one obtained by the $DS - rule$ using single hypothesis.

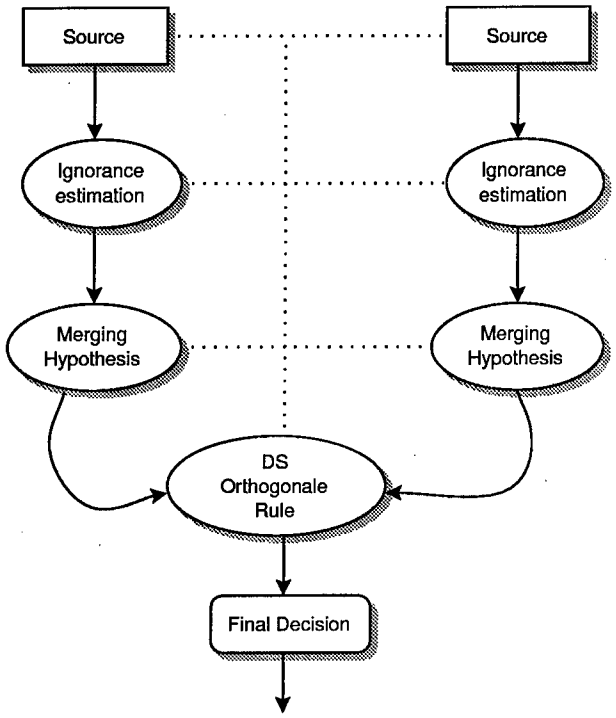


Figure 1: Merge and Split System

5 Decision Rules

The decision is taken from the hypothesis providing the maximum value of credibility, plausibility or pignijstic probability this latest is estimated only on the single hypothesis see [4].

- Maximum of Credibility:

$$Max_i(Cr(\Omega_i))$$

$$Cr(\Omega_i) = \sum_{X \subset \Omega_i} m(X)$$

The credibility is a the minimal belief degree assigned to a given hypothesis. For

a given hypothesis, the credibility is the sum of the mass of all the hypothesis supporting this hypothesis. The credibility is considered as a pessimistic decision. In our case the result maybe a singleton or a subset (multiple hypothesis) which verify the maximum of credibility, taking into account the different opinions of the sources.

- Maximum of Plausibility:

$$Max_i(Pl(\Omega_i))$$

$$Pl(\Omega_i) = \sum_{X \cap \Omega_i \neq \emptyset} m(X)$$

The Plausibility is a the maximal belief degree assigned to a given hypothesis. For a given hypothesis, the Plausibility is the sum of the mass of all the hypothesis not in contradiction with the hypothesis . The Plausibility is considered as a optimistic decision.

- Maximum of Pignijstic Probability The credibility and plausibility are approximate by a probability measure by assigning the mass placed in each multiple hypothesis between the single hypothesis

$$Max_i(Pign(\Omega_i))$$

$$Pign(\Omega_i) = \sum_{A \in 2^D, \Omega_i \subset A} \frac{1}{|A|} m(A)$$

This approximation is similar to one applied in the bayesian inference when the lack of information. A uniform distribution is used on the hypothesis, and the probability issued is called "pignistic probability". The decision rule adopted here lead to choose a single hypothesis with the highest value of $Pign(\Omega_i)$

6 Example

Let S_1 and S_2 be two different sources providing their opinion on six singles hypothesis (classes), Let for instance m_1 and m_2 be the belief function expressing these measures:

$$\theta = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$$

$$S_1 : m_1(0.063, 0.712, 0.004, 0.927, 0.468, 0.904)$$

$$S_2 : m_2(0.852, 0.651, 0.947, 0.592, 0.354, 0.237)$$

From these belief functions, the mass of the set of single hypothesis ω_i and the ignorance(θ_1, θ_2) for S_1 and S_2 are estimated and the mass function will be mapped to

$$m_1(\omega_i, \theta_1) = (0.02, 0.2, 0.00, 0.26, 0.13, 0.25, 0.14)$$

$$m_2(\omega_i, \theta_2) = (0.21, 0.16, 0.23, 0.15, 0.09, 0.06, 0.1)$$

The mass assigned to ignorance for the sources S_1 and S_2 are respectively 0.14 and 0.1

6.1 Fusion with simple hypothesis

The DS orthogonal sum with simple hypothesis(singleton) applied to the above example gives:

$$m_{1 \oplus 2}(\omega_i, \theta) = (0.11, 0.23, 0.1, 0.26, 0.1, 0.15, 0.04)$$

$$Cr(\omega_i, \theta) = Pl(\omega_i, \theta) = Pign(\omega_i, \theta) = m_{1 \oplus 2}(\omega_i, \theta)$$

The conflict degree among the two source $k = 0.67$, is relatively great. The class ω_4 gives the maximum value of credibility, plausibility and pignijstic probability , so if a decision must be taken, it will surely be the class ω_4, knowing that the the conflict degree is high.

6.2 Fusion with merging hypothesis

In a second case, if we decide to use the DS rule after merging the hypothesis, we will have the following results with $\epsilon = 0.025$ and $q = 0.5$:

1. Merging the hypothesis of the first source S_1 gives :

$$m_1(\Omega_{13}, \Omega_2, \Omega_{46}, \Omega_5, \theta_1) =$$
$$(0.01, 0.20, 0.26, 0.13, 0.40)$$

2. Merging the hypothesis of the second source S_2 gives:

$$m_2(\Omega_{13}, \Omega_{24}, \Omega_5, \Omega_6, \theta_2) =$$
$$(0.22, 0.15, 0.09, 0.06, 0.48)$$

3. The DS orthogonal rule from these values gives:

   - $m_{1\oplus 2}(\Omega_{13}, \Omega_2, \Omega_4, \Omega_6, \Omega_{46}, \Omega_5, \Omega_{24}, \theta)$= (0.12,0.16,0.05,0.05,0.16,0.14,0.08,0.25);
   -$Cr(\Omega_{13}, \Omega_2, \Omega_4, \Omega_6, \Omega_{46}, \Omega_5, \Omega_{24}, \theta) =$ $(0.12, 0.16, 0.05, 0.05, 0.25, 0.14, 0.29, 0.25)$
   -$Pl(\Omega_{13}, \Omega_2, \Omega_4, \Omega_6, \Omega_{46}, \Omega_5, \Omega_{24}, \theta) =$ $(0.12, 0.24, 0.28, 0.20, 0.33, 0.14, 0.44, 0.25)$
   -$Pign(\Omega_{13}, \Omega_2, \Omega_4, \Omega_6, \Omega_{46}, \Omega_5, \Omega_{24}, \theta) =$ $(0.00, 0.20, 0.17, 0.13, 0.00, 0.14, 0.00, 0.25)$
   and the conflict degree $k = 0.21$

The maximum value of credibility, plausibility is assigned to the the double hypothesis Ω_{24}. The interpretation of this result, is that with the available information of the actual sources, nothing allows us to decide between class ω_2 or ω_4. The system requires more information to discriminate between the two classes. If a decision must be taken on a singleton hypothesis, the pignijstic probability of class ω_2 gives the highest value. Note also that the conflict degree $k = 0.21$ is relatively low compared to conflict degree k of the previous case. One of the advantage of merging and splitting hypothesis is that the conflict degree is low and enables to take a less risky final decision.

7 Conclusion

In this paper a method of updating the mass of hypothesis and ignorance is proposed, this method is integrated in a process of mass functions initialization in the case of multiple (2^θ) hypothesis in order to better use the *Orthogonal Dempster – Shafer Rule* to combine data providing from several heterogeneous sources in a decision system. The method allows to combine data providing from several sources, and before taking any decision, the system gives additional information concerning the sources involved (reliable or not), and the conflict produced by combining such sources. we are actually testing the method on multispectral image satellite with various algorithms of classification considered as data sources more or less reliable and precise. The first results seem to be very promising.

References

[1] Barnett J.a : *"Computational methods for mathematical theory of evidence"*: Proc of the 7th IJCAI, 868,875, Vancouver, 1981.

[2] Bloch I: " *Fusions de donnes, ensemble flous et morphologie mathmatique en traitement d'images, Application l'imagerie mdicale crbrale et cardio vasculaire multi-modalits"*: ENST- Paris Avril 1995.

[3] Mascarilla L, Frélicot C , Zahzah E-h: *"Combining : An Alternative For Choosing Between Reject-First And Accept-First Classifiers"*, NAFIPS'99, 10-12 june '99, NYC, USA.

[4] P. Smets. and R. Kennes. *"The transferable belief model"*. *Artificial Intelligence* 66:191–243, 1994

[5] Zahzah E-h . *"Contribution la reprsentation des connaissances et leur utilisation pour l'interprtation automatique des images satellite"*: Thse de doctorat de l'universit Paul Sabatier Toulouse III. Septembre 1992

[6] Rocha L M : *"Evidence States: Modeling Subjective categories"*: Int Journal of General system. Vol. 28.

[7] Wilson N : *"Algorithms for Dempster-Shafer Theory"*: Report, Dept of Computing and mathematical Sciences, Oxford Brookes University. Sept 98.

A classification method based on the Dempster-Shafer's theory and information criteria

E. LEFEVRE
PSI, UNIVERSITE/INSA de ROUEN
Place Emile Blondel, BP 08
76131 Mont-Saint-Aignan Cedex, France

O. COLOT
PSI, UNIVERSITE/INSA de ROUEN
Place Emile Blondel, BP 08
76131 Mont-Saint-Aignan Cedex, France

P. VANNOORENBERGHE
PSI, UNIVERSITE/INSA de ROUEN
Place Emile Blondel, BP 08
76131 Mont-Saint-Aignan Cedex, France

Abstract ***Within the framework of pattern recognition, many methods of classification were developed. More recently, techniques using the Dempster-Shafer's theory or evidence theory tried to deal with the problem related to the management of the uncertainty and data fusion. In this paper, we propose a classification method based on the Dempster-Shafer's theory and information criteria. After an original basic belief assignment, we introduce an attenuation factor based on the dissimilarity between probability distributions.***

Keywords: Data Fusion, Dempster-Shafer's theory, Information Criteria, Classification.

1 Introduction

Data analysis and processing are two important tasks in today's information society. The data management becomes essential when the information is imperfect, that is to say imprecise and uncertain. Traditionally, probability theory, which is inadequate in some cases as well known [1], is used for dealing with imperfect data. In the recent past, other models have been developed for handling imprecise knowledge (theory of fuzzy sets [2], possibility theory [3, 4]) or uncertain information (theory of belief functions [5]). In this paper, we deal with a classification method of imperfect data sets using evidence theory [5, 6, 7]. Recently, in this context, a new approach using neighbourhood information has been developed [8]. Each nearest neighbour of a pattern to be classified is considered as an item of evidence. The resulting belief assignment is also defined as a function of the distance between the pattern and its neighbour. We propose an alternative solution to this classification method in initializing the belief functions using information criteria. This paper is organized as follows. In section 2, we introduce notations allowing to describe the Dempster-Shafer's Theory of evidence. Section 3 is devoted to present the proposed methodology. This work is applied to synthetic and real data (section 4).

2 Dempster-Shafer's Theory

In this section, a brief overview of the Evidence's Theory [5] is provided. Let Θ represents the set of hypotheses H_n, called the frame of discernment. The knowledge about the problem induces a basic belief assignment which allows to define a belief function m from

2^Θ to $[0,1]$ such as :

$$m(\emptyset) = 0 \quad (1)$$

$$\sum_{H_n \subseteq \Theta} m(H_n) = 1. \quad (2)$$

Subsets H_n of Θ such that $m(H_n) > 0$ are called focal elements of m. From this basic belief assignment m, the credibility $Bel(H_n)$ and plausibility $Pl(H_n)$ can be computed using the equations :

$$Bel(H_n) = \sum_{A \subseteq H_n} m(A) \quad (3)$$

$$Pl(H_n) = \sum_{H_n \cap A \neq \emptyset} m(A). \quad (4)$$

The value $Bel(A)$ quantifies the strength of the belief that event A occurs. These functions (m, Bel and Pl) are derived from the concept of lower and upper bounds for a set of compatible probability distributions. In addition, Dempster-Shafer's theory allows the fusion of several sources using the Dempster's combination operator. It is defined like the orthogonal sum (commutative and associative) following the equation :

$$m(H_n) = m_1(H_n) \oplus \cdots \oplus m_M(H_n). \quad (5)$$

For two sources S_i and S_j, the aggregation of evidence for a hypothesis $H_n \subseteq \Theta$ can be written :

$$m(H_n) = \frac{1}{\mathcal{K}} \sum_{A \cap B = H_n} m_i(A).m_j(B) \quad (6)$$

where $\mathcal{K}$ is defined by :

$$\mathcal{K} = 1 - \sum_{A \cap B = \emptyset} m_i(A).m_j(B). \quad (7)$$

The normalization coefficient $\mathcal{K}$ evaluates the conflict between two sources. An additionnal aspect of the Dempster-Shafer's theory concerns the attenuation of the basic belief assignment m_j by a coefficient α_j for a source S_j. For all $H_n \subseteq \Theta$, the attenuated belief function can be written as :

$$m_{(\alpha,j)}(H_n) = \alpha_j.m_j(H_n) \quad (8)$$

$$m_{(\alpha,j)}(\Theta) = 1 - \alpha_j + \alpha_j.m_j(\Theta). \quad (9)$$

3 Methodology of classification process

The proposed methodology can be decomposed in three steps. The first one corresponds to the basic belief assignment based on analysis of the learning set (see section 3.1). The second one consists in attenuating the belief structure by means of a coefficient α_j derived from the Hellinger's distance between probability distributions. This one has a lower bound equal to 0 and an upper bound equal to 1. This distance allows to estimate the similarity between two probability distributions and, in particular to check if the gaussian assumption is correct (see 3.2). Finally, the belief structures defined for each source of information are aggregated in order to decrease significantly the uncertainty for the later classification process (see 3.3).

3.1 Basic Belief Assignment

An important aspect of the classification concerns learning knowledge using data. In evidence theory, this problem leads to initialize the belief functions m. We make the hypothesis that the data extracted from one information source S_j among M sources can be represented as a gaussian distribution. This assumption is obtained by means of the study of the learning database defined as : $\mathcal{X} = \{\mathcal{X}_{(n;1)}, \ldots, \mathcal{X}_{(n;M)}\}$ where $\mathcal{X}_{(n;j)} = \{X_{(n;j)}\}$ represents the set of vectors X_n classified in the hypothesis H_n. For the value x_j, we determine the membership probability according to the hypothesis as :

$$P(x_j/H_n) = \frac{1}{\sigma_{(n;j)}\sqrt{2\pi}} e^{\frac{-(x_j - \mu_{(n;j)})^2}{2\sigma^2_{(n;j)}}} \quad (10)$$

that is to say:

$$P(x_j/H_n) = \mathcal{N}(\mu_{(n;j)}, \sigma_{(n;j)}). \quad (11)$$

The pair $(\mu_{(n;j)}, \sigma_{(n;j)})$ represents respectively the mean and the standard deviation computed after the learning step for each hypothesis H_n and each source S_j. In addition, we compute

a third gaussian distribution representing the conjunction of the two hypotheses. This new distribution has the following mean and standard deviation :

$$\mu_{((n,n');j)} \triangleq \frac{\mu_{(n;j)} + \mu_{(n';j)}}{2} \qquad (12)$$

$$\sigma_{((n,n');j)} \triangleq \max(\sigma_{(n;j)}, \sigma_{(n';j)}). \qquad (13)$$

This assumption allows to generate the belief functions. Let X' a M component vector to be classify with $X' = [x'_1, \ldots, x'_M]^t$. The belief given for each hypothesis $H_n \in 2^\Theta$ depends on the membership probability with respect to :

$$m_j(H_n) = R_j * P(x'_j/H_n). \qquad (14)$$

The coefficient R_j is a normalization factor. It allows to verify the condition given by equation (2). It is defined for a hypothesis $H_n \in 2^\Theta$ as :

$$R_j = \frac{1}{\sum_{H_n \in 2^\Theta} P(x'_j/H_n)}. \qquad (15)$$

3.2 Belief function attenuation

After this learning step, the main idea is to resume the information contained in each source S_j by means of an optimum histogram computed on the set $\bigcup_{i \in H_n} \mathcal{X}_{(i;j)}$ in the sense of the maximum likelihood and of a mean square cost. This histogram will be used in order to establish the relevance of a source of information. First, we have to build an approximation of the unknown probability distribution with only the N-sample given in each source. That is done by means of a histogram building which is led by the use of an information criterion. We will see that different information criteria initially designed for model selection can be used.

3.2.1 Probability density approximation

Let be $A_1 A_2 \ldots A_p \ldots A_q$ an initial partition Q of an unknown distribution λ with $q = Card(Q)$. The aim is to approximate λ with a histogram built on a subpartition $C = B_1 B_2 \ldots B_c$ of Q with c bins such as $c \leq q$. The probability distribution $\hat{\lambda}_C$ built with C is an optimum estimation of λ according to a cost function to define. C results from an information criterion called IC issued from the basic Akaike's information criterion (AIC) [9], AIC^* or ϕ^* [10] which are respectively Hannan-Quinn's criterion and Rissanen's criterion. These criteria have the following form :

$$IC(c) = g(c) - \sum_{B \in C} \hat{\lambda}_c \ln \frac{\hat{\lambda}_c(B)}{\nu_c(B)} \qquad (16)$$

where $g(c)$ is a penalty which differs from one criterion to another one. Let us note ε a random process of a probability distribution λ supposed absolutely continuous to an *a priori* given probability distribution ν. Let ω be the set of all values taken by ε. The probability density f of λ is given by the Radon-Nycodim's derivative such as :

$$\forall \epsilon \in \omega \qquad f(\lambda, \epsilon) = \frac{d\lambda}{d\nu}(\epsilon). \qquad (17)$$

The probability density f is approximated from N samples (ϵ_k) of ε by means of a histogram with c bins obtained with these N values. An optimum histogram to approximate the unknown probability distribution λ is obtained in two steps. The first one consists in merging two contiguous bins in a histogram with c bins among the $(c-1)$ possible fusions of two bins. This is made by minimizing the IC criterion. The second one consists in finding the "best" histogram with c bins. The optimum histogram with $c = c_{opt}$ bins is the one which minimizes IC.

3.2.2 Maximum likelihood estimator for a partition Q

Let Q be a partition with q bins and let $\epsilon_1 \ldots \epsilon_N$ be a N-observation sample and let be λ_Q the probability distribution according to Q. The maximum likelihood estimator $\hat{\lambda}_Q$ of λ_Q is given by the following equation :

$$\forall p \in \omega \qquad \hat{\lambda}_Q(A_p) = \frac{1}{N} \sum_{\epsilon_k \in A_p} \epsilon_k \qquad (18)$$

where A_p is a bin of the partition Q. This result derives from the density expression of λ_Q :

$$\forall \epsilon \in \omega \qquad f(\lambda_Q, \epsilon) = \sum_{A \in Q} \frac{\hat{\lambda}_Q(A)}{\nu(A)} 1_A(\epsilon) \quad (19)$$

with $1_A(\epsilon) = 1$ if $\epsilon \in A$ and 0 otherwise.

3.2.3 Selection of the bin number of a histogram

The obtaining of the optimum histogram is based on the use of an information criterion IC which gives the number of bins optimal thanks to a cost function based on the Kullback's contrast or the Hellinger's distance. We define the cost to take $\hat{\lambda}$ when λ is the true probability density by :

$$W(\lambda, \hat{\lambda}) = E_\lambda \left(\psi \left[\frac{f(\hat{\lambda}, \epsilon)}{f(\lambda, \epsilon)} \right] \right) \quad (20)$$

where E_λ is the mathematical expectation according to λ and ψ is a convex function. According to the expression of ψ the cost function leads to different information criteria to choose the histogram with c bins. So, if ψ is the Hellinger's distance we get :

$$AIC(c) = \frac{2c-1}{N} - 2 \sum_{B \in C} \hat{\lambda}_c(B) \ln \frac{\hat{\lambda}(B)}{\nu(B)}. \quad (21)$$

It can be seen that it is identical to the classical Akaike's information criterion. If the cost function $W(\lambda, \hat{\lambda})$ is expressed according to the KullBack's contrast, we obtain two new criteria such as :

$$\phi^*(c) = \frac{c(1 + \ln(\ln N))}{N} - 2 \sum_{B \in C} \hat{\lambda}_c(B) \ln \frac{\hat{\lambda}(B)}{\nu(B)} \quad (22)$$

$$AIC^*(c) = \frac{c(1 + \ln N)}{N} - 2 \sum_{B \in C} \hat{\lambda}_c(B) \ln \frac{\hat{\lambda}(B)}{\nu(B)}. \quad (23)$$

These criteria can be used to select the optimum histogram with c bins to approximate the unknown probability density of a N-sample. Detailed demonstrations are available in [9, 10].

3.2.4 Optimum histogram building process

At first, an initial histogram with $q = Card(Q) = 2.In[\sqrt{N} - 1]$ bins is built giving the partition Q, where $In[]$ denotes the integer part [11]. Then, a partition with $(q-1)$ bins is considered. For each possible fusion of two contiguous bins among $(q-1)$ the criterion $IC(q-1)$ is computed. The choice of the best fusion is made according to the minimization of $IC(q-1)$. When it is done, we look for the best partition with $(q-2)$ bins according to the same rule. Finally, the histogram with c bins such as $IC(c)$ for $c \in \{1, \ldots, q\}$ is retained. Figure 1 shows an initial histogram built with a N-sample ($N = 90$) randomly generated according to a gaussian distribution with mean equal to 0 and with a variance equal to 1. This initial histogram is made of $16 = 2.In[\sqrt{90} - 1]$ bins. Final histograms according to respectively

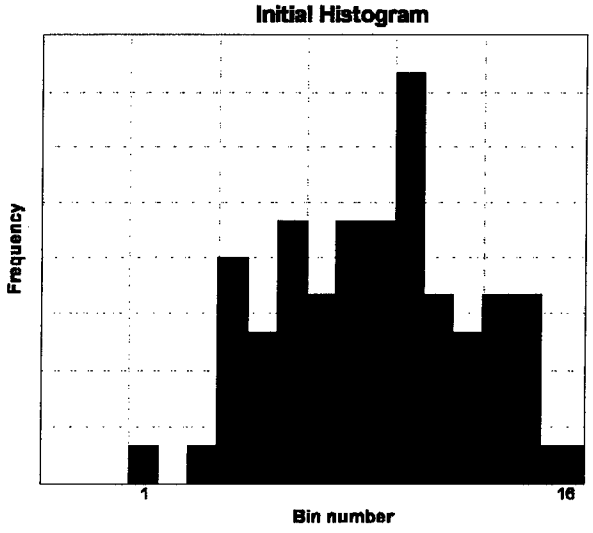


Figure 1: Initial histogram

AIC, AIC^* and ϕ^* are given in figures 2, 3, 4. Figure 5 gives the behaviour of the three

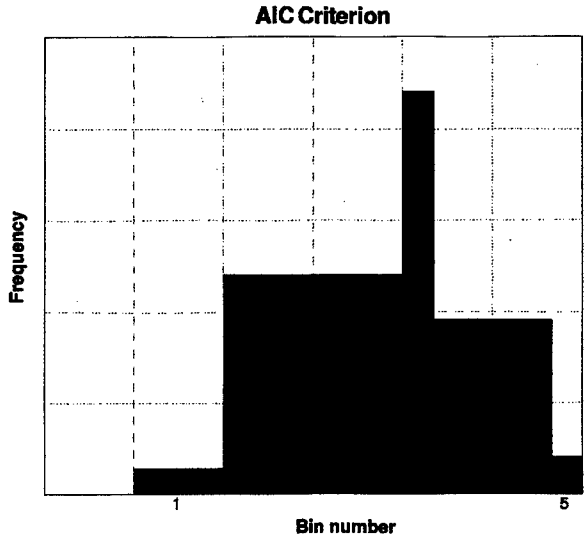


Figure 2: Optimum histogram with AIC

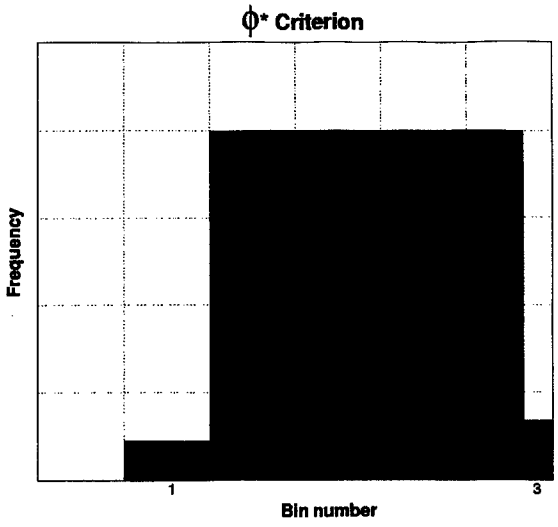


Figure 4: Optimum histogram with ϕ^*

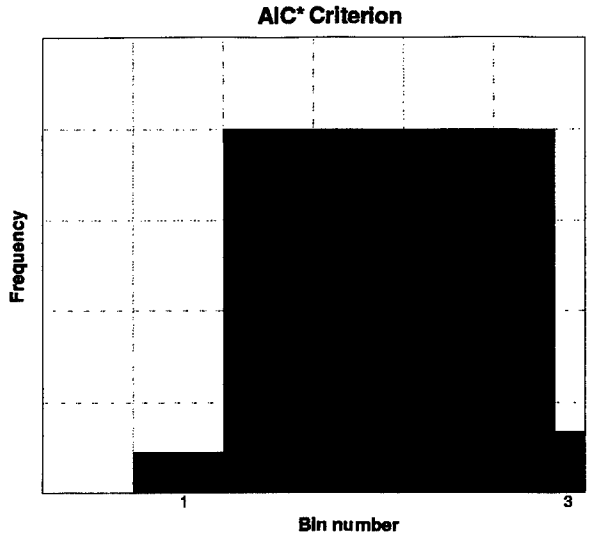


Figure 3: Optimum histogram with AIC^*

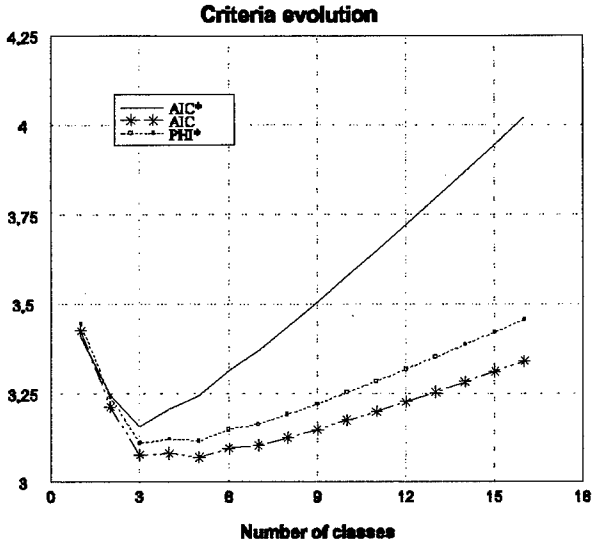


Figure 5: Criteria evolutions

criteria. It can be seen that AIC^* and ϕ^* give the same final histogram. AIC gives a final histogram with an upper bin number. This difference is linked to the type of convergence for each information criterion [10].

The optimum histogram is computed on the set $\bigcup_{i \in H_n} \mathcal{X}_{(i;j)}$. Once this histogram is obtained, we use the Hellinger's distance between the approximated distribution $\hat{\lambda}_C$ computed on the set $\mathcal{X}_{(n;j)}$ and the approximated distribution $\hat{\lambda}'_C$ computed on the set $\mathcal{X}_{(n';j)}$. This distance gives a dissimilarity between the two probability densities that is to say the ability of the source to distinguish the two hypotheses H_n and $H_{n'}$.

3.3 Information sources aggregation and decision

We attenuate the belief structures according to the equations (8) and (9) where α_j is the

Hellinger's distance :

$$\alpha_j = \psi\left[\frac{f(\hat{\lambda},\epsilon)}{f(\lambda,\epsilon)}\right] = 4\sum_{B\in C}\lambda(B)\left[\sqrt{\frac{\hat{\lambda}(B)}{\lambda(B)}}-1\right]^2. \quad (24)$$

The information sources S_j are then aggregated using the Dempster's combination rule (see equation (5)). Finally, the decision is made by assigning the vector X' to the hypothesis H_n with the maximum credibility. The decision rule is based on the decision function δ which assignes a vector X' to the hypothesis H_n following :

$$\delta(X', H_n) = n \text{ iif } H_n = \arg\max_{H_i\in 2^\Theta}(Bel(H_i)) \quad (25)$$

4 Simulations

The proposed method has been applied to several sets of artificial and real data in order to perform an evaluation of the algorithm.

4.1 Synthetic data

In this section, we present results obtained on synthetic data. For the simulations, we have generated three gaussian distributions such as : $\mu_1 = [1,1,1]^t$ and $\sigma_1^2 = 1$; $\mu_2 = [-1,1,0]^t$ and $\sigma_2^2 = 4$; $\mu_3 = [0,-1,1]^t$ and $\sigma_3^2 = 3$. The first learning set is made of $N = 90$ elements (30 for each class) and the second one is made of $N = 200$ elements (70 elements in the first class, 50 elements in the second class and 80 elements in the third class). The test base is made of 600 elements. Our method is compared to the method proposed in [12]. The results are given in the two following tables for the first learning set.

For the method proposed by Zouhal, the good classification rate is of 59.16% and 62.50% for our method. According to the second learning set, we get the following results (see tables 3 and 4).

For the method proposed by Zouhal, the good classification rate is of 57.83% and 60.33% for our method.

Table 1: Results of method [12]

	Classified		
Presented	C_1	C_2	C_3
C_1	81	7.5	11.5
C_2	29.5	43.5	27
C_3	31	16	53

Table 2: Results of our method

	Classified		
Presented	C_1	C_2	C_3
C_1	87	3.5	9.5
C_2	33	42.5	24.5
C_3	27	15	58

4.2 Real data

A second database concerns a set of 16 characteristics extracted from 122 images of dermatological lesions. Details concerning the features can be found in [13]. The database is composed of 101 naevi (no pathological lesions) and 21 melanoma. Final results are presented in the following tables (Tables 5 and 6). The proposed method allows to obtain 81.1% of good classification towards 75.7% for the method presented in [12].

5 Conclusion

In this paper, we have presented an original method of classification using both information criteria and Dempster-Shafer's theory. The proposed methodology consists in initializing the belief functions with probability densities obtained by learning. By means of informa-

Table 3: Results of method [12]

	Classified		
Presented	C_1	C_2	C_3
C_1	78.5	4.5	17
C_2	29.5	47	23.5
C_3	26	26	48

Table 4: Results of our method

	Classified		
Presented	C_1	C_2	C_3
C_1	83.5	6.5	10
C_2	38	41	21
C_3	27.5	16	56.5

Table 5: Results of method [12]

	Reality	
Decision	Naevus	Melanoma
Naevus	99	1
Melanoma	47.6	52.4

tion criteria, we determine the attenuation of the belief assignment based on the dissimilarity between probability distributions. Results on artificial and real data demonstrate the effectiveness of the proposed method. Concerning the real-world data (diagnosis in dermatology), tests on a larger base are processing at this time. Future work is concerned with analysis of several decision rules using uncertainty measures proposed by Klir [14, 15].

6 Acknowledgements

The authors would like to acknowledge the support of the Normandy regional council. The co-operation of P. Joly of the clinic of dermatology of Charles Nicolle's hospital of Rouen in supplying diagnosis on melanoma and acquisition of image lesions, is also gratefully acknowledged. Finally, authors would like to express their appreciation to Professor D. De Brucq, whose comments led to a significant improvement of the paper.

Table 6: Results of our method

	Reality	
Decision	Naevus	Melanoma
Naevus	86.1	13.9
Melanoma	23.8	76.2

References

[1] J.C. Bezdek. Fuzziness vs. probability - the n-th round. *IEEE Trans. on Fuzzy Systems*, 2(1):1–42, 1994.

[2] L.A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.

[3] D. Dubois. Belief structures, possibility theory and decomposable confidence measures on finite sets. *Comput. Artif. Intell.*, 5(5):403–416, 1986.

[4] D. Dubois, J. Lang, and H. Prade. Automated reasoning using possibilistic logic: Semantics, belief revision, and varaible certainty weights. *IEEE Trans. on Knowledge and Data Engineering*, 6:64–71, 1994.

[5] G. Shafer. *A Mathematical Theory of Evidence.* Princeton University Press, 1976.

[6] P. Smets. Belief functions: The disjonctive rule of combination and the generalized bayesian theorem. *Int. J. of Approximate Reasoning*, 9:1–35, 1993.

[7] P. Smets and R. Kennes. The transferable belief model. *Artificial Intelligence*, 66:191–234, 1994.

[8] T. Denoeux. Analysis of evidence-theory decision rules for pattern classification. *Pattern Recognition*, 30(7), 1998.

[9] O. Colot. *Apprentissage et détection automatique de changements de modèles - Application aux signaux électro-encéphalographiques.* PhD thesis, Université de Rouen, 1993.

[10] O. Colot and al. Information criteria and abrupt changes in probability laws. In M. Holt, C. Cowan, P. Grant, and W. Sandham, editors, *Signal Processing VII: Theories and Applications*,

pages 1855–1858. EUSIPCO'94, September 1994.

[11] Y. Sakamoto, M. Ishiguro, and G. Kitagawa. *Akaike Information Criterion Statistics.* KTK Scientific Publishers, Tokyo, 1986.

[12] L. M. Zouhal and T. Denoeux. An adaptative k-NN rule based on dempster-shafer theory. In *Proc. Of the 6th Intern. Conf. On Comput. Analys. Of Imag. and Pattern*, pages 310–317. ICAIP'95, Springer Verlag, September 1995.

[13] O. Colot and Al. *A Colour Image Processing Method for Melanoma Detection.* Lecture Notes in Computer Science, 1496. Springer-Verlag, October 1998.

[14] G.J. Klir and T.A. Folger. *Fuzzy Sets, Uncertainty, and Information.* Prentice Hall P T R, Englewood Cliffs, New Jersey 07632, 1988.

[15] Z. Wang and G.J. Klir. *Fuzzy Measure Theory.* Plenum Publishing Corporation, 233, Spring Street New York, 1992.

Recursive Composition Inference for Force Aggregation

Jason K. Johnson
Information Technology Division
Alphatech, Inc.
Burlington, MA

Ronald D. Chaney
Information Technology Division
Alphatech, Inc.
Burlington, MA

Abstract - *This paper describes recursive composition inference techniques (identified with inference on a Bayesian network) to efficiently and optimally reason about military units from partial observations of constituent vehicles. This model-based approach requires statistical models of the composition of military units as well as of the processes of vehicle detection and observation. Given these statistical models the Bayesian network both infers unit-type and refines vehicle-type conditioned upon partial observations of the detected vehicles. Monte Carlo experiments are performed to provide performance estimates under variable conditions. The main innovation of this approach is the recursive decomposition of composition hypotheses to efficiently infer between vehicle-level hypotheses and unit-level hypotheses. The primary benefit of this work is that it provides a means for reasoning about aggregates of objects which is simple, robust and rigorous.*

Key Words: force aggregation, situation refinement, Bayesian inference, Bayesian networks.

1. Introduction

Military units organize themselves in hierarchical structures to enable efficient deployment, operation, and command and control. Determining the hierarchical structure of an opposing force is critical to determining its capability, threat, and ultimately its intent. Sensor systems can detect and take measurements on individual entities, such as vehicles and installations. These measurements can be used to infer the class of individual entities. However, few collection assets provide direct measurements on the hierarchical force structure of units that the entities comprise. Consequently, it is desirable to develop the capability to draw inferences on the hierarchical structure of military units based on inferences and measurements of individual entities and sub units.

This paper describes a technique for assessing the relative merit of force aggregate hypotheses from partial observations of a set of entities. That is, given partial observations of entities that comprise military units, the technique draws inferences about the type of military unit that is present. Furthermore, drawing inference about the type of military unit provides contextual information that enables improved inference about the type of individual vehicles.

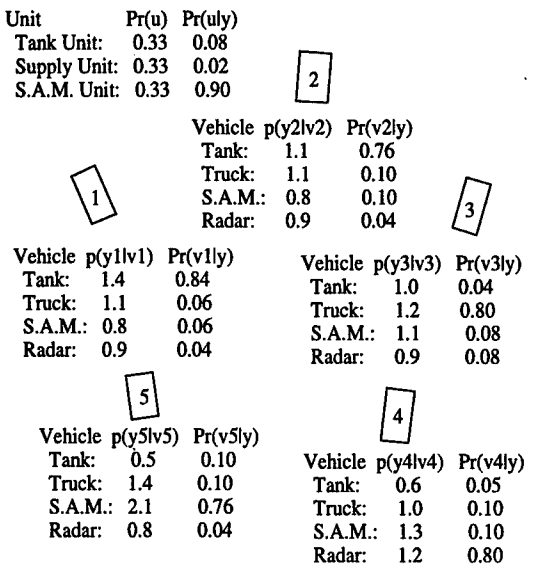


Figure 1. Illustration of a notional estimation scenario.

2. Model Specification

The force aggregation problem solved in this paper may be stated as follows. The available data are partial observations of a military unit in the form of target reports for detected targets. Each target report provides partial information as to the classification of the detected target. The objective is then twofold. First, fuse these target reports to infer unit-type. Second, exploit context to refine vehicle-type. Figure 1 provides a notional scenario to illustrate this problem. The solution described in this paper operates strictly by reasoning about the composition of military units.

In order to perform probabilistic inference we must formulate a probabilistic model for the underlying process being observed. The model we will employ in this paper is developed below. This model consists of three tiers: the composition model of the military units; the detection model of the vehicle detector; and the measurement model of the process by which the type of each detected vehicle is partially observed. A diagram of the overall structure of the model is provided in Figure 2.

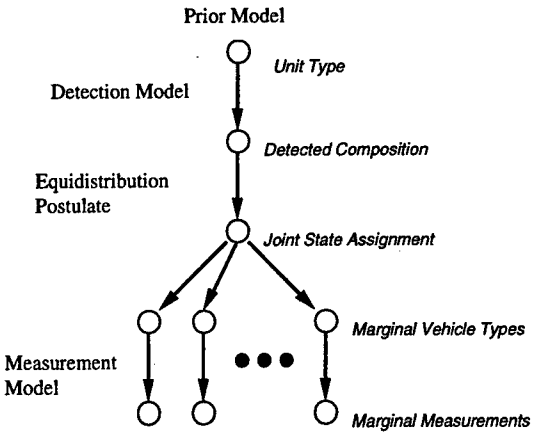


Figure 2. Diagram of Bayes net used to model fusion problem. Note, the joint-state assignment node is replaced by a recursive binary tree model in section 3.2 (Figure 5).

2.1 The Composition Model

The composition model specifies what unit-types may occur, the composition of these units and the prior probability of the occurrence of each unit-type. Unit-type is modeled by the random variable U assuming the values $u=1,\ldots,q$ with prior probability $\Pr(U=u)$. Vehicle-type is modeled by the random variable V assuming the values $v=1,\ldots,r$. The composition of each unit-type is then specified as the number of instances of each vehicle-type present within the unit. The number of vehicles of type v present within units of type u is denoted $n(v;u)$ for $v=1,\ldots,r$ and $u=1,\ldots,q$. Hence, specification of the composition model consists of an r x q matrix of non-negative integers (vehicle counts) and a q-length vector of prior probabilities (which sum to unity). Note, this composition model could easily be extended to also model uncertainty in the composition of a given unit-type. However this extra layer of uncertainty is omitted in this paper.

2.2 The Detection Model

The detection model is a statistical model of the process of detecting military vehicles in the signal-level data generated by some sensor. There are three sources of uncertainty that tend to obscure the identity of the military unit: undetected vehicles, extraneous clutter vehicles and false alarms. The detection model is included so as to provide robustness against these anticipated ATR operating conditions. The occurrence of these failures of detection are statistically modeled by the Bernoulli probability of detection as a function of vehicle-type $P_D(v)$ the Poisson clutter rate also as a function of vehicle type $\lambda_C(v)$ and the Poisson false-alarm rate λ_{FA}. A diagram illustrating the cumulative effect of these three detection processes is shown in Figure 3. The "null hypothesis" $v=0$ is introduced to denote false alarms.

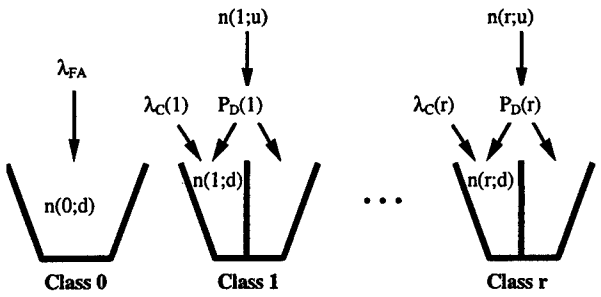


Figure 3. Diagram of the detection model.

As indicated in Figure 2, the role of the detection model is to specify the transition probabilities from the unit-type (of known composition) to the detected composition (the composition of the set of detected vehicles). However, these transition probabilities depend upon the number of detected vehicles n (implicitly part of our observation of the military unit). Once the number of detections is received the transition probabilities may then be computed as shown below. First, due to the independence of the various detection processes this transition probability from unit composition u to detected composition d is separable by vehicle-type (the number of objects accumulated in each of the left-hand bin partitions of Figure 3 are conditionally independent given the unit-type).

$$(1) \qquad \Pr(d \mid u) = \Pr(n(0;d)) \prod_{v=1}^{r} \Pr(n(v;d) \mid n(v;u))$$

Note, the notation $n(v;d)$ indicates the number of detected vehicles of class v. The probability of the number of detected vehicles of each type v is then computed as

$$(2) \qquad \Pr(n(v;d) \mid n(v;u)) = \sum_{k=0}^{n(v;u)} b(k; n(v;u), P_D(v)) \times p(n(v;d)-k; \lambda_C(v))$$

where $b(k;n,p)$ are the binomial probabilities

$$(3) \qquad b(k;n,p) = \begin{cases} \binom{n}{k} p^k (1-p)^{k-n}, & k \in [0,n] \\ 0, & k \notin [0,n] \end{cases}$$

and $p(k;\lambda)$ are the Poisson probabilities

$$(4) \qquad p(k;\lambda) = \begin{cases} \lambda^k e^{-\lambda}/k!, & k \geq 0 \\ 0, & k < 0 \end{cases}$$

The number of false alarms generated by the detector is also modeled by (4). These transition probabilities will be used in Section 3 to infer between unit-type and detected composition.

2.3 The Measurement Model

The measurement model is a statistical model of the process of observing a detected vehicle so as to produce a measurement of that vehicle. Each vehicle measurement y is modeled as being a random function of the observed vehicle's type v. This random function is statistically characterized by a measurement probability density function $p(y|v)$ for each of the possible vehicle types $v=1,\ldots,r$. Also, if false-alarms are modeled then a measurement model must be provided for those as well.

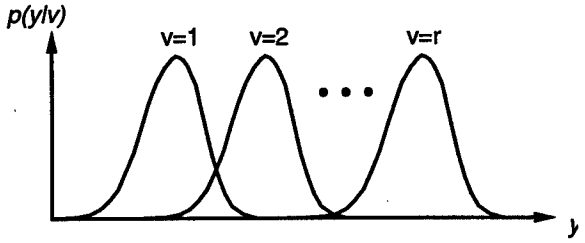


Figure 4. Plot of a notional measurement model.

3. Composition Inference Algorithms

The statistical models specified above provide the basis for implementing rigorous probabilistic inference algorithms for the optimal estimation of both unit-type and vehicle-type from partial observations of the detected vehicles. The data available to the inference algorithms are the number of detected vehicles n as well as a measurement of each of these vehicles y_k for $k=1,\ldots,n$. For the remainder of this paper y will denote the set of all such observations. Given these data the objective of the inference algorithm is then twofold. First, calculate the unit-type probabilities conditioned upon all observations $\Pr(U=u \mid y)$ for $u=1,\ldots,q$. Second, calculate the vehicle-type probabilities conditioned upon all observations $\Pr(V_k=v \mid y)$ for $k=1,\ldots,n$ and $v=0,1,\ldots,r$. These conditional probabilities then provide the basis for rendering optimal marginal estimates of the unit-type of the observed unit and of the vehicle-type of the detected vehicles. These inference calculations are outlined below.

To infer unit-type from vehicle observations, first the measurement data are pre-conditioned under the measurement model. This involves computation of the measurement likelihood conditioned upon vehicle-type $p(y_k \mid V_k=v)$ for each vehicle $k=1,\ldots,n$ as a function of vehicle-type $v=0,1,\ldots,r$. At this point the measurements themselves may be discarded as these likelihoods are sufficient statistics of those measurements for the purposes of inferring both unit-type and vehicle-type.

Next, the composition of detected vehicles is inferred by either of the methods discussed below in sections 3.1 and 3.2. This involves calculation of the measurement likelihood $p(y|d)$ as a function of detected composition d. The fundamental postulate underlying either method is the Equidistribution Postulate stated below.

Equidistribution Postulate: *Conditioned upon the composition of some set of vehicles, all joint-state assignments of vehicle-type to those vehicles consistent with the composition constraint are equally probable.*

To state this mathematically the set of all joint-state assignments (of vehicle classes to vehicles) consistent with a specified detected composition hypothesis d is denoted as $\Omega(d)$. Hence (according to the Equidistribution Postulate) the conditional probability of an assignment a conditioned upon a composition d is given by

$$(5) \qquad \Pr(a \mid d) = \begin{cases} 1/|\Omega(d)|, & a \in \Omega(d) \\ 0, & a \notin \Omega(d) \end{cases}$$

where $|\Omega(d)|$ is the degeneracy of the composition (the number of assignments of vehicle types to vehicles consistent with the composition) given by the multinomial coefficients.

$$(6) \qquad |\Omega(d)| = \binom{n(d)}{n(1;d),\ldots,n(r;d)}$$

Once the measurement likelihood as a function of the detected composition has been inferred, the measurement likelihood conditioned upon unit-type may then be inferred employing the transition probabilities from unit composition to detected composition as computed under the detection model (1-4).

$$(7) \qquad p(y \mid u) = \sum_{d \in D(n,r+1)} p(y \mid d) \Pr(d \mid u)$$

This sum is taken over the set $D(n,r+1)$ of all possible distributions of the n detected vehicles into the $r+1$ vehicle-classes (included the false-alarm class). Finally, the unit-type probabilities are computed according to Bayes rule from the measurement likelihoods $p(y \mid u)$ and the prior model $\Pr(u)$.

$$\Pr(u \mid y) = \frac{p(y \mid u)\Pr(u)}{p(y)} \tag{8}$$

The denominator $p(y)$ is the likelihood of all measurements which simply normalizes the relative probabilities computed by the numerator.

$$p(y) = \sum_{u=1}^{q} p(y \mid u)\Pr(u) \tag{9}$$

This is the basic structure of the unit-type inference algorithm (except for the composition inference step which is discussed in sections 3.1 and 3.2).

To refine vehicle-type the detected composition (implicitly conditioned upon the number of detected vehicles) is inferred from unit type using the prior model and the detection model.

$$\Pr(d) = \sum_{u=1}^{q} \Pr(d \mid u)\Pr(u) \tag{10}$$

Then vehicle-type probabilities $\Pr(V_k=v \mid y)$ for each vehicle $k=1,\ldots,n$ as a function of vehicle-type $v=0,\ldots,r$ are inferred from the detected composition probabilities $\Pr(d)$ and the Equidistribution Postulate. This calculation is deferred to section 3.2.

3.1 Brute Force Composition Inference

Before developing the recursive composition inference techniques in the next subsection a simpler brute force approach is considered. This calculation operates by performing inference with respect to the set of all possible joint-states of the detected vehicles. Due to the independence of the vehicle measurements, the measurement likelihood conditioned upon the joint state assignment is simply the product of the marginal measurement likelihoods conditioned upon the respective marginal vehicle types.

$$p(y \mid a) = \prod_{k=1}^{n} p(y_k \mid v_k) \tag{11}$$

The likelihood of all measurements conditioned upon the detected composition may then be computed from the assignment likelihoods (11) and the transition probabilities (5).

$$\begin{aligned} p(y \mid d) &= \sum_{a} p(y \mid a)\Pr(a \mid d) \\ &= \frac{1}{|\Omega(d)|} \sum_{a \in \Omega(d)} p(y \mid a) \end{aligned} \tag{12}$$

This likelihood computation is performed for every possible distribution d of the n detected vehicles into the $r+1$ vehicle-classes. Once this likelihood function has been computed the unit-type may then be inferred as described previously (7-10).

Under this approach the likelihood calculation (11) must be computed for each of the $(r+1)^n$ possible joint-state assignments of the detected vehicles such that the complexity of these calculations is $O(n(r+1)^n)$. This prohibitive computational complexity is the motivation for the recursive composition inference techniques developed in the next section.

3.2 Recursive Composition Inference

In this section we develop recursive composition inference techniques which offset the computation burden of the simplistic approach described above. This technique avoids considering the set of $(r+1)^n$ joint-states of the detected vehicles by recursively partitioning this set of vehicles into half-sets and inferring between the composition of the halves and the composition of the whole. Figure 5 illustrates the Bayesian structure of this decomposition. Note that Figure 5 is an alternate expansion of the middle three tiers of the Bayes net shown in Figure 2 (replacing the "brute force" inference technique depicted there).

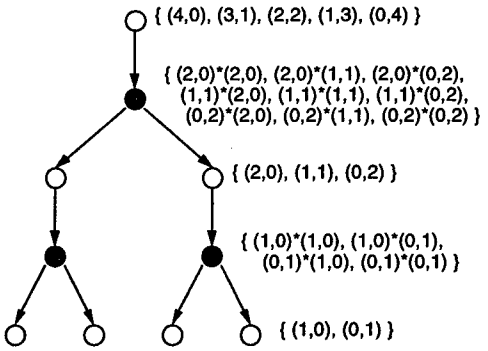


Figure 5. Bayes net to recursively infer composition. For this example there are four vehicles which submit to two classifications.

There are two types of nodes depicted in Figure 5. The state-space of a white node corresponds to the set

of all possible compositions of the vehicles beneath that node (the leaf nodes). For example, the states of a leaf node correspond directly to the vehicle-type of the associated vehicle, also the states at the root node correspond to the composition of all detected vehicles. In order to statistically relate the composition of a set of vehicles to the composition of its two subsets the grey nodes are introduced such that the state of the subtrees are conditionally independent given the state of the grey node. This is achieved simply by choosing the states of the grey node to be the joint-state of its two white-node children. Hence the states of a grey node essentially represent hypotheses as to how the composition of the whole is distributed between the halves. The composition of the two halves may then be considered independently.

The transition probabilities from a white-node state d_T (specifying the total composition of the n_T vehicles contained within that subtree) to a grey-node state (d_L, d_R) (specifying the respective composition of the n_L and n_R vehicles in the left and right subtrees) are determined by the Equidistribution Postulate.

$$\Pr(d_L, d_R \mid d_T) = \begin{cases} \dfrac{|\Omega(d_L)| \times |\Omega(d_R)|}{|\Omega(d_T)|}, & d_L + d_R = d_T \\ 0, & d_L + d_R \neq d_T \end{cases} \tag{13}$$

The requirement that $d_L+d_R=d_T$ indicates that the sub-compositions must be consistent with the total composition which is to say that $n(v;d_L)+n(v;d_R)=n(v;d_T)$ for $v=0,\ldots,q$. The transition probabilities from grey nodes to white nodes are trivial since the state of the grey node is just the joint-state of its two white-node children.

Below the standard Bayesian inference algorithms (see [2]) are adapted to take advantage of the sparsity of the transition probabilities described above (for the sake of efficient computation) as well as to exploit the special structure of (13) (to simplify the prediction operations).

Recursive Composition Inference

The objective of recursive composition inference is to propagate the information contained in the measurements at the leaves of the tree upwards fusing this information at each node s to yield statistics concerning the set of all measurement on that subtree y_s. Rather then propagating the traditional likelihood function $p(y_s|d_s)$ we instead propagate the likelihood function weighted by the degeneracy of the composition $|\Omega(d_s)|$ which is denoted by $\lambda_s(d_s)$.

$$\lambda_s(d_s) = p(y_s \mid d_s) \times |\Omega(d_s)| \tag{14}$$

The advantage of propagating these functions is that the degeneracy weights cancel with the transition probabilities (13) such that prediction and merging of these messages from the two subtrees simplifies to the formula below.

$$\lambda_T(d_T) = \sum_{\substack{d_L, d_R \\ d_L + d_R = d_T}} \lambda_L(d_L) \times \lambda_R(d_R) \tag{15}$$

Here the sum is taken over all possible compositions of each of the two subtrees (d_L and d_R respectively) under the constraint that their composition sum to the composition d_T. An algorithm for accomplishing this calculation is outlined below.

```
Inference Code
Inputs: λ_L, λ_R
Outputs: λ_T
Initialize λ_T
for d_T ∈ D(n_T, r+1)
  λ_T(d_T) = 0
end
Accumulate λ_T
for d_L ∈ D(n_L, r+1)
  for d_R ∈ D(n_R, r+1)
    let d_T = d_L + d_R
    λ_T(d_T) = λ_T(d_T) + λ_L(d_L) × λ_R(d_R)
  end
end
```

This code fragment is recursively applied to the Bayes net starting at the bottom of the tree and propagating statistics up the tree until the root node is reached. This recursion is initialized at the leaves according to the measurement model (note, there is no degeneracy at the leaf nodes). At the root node of the tree the likelihood function $p(y|d)$ is recovered from $\lambda(d)$ by dividing by the degeneracy function (6) in accordance with (14).

Once these measurement likelihoods are computed the unit-type may then be inferred by formulas (7-9). These probabilities may then be used to render an optimal marginal estimate of unit-type.

Recursive Composition Refinement

The objective of recursive composition refinement is to propagate prior information down the tree as well as to redistribute the information passed up from each subtree to the other subtree so as to compute statistics

at each node s conditioned upon all measurements not on that subtree $\bar{y}_s = y \setminus y_s$. Rather than propagating the traditional probability mass function $\Pr(d_s \mid \bar{y}_s)$ we instead propagate a function $\pi_s(d_s)$ proportional to the probability mass function weighted by the inverse of the degeneracy $|\Omega(d_s)|$.

$$(16) \qquad \pi_s(d_s) \propto \frac{\Pr(d_s \mid \bar{y}_s)}{|\Omega(d_s)|}$$

Again the degeneracy factor is included so as to cancel with the factors of the transition probabilities (13). The prediction of this function from parent node T to child nodes L and R then simplifies to the formula below.

$$(17) \qquad \begin{aligned} \pi_L(d_L) &= \sum_{d_R \in D(n_R, r+1)} \lambda_R(d_R) \times \pi_T(d_L + d_R) \\ \pi_R(d_R) &= \sum_{d_L \in D(n_L, r+1)} \lambda_L(d_L) \times \pi_T(d_L + d_R) \end{aligned}$$

An algorithm for accomplishing these calculations is outlined below. Note, this algorithm requires that the composition inference has already been executed such that the statistics $\lambda_s(d_s)$ are available for each node s.

```
Refinement Code
Inputs: π_T, λ_L, λ_R
Outputs: π_L, π_R
Initialize π_L
for d_L ∈ D(n_L, r+1)
  π_L(d_L) = 0
end
Initialize π_R
for d_R ∈ D(n_R, r+1)
  π_R(d_R) = 0
end
Accumulate π_L and π_R
for d_L ∈ D(n_L, r+1)
  for d_R ∈ D(n_R, r+1)
    let d_T = d_L + d_R
    π_L(d_L) = π_L(d_L) + λ_R(d_R) × π_T(d_T)
    π_R(d_R) = π_R(d_R) + λ_L(d_L) × π_T(d_T)
  end
end
```

By applying this code fragment to the Bayes net in a recursive down-sweep manner (start at the root node, compute statistics of children and recurse on subtrees) the statistics $\pi_s(d_s)$ are computed at every node s. This recursion is initialized at the root node according to (16) from the prior composition statistics (10) (conditioned upon the number of detections) and the degeneracy function (6). The final probability computation conditioned upon all observations as a function of the state of the node $\Pr(d_s|y)$ is computed by merging these statistics $\pi_s(d_s)$ with those computed during the composition inference algorithm $\lambda_s(d_s)$.

$$(18) \qquad \Pr(d_s \mid y) = \frac{\lambda_s(d_s) \times \pi_s(d_s)}{p(y_s \mid \bar{y}_s)}$$

The likelihood ratio shown in the denominator is simply the norm of the relative likelihoods computed by the numerator.

$$(19) \qquad p(y_s \mid \bar{y}_s) = \sum_d \lambda_s(d) \times \pi_s(d)$$

Formulas (18) and (19) are used to compute the refined vehicle-type probabilities which may then by used to render optimal marginal estimates of vehicle-type.

While the complexity of these recursive calculations is much improved over the brute-force approach, the complexity of the recursive calculations nevertheless grows rapidly with n ($O(n^2)$ for r=2). For this reason we briefly comment on the possibility of employing hypothesis pruning to limit the computational complexity of the recursive approach to $O(n)$ (at the expense of performing sub-optimal inference). This is accomplished by implementing a pruning operation in the composition inference upsweep such that only those N compositions maximizing the function $\lambda_s(d)/|\Omega(d)|$ (a lower-bound approximation of $p(y|d)$) are retained.

4. Monte Carlo Performance Estimation

In this section Monte Carlo simulation techniques are used to characterize the performance of the inference techniques developed above under various circumstances. In this section we characterize the difficulty of the problem by the parameters listed below. Unless otherwise stated, these parameters are chosen as follows.

n =3	number of vehicles per unit
q =2	number of unit-types
r =4	number of vehicle types
P_D =0.9	probability of detection
λ_C =0.25	vehicle clutter rate
λ_{FA} =1	detector's false alarm rate
J =4	divergence between vehicle-types

For each Monte Carlo trial we randomly select the q unit composition models by sampling the vehicle-type of each of the unit's n vehicles uniformly from the r vehicle-types. The true unit-type is then selected from among these q unit-types with uniform probability. The detection model is then applied to the composition of the true unit-type to simulate the number of detected vehicles of each target class (including the false-alarm class). Finally, the measurement likelihood function is simulated for each detected vehicle conditioned upon the vehicle-type as described below.

The divergence J is an information-theoretic measure (see [3]) of how easily two hypotheses 0 and 1 may be discriminated on average given a measurement y.

$$(20) \qquad J(0,1) = \int_y (p(y|1) - p(y|0)) \log \frac{p(y|1)}{p(y|0)} dy$$

For Gaussian measurement models having a common covariance P and separation between the means Δ the divergence is $J=\Delta^T P^{-1} \Delta$. In this case, the distribution of the log-likelihood-ratio $L(1|0)=\log(p(y|1)/p(y|0))$ is normally distributed having variance $J^{1/2}$ and mean values $-J/2$ and $+J/2$ under the respective hypotheses 0 and 1. Hence, the likelihood function may be simulated (up to an undetermined scale factor) under hypothesis 1 as $(\lambda(0)=e^L, \lambda(1)=1)$ for $L \sim N(-J/2, J^{1/2})$. We take the liberty of generalizing this model to the N-ary hypothesis case by independently sampling L for each confusor class (all vehicle-types except the true vehicle-type). The goal here is to simply characterize the uncertainty of the measurement process by the single parameter J in order to facilitate generic performance estimation.

Given the simulated target-reports we may then apply the inference techniques of section 3 to infer the hidden unit-type and the hidden vehicle-type of detected vehicles. These estimates are then compared to the true values and estimates of the probability of correct-classification are accumulated for both unit-level and vehicle-level estimates. Unless otherwise stated, each of the performance plots provided below are based upon 1000 independent iterations of the above simulation for each data-point.

First the performance of unit-type estimation is examined by plotting the probability of correct-classification as a function of the divergence between measurement models (Figure 6).

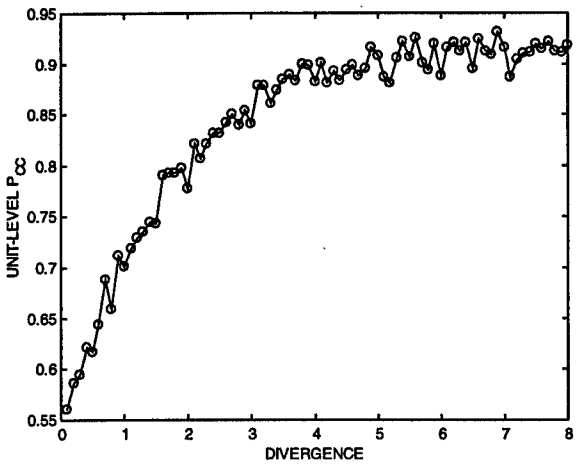


Figure 6. Plot of unit-type P_{CC} vs. J.

Likewise, the performance of vehicle-type estimation is measured by the vehicle-level probability of correct-classification plotted as a function of the divergence between measurement models (Figure 7). Here we plot the estimation performance for both the unrefined estimates (each vehicle's type is estimated so as to maximize the likelihood of just that vehicles measurement) and the refined estimates (conditioned upon all measurements).

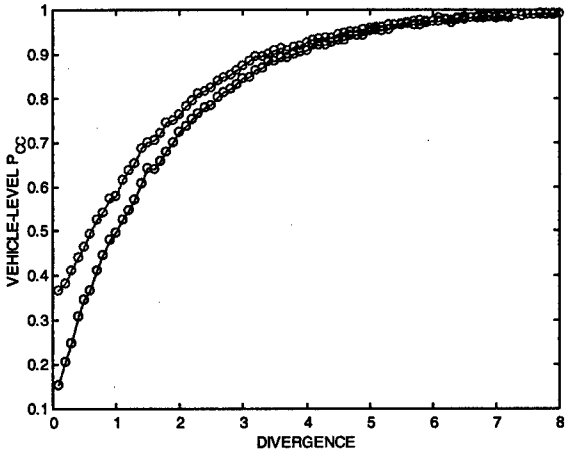


Figure 7. Plot of vehicle-type P_{CC} for refined (top) and unrefined estimates (bottom) vs. J.

It is also interesting to examine the unit-level classification performance as a function of the uncertainty of the detection process. Hence the unit-level P_{CC} is plotted as a function of P_D (Figure 8) and the clutter rate λ_C (Figure 9)

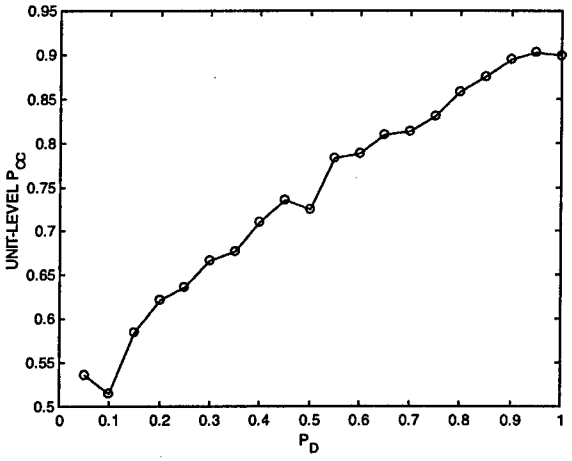


Figure 8. Plot of unit-type P_{CC} vs. P_D of detector.

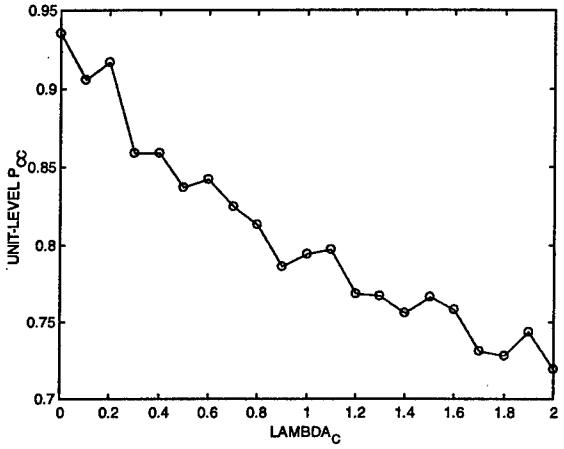


Figure 9. Plot of unit-type P_{CC} vs. clutter-rate λ_C.

Finally, the trade-off between classification performance and run-time introduced by hypothesis pruning is examined in Figure 10. This is a plot of P_{CC} vs. average run-time traced out by varying the maximum number of hypotheses control parameter N from 1 to 19 (both performance and run-time level off for larger values of N). For this simulation we chose the parameters q=2, n=3, r=6, P_D =1, λ_C=0, λ_{FA}=0, and J=1. Each of these data points represent 10000 Monte Carlo trials.

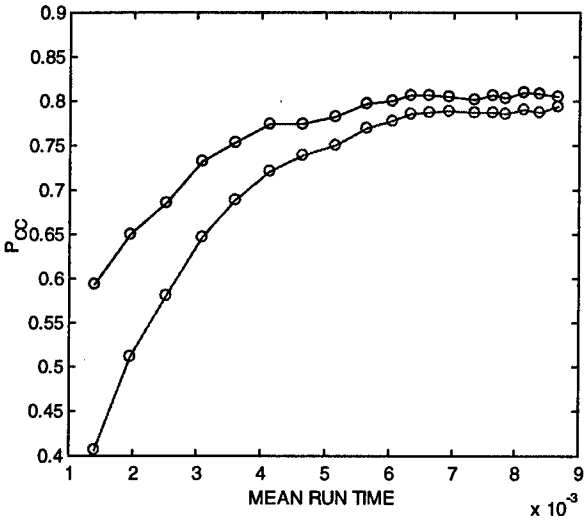


Figure 10. Plot of P_{CC} (upper curve is unit-type, lower-curve is vehicle-type) vs. average run-time (ms) traced out by varying the pruning control parameter N from 1 to 19.

5. Example Application

This section briefly introduces the application of the inference techniques developed in this paper to hierarchical force structures. The key idea here is that the same inference techniques used to infer unit-type from observations of the unit's constituent vehicles may also be employed to infer the type of a complex unit (consisting of multiple subunits) from knowledge of its subunits. This concept may be applied recursively to analyze an arbitrarily complex hierarchical force structure. This concept is illustrated below (Figure 11).

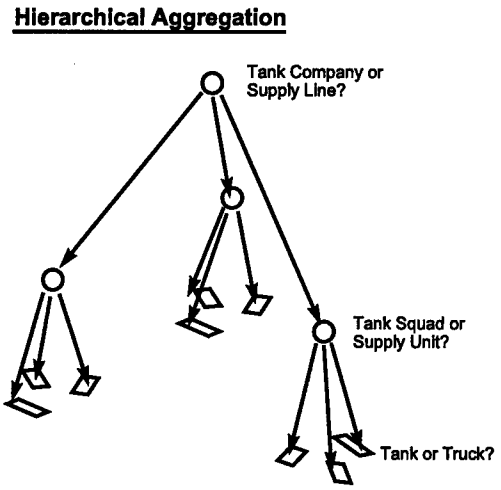


Figure 11. Illustration of hierarchical force inference.

An example of this technique is shown if Figures 12 and 13. Figure 12 depicts a simulated hierarchical force deployment. Partial observations of the vehicles within this hierarchy were simulated (J=1, P_D=1, λ=0) and provided to the hierarchical estimator. The resulting estimates are shown in Figure 13. All entities were classified correctly except that the *TelSquadConvoy* was classified as *TelSquadDefend.* This is because the composition-based inference cannot distinguish between the different configurations of a unit.

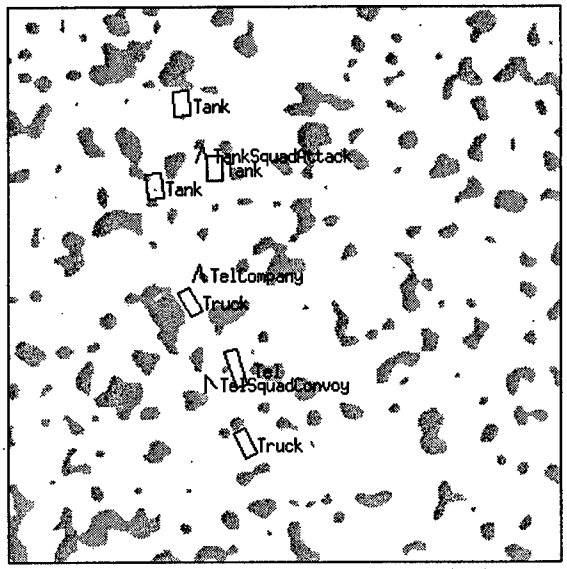


Figure 12. A simulated hierarchical force deployment.

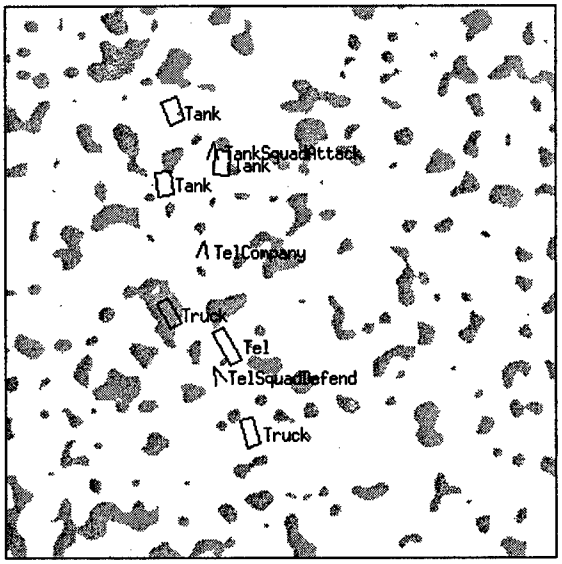


Figure 13. Estimates resulting from inference within this force structure inference.

6. Conclusions

In this paper we have described a simple, robust and rigorous technique for inferring the unit-type of a partially observed military unit as well as refining the vehicle-types of the detected vehicles. The simplicity of the model insures its generality, robustness and ease of model identification. Monte Carlo simulations demonstrate the utility and robustness of these techniques. In future research we plan to extend these techniques to model the spatial deployment of military units, to utilize terrain information, and to provide for automatic clustering of detected vehicles by searching for the clustering which maximizes the likelihood of the observations.

References

[1] W. Feller, "An Introduction to Probability Theory and Its Applications," John Wiley, 1968.
[2] J. Pearl, "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference," Morgan Kaufman, 1988.
[3]. S. Kullback, "Information Theory and Statistics," Dover Publications, Inc. 1968.

Using Hierarchical Classification to Exploit Context in Pattern Classification for Information Fusion

A. Bailey and C. J. Harris
ISIS Research Group, Department of Electronics and Computer Science
University of Southampton,
Southampton, SO17 1BJ, UK.
e-mail: ab96r@ecs.soton.ac.uk

Abstract *In data fusion applications it is important that only the minimum set of relevant features are combined at any one stage in the fusion process. A hierarchical classification methodology is described which handles features at different levels of abstraction to produce a more robust and interpretable classifier. This is achieved by dividing the classes into contextual subgroups, which are further divided to produce a tree structure defining relationships between classes.*

A novel approach is proposed for the class structure design which is formulated as a constrained search in the structure space. This can be performed via a forward search algorithm driven by a cost function dependent on the performance of the class structure and constraints on the required solution.

Keywords: Statistical Pattern Classification, Context, Hierarchical

1 Introduction

In data fusion applications it is important that only the minimum set of relevant features are combined at any one stage in the fusion process. A hierarchical classification methodology is described which handles features at different levels of abstraction to produce a more robust and interpretable classifier. Classification is important in many areas of data fusion such as automatic target recognition, situation assessment and ballistic missile defence [1].

In pattern classification tasks of many classes in a high-dimensional input space, e.g. document classification, remote sensing and automatic target recognition, traditional flat classifiers tend to suffer from the curse of dimensionality [2]. Even after feature selection, a large number of inputs is still needed to discriminate the large number of potential classes. A hierarchical classifier overcomes this problem by dividing the classes into contextual subgroups, which are then further divided to produce a tree structure defining relationships between classes.

It can be shown that by using arbitrary probabilistic classifiers to discriminate between subgroups at each node of the tree, posterior probabilities can be output equivalent to a standard flat classifier but the feature space for each classifier is significantly reduced [3]. In this case each classifier node can be more robustly estimated as they perform simpler discrimination tasks in lower dimensional spaces.

Previous methods have used various clustering schemes to build the most appropriate class structure in terms of interpretation and accuracy of classification [4]. Once the class structure has been elicited, standard feature selection and parameter estimation techniques can be used to specify each classifier node.

Rose *et al.* [5] present a constrained hierarchical clustering method using simulated annealing. A small number of coarse clusters are identified at high temperature. As the temperature is lowered coarse clusters split into more detailed clusters. This leads to a course to fine hierarchy of clusters extracted during the an-

nealing schedule. Kim and Landgrebe [2] define a hierarchical classifier and a hybrid hierarchy design process combining both agglomerative and divisive clustering techniques.

A novel approach to identifying the class hierarchy by discrete optimization search is proposed. Constraints on the structure of the hierarchies in the search space are defined.

2 Context

Dealing with situations in their proper context is an everyday notion that allows people to select the right information in completing complex tasks. Researchers have previously identified the use of context as an advantage in problem solving.

Toussaint [7] suggests that to solve a problem instead of increasing the depth of analysis, one should widen the field of context in which the problem is viewed.

Whereas Antony [8] refers to context in tactical data fusion applications as including 'current friendly force disposition, existing weather conditions, natural domain features (terrain/elevation, surface materials, vegetation, rivers, drainage regions), and cultural features (roads, airfields, mobility barriers)'.

Context may be used in different ways to aid a classifier system. The above two references refer to widening the field of context by searching for extra information that may aid discrimination between classes. In terms of classification, this can be thought of as an intelligent method to generate more features for a classifier, given knowledge of the problem domain.

An alternative standpoint is that one can use the context of a situation to refine and abstract only the relevant information in the context of the classification problem at hand. This can be thought of as decomposing a problem and focusing attention only on those features of importance.

This paper is primarily concerned with the second view on context and a classification scheme is proposed that will fully exploit this notionl.

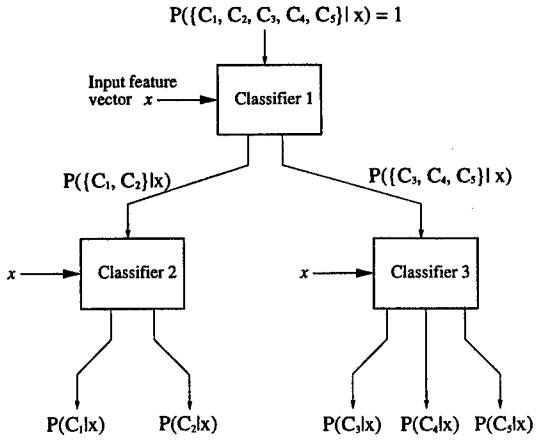


Figure 1: Hierarchical classifier structure

3 Hierarchical classifier

The framework described in this section is based on Schurmann [3]. Let Ω^{all} be the set of all the possible classes. This can be split, initially arbitrarily, into a number of subsets. The top level classifier provides posterior probabilities for each subset given the input vector, x. The second level classifiers then work on the individual subsets, breaking them down into smaller subsets, and incorporating the posteriors from the level above. This is performed in a tree-structured hierarchy until a posterior has been calculated for each individual class. This is shown in figure 1.

To formalise this hierarchical process consider a single classifier node as depicted in figure 2. The classifier need only be concerned with the subset of classes that it has been designated from the level above, this is called the input set, Ω^{in}. The input set is split into S output subsets, $\Omega_1^{\text{out}}, \Omega_2^{\text{out}}, ..., \Omega_S^{\text{out}}$. Each output subset is unique and all elements in the input subset are assigned to only one output subset. Formally,

$$\Omega_i^{\text{out}} \subset \Omega^{\text{in}} \; \forall i \text{ and } \Omega_{\text{i}}^{\text{out}} \cap \Omega_{\text{j}}^{\text{out}} = \emptyset \; \forall \text{i} \neq \text{j}, \quad (1)$$

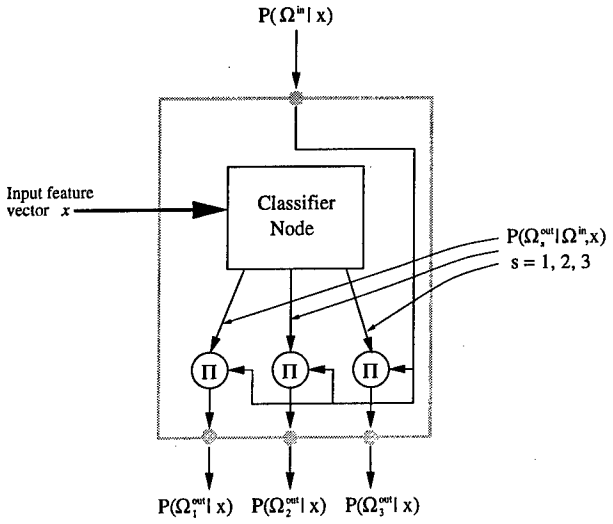


Figure 2: Individual classifier node

$$\bigcup_{s=1}^{S} \Omega_s^{\text{out}} = \Omega^{\text{in}}. \tag{2}$$

The goal of the complete classification process is to calculate posterior probabilities for each class given the input vector. However, each classifier in the hierarchical structure can only generate posteriors that are valid given the input set Ω^{in}:

$$P(\Omega_s^{\text{out}}|\Omega^{\text{in}}, x) \; s = 1, ..., S. \tag{3}$$

Here we can use the law of conditional probability to determine the relation between the input and output sets of the individual classifiers:

$$P(\Omega_s^{\text{out}}|\Omega^{\text{in}}, x) = \frac{P(\Omega_s^{\text{out}}, \Omega^{\text{in}}, x)}{P(\Omega^{\text{in}}, x)} \tag{4}$$

$$\begin{aligned} P(\Omega_s^{\text{out}}|\Omega^{\text{in}}, x) &= \frac{P(\Omega_s^{\text{out}}, x)}{P(\Omega^{\text{in}}, x)} \\ &= \frac{P(\Omega_s^{\text{out}}|x)p(x)}{P(\Omega^{\text{in}}|x)p(x)} \\ &= \frac{P(\Omega_s^{\text{out}}|x)}{P(\Omega^{\text{in}}|x)} \end{aligned} \tag{5}$$

Using this relationship it is easy to demonstrate that to calculate the required class posteriors, all is required is to multiply the posteriors given at each level while following the route down the tree to the required class.

3.1 Comments

No assumptions have been made on the type of classifiers used, save that they return a posterior probability for each output. No assumptions have been made at all concerning the structure of the class subsets and the elements of the input vector, x. Schurmann [3] states that, given optimal discrimination at each node, optimal discrimination will be achieved for each final class regardless of the choice of tree structure. Evidently optimal discrimination is not always possible and a sensible choice of the class subsets that defines the tree structure will maximise discrimination as each node and therefore maximise global discrimination.

3.2 Difference from Decision Tree Classifiers

The principle differences between hierarchical classifiers and decision tree classifier (such as Quinlan's C4.5 [9]) are the use of soft decision making at each node, allowing posteriors to be output for all classes, the use of arbitrary classifiers at each node, and the manipulation of the set of class hierarchies.

Although refence is made to the data in the input space via the distance metrics in the objective function, the space that is being searched in the construction of the hierarchy is not the set of possible splits in input space, but the combination of class subsets in the set of all possible class hierarchies.

A given class is always considered complete with all points belonging to that class irrespective of any class overlap in input space. Under the assumption of of normally distributed classes the collection of points in al class may be represented by their mean and covariance.

3.3 Existing methods for class hierarchy design

Standard methods exist for generating class hierarchies. These tend to be motivated from hierarchical clustering techniques which are unsupervised methods. Two distinct approaches are agglomerative and divisive methods, depending on whether the hierarchy is grown from the top down, or from the bottom up.

3.3.1 Agglomerative clustering

When designing a class structure from the bottom up, the goal is to group the classes together according to a similarity measure. The two most similiar classes are joined to form an intermediate class. Each new intermediate class replaces its members and the process is repeated until the last two classes are joined.

The similiary measures used can be chosen depending on the assumptions made on the class distributions. Suitable distance measures are the Mahalanobis distance, or the Battacharyya distance, both of which assume normally distributed classes with arbitrary covariance matrices.

3.3.2 Divisive clustering

The top-down approach to structure design works in an opposite fashion by considering the dataset as a whole initially and then splitting it into smaller groups according to a performance criteria. Each group can then be considered for splitting itself and a hierarchy is built in a recursive manner.

The technique of splitting data into homogeneous groups is well researched in the literature on clustering.

4 Contextual features

In most pattern recognition problems the dataset is defined by a set of points in an n-dimensional space. Each point represents an entity to be classified and each dimension represents a measureable feature of that object. If a feature is not present for all points in the dataset then methods are available to fill in these points with estimates of the most likely missing values [10]. Most classification algorithms require that all features are present for every point before training can take place. However, for the set of decision tree classifiers, including C4.5-type algorithms and the hierarchical classifier described here, certain features may not have to be present for all classes. A feature may be used to discriminate between only a subset of the available classes, and in this case its value for points outside that subset are not important, and even need not be defined.

In fact, the presence of such features, that may occur naturally in many real-world datasets such as in target recognition or medical diagnosis applications, can aid the design process of a hierarchical classifier. Regardless of any discriminative information that may be contained within such features their presence or absence can suggest a suitable class hierarchy. For example, studying the dataset in figure 3, one can determine the class hierarchy given alongside.

Although it may be possible to generate a class hierarchy simply from the pattern of such contextual features, it is again a difficult search problem. It is much easier to verify that a given class hierarchy matches a feature pattern by checking for consistency of contextual features across class subsets. A feature may only be entirely present or absent for a given class subset. If this constraint is broken then the hierarchy can be assumed to be an unlikely solution for that dataset. This constraint can easily be incorporated into the search procedure proposed below.

5 Solution requirements

It is proposed that the class hierarchy can be found by a search in solution space (i.e. the set of all possible class hierarchies). The search strategy must allow for four specific requirements on the eventual solution sought. That is:

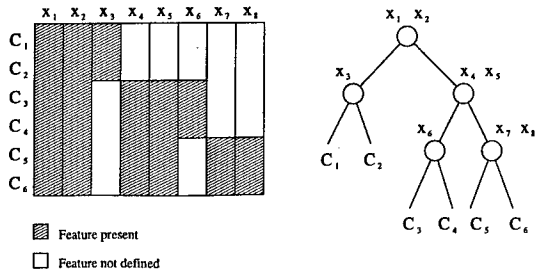


Figure 3: Dataset with contextual features and matching hierarchy

- Accuracy - Evidently the final classifier should produce accurate classification results.
- Degrees of freedom - The number of parameters for the final model should be controlled to prevent overfitting and the curse of dimensionality.
- Smoothness - The models must adhere to predefined constraints on the structure to prevent unbalanced trees.
- Prior Knowledge - If a prior model is given, the information represented in that prior must be respected.

5.1 Accuracy

To measure the accuracy of a classifier the most direct measure is the misclassification rate, as this truly reflects the desired performance. However this can be costly to compute, as it involves the complete training of each classifier node in the hierarchy. A more efficient and reasonable approximation is a sum of Mahalanobis distances between class subsets:

$$J = \sum_{n \in S} (\mu_{l_n} - \mu_{r_n})^T \Sigma^{-1} (\mu_{l_n} - \mu_{r_n}) \qquad (6)$$

where S is the set of all nodes in the hierarchy, μ_{l_n} and μ_{r_n} are the means of the points in the left and right class subsets of node n respectively and Σ is the covariance matrix for the points in both class subsets.

This distance measure reflects the separation between the left and right class subgroups chosen at each level. This value is to be minimised, since the classes at the leaves of the tree are required to be close together so that they may sensibly be contained by the class superset at the parent node. At levels higher in the tree the requirement of proximity becomes less important and this is satisfied naturally since there will always be more nodes at each descending level of the tree. The summation therefore gives greater significance to the leaf classes due to their number.

5.2 Degrees of freedom

The dangers of overfitting and the curse of dimensionality have been known for some time now [10]. Controlling the effective number of parameters in a model is important, especially in terms of small datasets where the parameters need to be specified by a few data points in a manner that is statistically significant.

The number of nodes in the class hierarchy is fixed, due to the binary nature of the tree and the requirement that all classes should be present. The complexity of the model can be controlled by using the optimum minimal feature set at each classifier node.

However, distance measures such as in equation (6) can only be used to compare features sets of the same dimensionality. This presents a problem since performing feature selection for each state to be evaluated in solution space will result in models of different dimensionality across nodes, rendering the summation meaningless. Again a solution to this would be to use the final classification accuracy as a performance measure but this is too computationally expensive. As an initial investigation, the search is performed on models containing all features, allowing equation (6) to be used.

Feature selection and training can be performed on the model given by the final class hierarchy.

5.3 Smoothness

For a class hierarchy to be meaningful and interpretable, it is likely that is should be well balanced (i.e. no particular branch should be significantly deeper than the others). A constraint on the maximum allowable depth of a single branch can be imposed on the required solution.

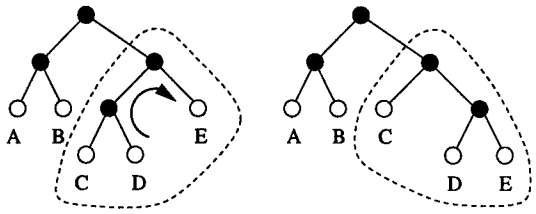


Figure 4: Example of branch shift operator

5.4 Prior Knowledge

The data is not usually the only available source of information on the desired model. Prior knowledge should be used wherever possible to aid the construction of a good model. If a class hierarchy is suggested through domain knowledge then it may be desirable to find a solution that does not differ significantly from the given hierarchy.

Tree comparison algorithms are available, but can be computationally expensive. In some cases the distance between two trees is described by the number of transformations required to transform one tree to the other. This distance is effectively the search depth if the search is initialised with a prior hierarchy. This can be easily and efficiently incorporated as a constraint on the search.

6 Hierarchy design by combinatorial optimization

The search strategies presented in this section strive to find a class hierarchy that gives a mimimum value of the objective function. It is expected that there will be many local minima due to the discrete nature of the search space (the set of possible class hierarchies). There are several combinatorial optimisation techniques that have become standard in the literature, some of which are designed to overcome such local minima. As ever, there is no one algorithm which can guarantee the global minimum is found, but techniques can be used that increase the chance of finding a good result.

The discrete optimisation problem can be defined as finding a solution from the set of all possible states such that minimises a given objective function within the given constraints. The constraints may be incorporated as hard constraints on the states allowed by the search procedure, the search rejects any states that do not adhere to the constrainst, or they may be incorporated as soft constraints by introducing extra terms in the objective function. The technique in this paper uses hard constraints, effectively reducing the search space. Soft constraints require free parameters to be estimated as coefficients for each term in the objective function which can be costly for large search spaces. The set of all possible states (state space) may be viewed as a graph, given operators that transform one state to another. Such graphs tend to grow exponentially with the size of a problem and optimal search techniques are NP-compete, that is the solution time increases exponentially with problem size for all algorithms. However, heuristic algorithms exist that can find suboptimal solutions in polynomial time. Examples are directed depth-first search (DFS), cost-bounded DFS (IDA*), depth-first branch-and-bound (DFBB) and best-first search (BFS). Many of these algorithms may be implemented on a parallel processor architecture [11].

A single operator can be defined on class hierarchies to generate a search tree. From a given node, a branch of a child node may be shifted to be the sibling of the opposite child node. This is illustrated in figure 4. Any valid hierarchy may generated from any other valid hierarchy using combinations of this operator. For simplicity only binary hierarchies are used.

7 Results

Initial results have shown that anything other than a directed depth-first search without backtracking produces a search tree that is far too large to evaluate. (Simply backtracking to just the next best branch produces a search space of 7219 nodes for an 8-depth search on a 12-class problem, compared to 504 nodes for a 30-depth search without backtracking. The deeper search consistently found a better solution.)

Depth-bounded directed depth-first searches were performed on a selection of simulated datasets with the number of classes ranging from 8 to 32 and the number of features ranging from 7 to 93. The searches were initialised using the solution from an agglomerative design procedure and all but one search showed an improvement on this initial solution, the improvements becoming more significant as the number of classes increased. Figure 5 shows the increase in performance as the number of classes increase.

Figure 6 shows for a dataset of 32 classes the improvement in the objective function as the search progresses. Figures 7 and 8 show the class hierarchies generated by the agglomerative clustering procedure and the improved hierarchy given by the search. The hierarchy in Figure 8 displays a more balanced and representative tree.

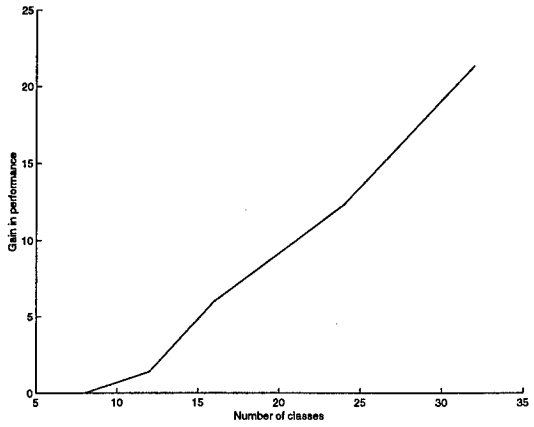


Figure 5: Gain in performance (sum of distances) against number of classes

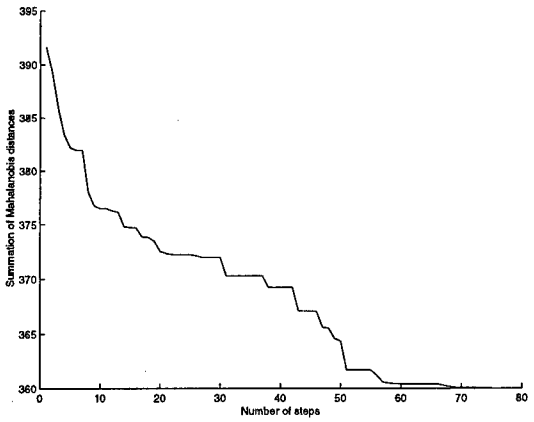


Figure 6: Search criterion against search steps

8 Conclusions

A novel approach to defining the class hierarchy for a hierarchical classifier using a search in solution space has been proposed and shown to improve upon results given by an existing method.

Contextual features have been defined and it has been shown how they may assist a search as described above by producing a constraint that can reduce the size of the search space.

9 Acknowledgements

The author would like to acknowledge IBM UK for their support of this work.

References

[1] C. J. Harris, A. Bailey, and T. J. Dodd. Multi-sensor data fusion in defence and aerospace. *The Aeronautical Journal*, 102(1015):229–244, 1998.

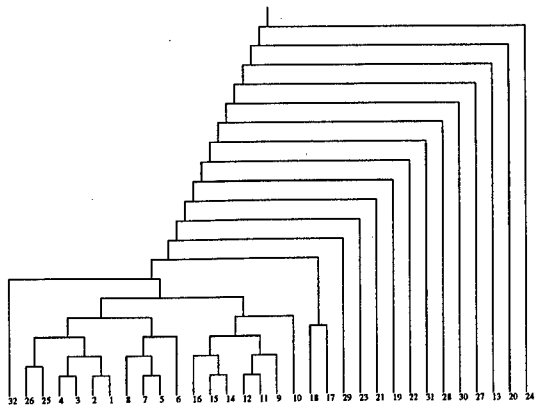


Figure 7: Hierarchy resulting from agglomerative design (depth of hierarchy = 20)

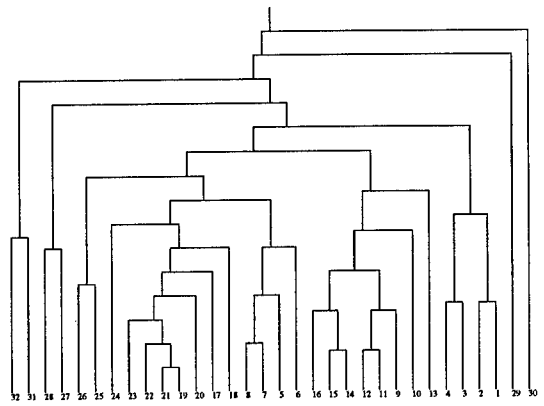


Figure 8: Hierarchy after search (depth of hierarchy = 16)

[2] B. Kim and D. A. Landgrebe. Hierarchical classifier design in high-dimensional numerous class cases. *IEEE Transactions on Geoscience and Remote Sensing*, 29(4):518–528, 1991.

[3] J. Schurmann. *Pattern Classification: A unified view of statistical and neural networks.* Wiley-Interscience, 1996.

[4] M. Sahami. Using machine learning to improve information access. Phd thesis, Stanford University, 1998.

[5] K. Rose, E. Gurewitz, and G. C. Fox. Constrained clustering as an optimization method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(8):785–794, 1993.

[6] S. R. Safavian and D. Landgrebe. A survey of decision tree classifier methodology. 21(3):660–674, 1991.

[7] G. C. Toussaint. The use of context in pattern recognition. *Pattern Recongnition*, 10:189–204, 1978.

[8] R. T. Antony. Database support to data fusion automation. *Proceedings of the IEEE*, 85(1):39 – 53, 1997.

[9] J. R. Quinlan. *C4.5 : Programs for machine learning.* Morgan Kaufmann, 1993.

[10] C. M. Bishop. *Neural Networks for Pattern Recognition.* Clarendon Press, Oxford, 1995.

[11] A. Grama and V. Kumar. State of the art in parallel search techniques for discrete optimization problems. *IEEE Transactions on Knowledge and Data Engineering*, 11(1):28–35, 1999.

Fusion of Information Multisensors Heterogeneous Using an Entropy Criterion

Bienvenu FASSINUT-MOMBOT, Mourad ZRIBI, Jean-Bernard CHOQUEL

Universite du Littoral Côte d'Opale (ULCO)
Laboratoire d'Analyse des Système du Littoral (LASL-EA 2600)
195, Rue du Pasteur Martin-Luther-King BP 649
62228 CALAIS CEDEX FRANCE
Bienvenu.Fassinut-Mombot@lasl-gw.univ-littoral.fr, Mourad.Zribi.@lasl-gw.univ-littoral.fr
Jean-Bernard.Choquel@lasl-gw.univ-littoral.fr

Abstract – *In this paper, we consider optimal information fusion (multisensors heterogeneous) problem from a Structural Analysis point of view. In this approach, Information Theory is used to elaborate an optimal fusion rule, i.e., to measure the performance of the data fusion structure given the sensors. A number of sensors transmit their observations about a phenomenon to a fusion center. But, in employing a criterion from Shannon's (mutual) Information, we assume that the probability distributions are known. These assumptions are not always true. In this case, parametric approach and non-parametric are presented to solve this problem. We present a design of data fusion algorithm.*

Keywords : Structural Analysis, Shannon mutual Information, data fusion heterogeneous, Maximum Likelihood and Maximum Entropy distributions, Fourier Series methods.

1 Introduction

In heterogeneous perception, sensors must recognize and identify rapidly, i.e., in real time, a set of agents acting on the activation of a process. Some sensors provide informations on abrupt evolution (real or erratic) and lead to an important modification of the model. Each sensor provides information of different nature that it is necessary to merge with intend to elaborate a decision. Data provided by sensors constitute the initial data table on which, we effect a Structural Analysis. Tools of the Structural Analysis [1] [2] [3] are mainly oriented towards :

- the identification of possible subsystems (structuring problems) ;
- the updating of redundancies, i.e., the internal organization of the system (explicative problem).

Structural Modelling of complex systems using the concepts of Information Theory presents the advantage to apply to different variables (quantitative, qualitative, set of structured modalities or non-structured,...) as well as to different relationship (linear, non-linear, fuzzy, ...).

This "system approach" is used, for example, to supervise complex industrial installations. In this case, a great number of sensors transmit their observations to a data fusion center where they are appropriately combined to obtain a global decision.

This paper is organized as follows. Section 2 contains the statement of the problem and the necessary notational definitions. In section 3, we consider the probability distributions estimations problem. In section 4, we study an optimum data fusion structure given the sensors. Finally, in section 5, we conclude and discuss of the major points of this paper.

2 Statement of the problem

We consider the data fusion problem with N sensors as shown Fig.1. The i^{th} sensor transmit a data string of length M to a central processor, denoted by $X_i^M \triangleq \{X_{ij}\}_{j=1}^M$, $i = 1, \ldots, N$. We use the following notational definitions :

Let $\Sigma = \{X, Y\}$ be the set of relevant classification and description variables of the system. Where :

- $X = \{X_1, X_1, \ldots, X_N\}$ is the global observations given by the sensors , and
- $Y = \{Y_1, Y_1, \ldots, Y_N\}$ is the classification function that associates a class at the phenomenon observed. See TAB. 1.

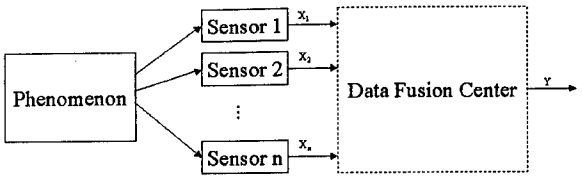


FIG. 1: Design data fusion

We assume that the sensor observations (raw data) are statistically independent and identically distributed (i.i.d). We further assume that the sensors at high rules and/or are geographically close to each other. The problem of combining the various sensor observations, that is, the data fusion problem involves simultaneous optimization the aggregation of the sensors (structuration problem) and fusion rule (explicative problem). It can be see as a classification problem subject to multiple hypothesis.

Thus goal of data fusion problem is to find a design of the form $Y = f(X)$, i.e, to process the raw data X in such manner that Y contains as much information as possible about the hypothesis. As a measure of information transfer, we shall use the Shannon's (mutual) information

$$I(Y;X) \triangleq H(Y) - H(Y/X) \qquad (1)$$

Where $H(Y)$ measure the uncertainty about Y before X is observed (entropy of Y) whereas $H(Y/X)$ the uncertainty after observations, i.e., the conditional entropy.

TAB. 1: Initial data Table

Ω Σ	X_1	X_2	X_i	...	X_M	Y
ω_1						
ω_2						
ω_j	...	...	$X_i(\omega_j)$	...		$Y_i(\omega_j)$
ω_K						

The optimal data fusion is obtained by maximizing the mutual information $I(Y;X)$. Maximizing the mutual information is equivalent to simultanemously maximizing $H(Y)$ and minimizing $H(Y/X)$. We choose the conditional entropy to be the criterion of the design performance. This quantity is given by

$$\begin{aligned} H(Y/X) \triangleq & H(X,Y) - H(X) \\ = & -\sum_{i=1}^{M} Pr(X = X_i) \cdot H(Y/X = X_i) \\ = & -\sum_{i=1}^{M} Pr(X_i) \\ & \times \left[\sum_{j=1}^{K} Pr(Y_j/X_i) \log Pr(Y_j/X_i) \right] \end{aligned} \qquad (2)$$

with $\log(x) = \frac{\ln(x)}{\ln(2)}$.
Before we proceed with solution to the problem, in the next section, we will estimate the probability densities.

3 Probabilities learning

The formula of the conditional entropy requires the computation of two type of probabilities : prior probabilities and conditional probabilities. These problems are investigated in the following subsections.

3.1 Prior probabilities

Maximum Entropy-Likelihood distributions

Shannon (1948) defines entropy (disorder or uncertainty) of density function $\pi(x)$ (with respect to Lebesgue measure) by

$$H(X) \triangleq -\int_{-\infty}^{+\infty} \pi(x) \log \pi(x)\, dx \qquad (3)$$

Maximizing H subject to various side conditions is well-known in the literature as a method for deriving the forms of minimal information prior distributions; e. g Jaynes (1968) and Zellner (1977). This problem, in its general form is the following:

$$(\mathrm{P}) \quad \begin{cases} \text{maximize} \quad H(X) = -\int \pi(x) \log \pi(x)\, dx \\ \text{subject to} \\ \quad E[\phi_k(x)] = \int \pi(x) \log \pi(x)\, dx = d_k \\ \qquad\qquad\qquad\qquad k = 0, \ldots, K \end{cases}$$

where $d_0 = 1, \phi_0(x) = 1$ and $\phi_k(x), k = 0, \ldots, K$ are k known arbitrary functions, and $d_k, k = 0, \ldots, K$ are the given expectation data. The classical solution of this problem is given, writing the Lagrangien and using variational calculus, by

$$\pi_{ME}(x) = \frac{1}{Z(\lambda)} \cdot \exp\left[-\sum_{k=0}^{K} \lambda_k \phi_k(x)\right] \quad (4)$$

where $Z(\lambda)$ is a normalization function

$$Z(\lambda) = \int \exp\left[-\sum_{k=0}^{K} \lambda_k \phi_k(x)\right] dx \quad (5)$$

and $\lambda = (\lambda_0, \ldots, \lambda_K)^T$ is the vector of Lagrangien parameters. They can be determined exactly by solving the $(K+1)$ nonlinear constraint equations given by (P) for $\pi_{ME}(x)$

$$G_k(\lambda) = \int \phi_k(x) \cdot \exp\left[-\sum_{l=0}^{K} \lambda_l \phi_l(x)\, dx\right] \quad (6)$$
$$k = 0, \ldots, K$$

These equations are equivalent of the problem of calculating the maximum likelihood (ML) estimates of the parameters of a family of regular exponential densities. In this ML problem, the parameters are given by

$$\frac{\partial \log Z(\lambda)}{\partial \lambda_k} = d_k \quad k = 0, \ldots, K \quad (7)$$

where

$$d_k = \frac{1}{M} \sum_{i=1}^{M} \phi_k(X_i) \quad (8)$$

is empirical average. In fact, in practice , we do not have any values of d_k but a random samples $X_1, \ldots, X_M$ independently drawn from a distribution $P = \{\pi(X)/E_\pi[\phi_k(X)] = d_k; k = 0, \ldots, K\}$.

So, d_k, $k = 0, \ldots, K$ are determined by their empirical estimates and we use the Maximum Entropy principle (ME) to find unique solution of the problem of density estimation subject to empirical constraints [6]. We have shown the equivalency between the ME distributions subject to empirical constraints and ML estimates of the Lagrange parameters in exponential families. In general, we use the standard Newton-Raphson method to solve the nonlinear equations (6) or (7). This method consists of expanding G_k in Taylor's series around the trivial values λ^0 of the λ's, drop quadratic and higher order terms, and solve resulting linear system iteratively [5]

$$\begin{aligned} d_k &= G_k(\lambda) \quad \text{for} \quad k = 0, \ldots K \\ &\cong G_k(\lambda^0) + (\lambda - \lambda^0)^T \Big[G_k(\lambda)\Big]_{\lambda=\lambda^0} \end{aligned} \quad (9)$$

Noting the vectors δ et ν by

$$\begin{aligned} \delta &= \lambda - \lambda^0 \\ \nu &= \left[G_k(\lambda^0) - d_0, \ldots, G_K(\lambda^0) - d_K\right]^T \end{aligned}$$

and the matrix $\mathbf{G}$ by

$$\mathbf{G} = \left(g_{kl}\right) = \left(\frac{\partial G_k(\lambda)}{\partial \lambda_l}\right)_{\substack{\lambda=\lambda^0 \\ k,l = 0, \ldots K}} \quad (10)$$

then equation (9) becomes

$$\mathbf{G} \cdot \delta = \nu \quad (11)$$

This system is solved for δ from which we drive $\lambda = \lambda^0 + \delta$, which becomes the new vector initial vector λ^0 and the iterations continue until δ becomes appropriately small.

Fourier's components case

In our problem, we are interested the case where the data are the Fourier components (complex data) of the distribution probability function $\pi_{ME}(x)$. In this case, the Lagrange parameters are solution of Hermitian Toeplitz system (11)

$$\mathbf{G} \cdot \delta = \nu$$

where $\mathbf{G} \triangleq \left(G_{l-k}\right)_{k,l}^{K} \in \mathbb{C}^{K \times K}$ is positive define and the element of the matrix $\mathbf{G}$ is given by

$$G_k(\lambda) = \int \pi_{ME}(x) \cdot \exp(-jk\omega_0 x)\, dx \quad (12)$$
$$k = 0, \ldots, K$$

This system can be solved by the conjugate gradient method (C-G method) [13].

Proposition 1 *Suppose $X_1, X_2, \ldots, X_N$ are i.i.d with exponential families densities $\pi(x) = \exp[-\sum_{k=0}^{K} \lambda_k \phi_k(x)]$, then, the problem of ME-ML distributions estimate has equivalent formulation*

$$(Q) \quad \begin{cases} \textit{maximize} & \sum_{k=0}^{K} \lambda_k d_k, \\ \textit{subject to} \begin{cases} \mathbf{G} \cdot \delta &= \nu \quad \textit{with} \\ \delta &= \lambda - \lambda^0 \end{cases} \end{cases}$$

The procedure has been implemented in MATLAB and C language, and they have been tested on a Sun SPARCstation and ULTRA Enterprise 450.

3.2 Conditional probabilities

In this subsection, we shown that the estimation of the conditional probabilities by a parametric approach is applicable only under certain conditions (for example form of law). We will show that a non-parametric approach is to better adapt.

Parametric approach

We propose initially to carry out series of test statistics on the sensor observations, in order to put for the assumptions on the law subjacents [8]. Next, we estimate these laws by a parametric approach.

Let then $X_1, \ldots, X_N$ be a sample of N i.i.d random variables from a distribution having unknown mean μ, and unknown variance σ^2 (aggregation of the sensor observations). We also assume N is large.

1. compute the quantities:
   Skewness :
   $$\gamma_1 = \frac{d_3}{d_2^6} = \frac{d_3}{\sigma^3} \tag{13}$$
   Kurtosis :
   $$\gamma_2 = \frac{d_4}{d_2^2} = \frac{d_4}{\sigma^4} \tag{14}$$
2. test:
   $$\text{if} \begin{cases} \gamma_1 = 0, & \text{symmetric laws} \\ & \text{otherwise asymmetric laws} \\ \gamma_2 \simeq 3, & \text{Gaussian hypothesis.} \end{cases}$$
   We assume $f_i \sim \mathcal{N}(\mu, \sigma)$. Hence :
3. test statistics:
   We wish to test estimated law parameters. The test consist to test the variances initially and if they are not significantly different to then test the means by supposing that $\tilde{\sigma} = d_2^2 = \sigma$ [7].
   (a) Fisher's variances test:
   The test statistic is given by
   $$F = \frac{\sigma}{\tilde{\sigma}} \tag{15}$$
   where under suitable conditions, this ratio has a probability distribution known.
   (b) Student's means test:
   We suppose that $\tilde{\sigma} = \sigma$. In this case, the appropriate test statistic is
   $$T = \frac{d_1 - \tilde{\mu}}{\sigma \Big/ \sqrt{N}} \tag{16}$$
   with $d_1 = \mu$ and where under suitable conditions, this ratio has a probability distribution known.

In summary, if overall tests are rejeted (for example, instability of the estimates parameters) then, we will use a non-parametric approach.

Non-parametric approach

The estimation of the probability density function (pdf) based on the Fourier analysis method is suitable in this context. Then an estimator of the pdf based on independent samples $X_1, \ldots, X_N$ with density f is given by

$$\tilde{f}_{K_N}(x) = \sum_{k=0}^{K_N} \tilde{a}_{k,N} \cdot e_k(x) \tag{17}$$

where $\tilde{a}_{k,N} = \frac{1}{N} \sum_{i=0}^{N} e_k(X_i)$ and $\left\{ e_k(x) \right\}_{k \in \mathbb{N}}$ is an orthonormal basis of the Hilbert space $\mathrm{L}^2([a,b])$ and K_N is an integer dependent on N, called the *truncation point.*
Different theorems are provided for the convergence rate of K the terms of mean integrated square error (MISE) or mean square error (MSE) [9][12]. The optimal choice for the MISE criterion is $K_N \simeq N^{\frac{1}{p}}$ with $p \geqslant 2$.

Description of the non-parametric EM (Esti mation-Maximisation)

The form of the conditional pdf's does not need to be known is this approach, since we propose to define y_j as $y_j : a_{0,j} \longrightarrow a_{K_{N_j},j}$. So, we denote $\tilde{f}_{K_{N_j}}(x) = f(x/y_j)$. The suggested algorithm consists of three following steps:

1. Initialization step:
   The number of classes K is assumed to be known. The parameters of mixture can be initialized as following:
   $$\tilde{\pi}_j^0 = \pi_{ME} \quad \text{prior probability} \tag{18}$$
   $$\tilde{a}_{k,j} = \frac{1}{\tilde{N}_j^0} \sum_{l=0}^{\tilde{N}_j^0} e_k(X_l) \tag{19}$$
   $$K_{\tilde{N}_j^0} = int\left[(\tilde{N}_j^0)^{\frac{1}{p}} \right] \tag{20}$$
   where $int[\cdot]$ is the largest integer less than the real number x and $\tilde{N}_j^0$ is the total number of observations in the class j.
2. Expectation step: In this step, we estimate the a posteriori probability $\tilde{\pi}_j^n(x_i)$ for the observation X_i belonging to the class j at the

n^{th} iteration :

$$\tilde{\pi}_j^n(x_i) = \frac{\tilde{\pi}_j^n\{(x_i/\tilde{y}_j^n)}{\sum_{l=1}^{K} \tilde{\pi}_l^n\{(x_i/\tilde{y}_j^n)} \qquad (21)$$

3. Maximization step : The a posteriori probability $\tilde{\pi}_j^n(x_i)$ of each observation X_i is computed. So at $(n+1)^{th}$ iteration, we have

$$\tilde{\pi}_j^{n+1} = \frac{1}{N}\sum_{l=0}^{N} \tilde{\pi}_j^n(x_l) \qquad (22)$$

$$K_{\tilde{N}_j^{n+1}} = int\left[(\tilde{N}_j^{n+1})^{\frac{1}{p}}\right] \qquad (23)$$

where $\tilde{N}_j^{n+1} = N\tilde{\pi}_j^{n+1}$ and,

$$\tilde{a}_{k,j}^{n+1} = \frac{\sum_{l=1}^{N} e_k(x_l)\tilde{\pi}_j^n(x_l)}{\sum_{l=1}^{N} \tilde{\pi}_j^n(x_l)} \qquad (24)$$

for $k = 0, \ldots, K_{\tilde{N}_j^{n+1}}$.

We now proceed with the optimization of the data fusion algorithm.

4 Data fusion

In this section, we will consider data fusion algorithm, where each sensor transit its observations about the phenomenon in to the central processor where a global fusion is then effected.We wish to minimize $H(Y/S)$, i.e., to obtain the optimum fusion rule $Y = f(S)$, where S is a optimal partition of the system $\Sigma = (X, Y)$ (see TAB.2).

The problem consists then, on the one hand to built the sets S and $f(S)$, and one the other hand to evaluate the quality or efficiency of the rule is derived.
Consider the following properties :

Property 1 (sub-additivity) *Let S_1 and S_2 two sets of variables (S_1, $S_2 \in \mathcal{P}(\Sigma)$). Then*

$$H(S_1 \cup S_2) \leqslant H(S_1) + H(S_2)$$

with egality if and only if S_1 and S_2 are statistically independent.

Property 2 (monotonicity) *Shannon's (mutual) Information verifies the following monotonicity properties:*

1. *if $X : \Omega \rightarrowtail M_X$ and $Y : \Omega \rightarrowtail M_Y$ 2 variables of Σ defines on Ω, then*

$$X \leq Y \Longrightarrow H(X) \geqslant H(Y)$$

2. *if S_1, $S_2 \in \mathcal{P}(\Sigma)$, then*

$$S_1 \subseteq S_2 \Longrightarrow H(S_1) \leqslant H(S_2)$$

where $\mathcal{P}(\Sigma) \subseteq \mathbb{P}(\Omega)$ the set of the partition of set of the samples producing the data (obsevation space) and M_X and M_Y the modalities of X respectively Y.

Optimum fusion rule

Considering property 2 [4], the local optimum research will be much by levels (classes).
Let $\mathbb{P}_k(\Sigma) = \{S_1, S_2, \ldots, S_k\}$ be the partition of Σ into k classes obtained via aggregations, with $card(S_i) = k_i$ and

$$S_i = \bigvee_{i=1}^{k_i} \bigwedge_{j=1}^{M} X_i \qquad (25)$$

where $\bigvee$ is operator of disjonction and $\bigwedge$ conjonction. The composed rule is expressed under the form

$$X \in S_i \longrightarrow Y = f(S_i) \qquad (26)$$

We show without difficulty that

$$H(Y/S_k) \leqslant H(Y/S_{k-1}) \qquad (27)$$

Algorithm

We propose a successive approximation algorithm combining an aggregative and desaggregative structure. we obtain two possible operators :

$$AH \circ DR : \mathbb{P}_k(\Sigma) \rightarrowtail \mathbb{P}_k(\Sigma)$$
$$S_k \rightsquigarrow AH \circ DR(S_k) = AH(DR(S_k)) \qquad (28)$$

$$DR \circ AH : \mathbb{P}_k(\Sigma) \rightarrowtail \mathbb{P}_k(\Sigma)$$
$$S_k \rightsquigarrow DR \circ AH(S_k) = DR(AH(S_k)) \qquad (29)$$

for $2 \leqslant k \leqslant N-1$
where

$$DR : \mathbb{P}_k(\Sigma) \rightarrowtail \mathbb{P}_{k+1}(\Sigma)$$
$$S_k \rightsquigarrow DR(S_k) = S_{k+1} \backslash X_i) \qquad (30)$$

with

$$\min_{X_i \in S_{k+1}} (H(Y/\{S_{k+1} \backslash X_i\})) \qquad (31)$$

$$AH : \mathbb{P}_k(\Sigma) \rightarrowtail \mathbb{P}_{k-1}(\Sigma)$$
$$S_k \rightsquigarrow AH(S_k) = S_{k-1} \cup X_i \quad (32)$$

with

$$\min_{X_i \in \Sigma \setminus S_{k-1}} (H(Y/\{S_{k-1}, X_i\})) \quad (33)$$

Proposition 2 *$S_k \in \mathbb{P}_k(\Sigma)$ is optimum partition (stability) if and only if*

$$\overleftrightarrow{T}(AH \circ DR(S_k)) = \overleftrightarrow{T}(DR \circ AH(S_k)) \quad (34)$$
$$= \overleftrightarrow{T}(S_k) \quad (35)$$
$$\forall k, \quad 2 \leqslant k \leqslant N-1$$

This algorithm research for each of levels of the treillis $\mathbb{P}(\Sigma)$ Fig.2, a partition verifying the proposition 2.

Efficiency of rule

We define the following expressions

$$m(Y/X) = 1 - \frac{H(Y/X)}{H(Y)} \quad (36)$$

and,

$$q(Y/S) = \frac{H(Y) - H(Y/S)}{H(Y) - H(Y/X)} = \frac{m(Y/S)}{m(Y/X)} \quad (37)$$

where $m(Y/X)$ measure the modelisation $Y = f(X)$ and $q(Y/S)$ the quality of the optimum design $Y = \tilde{f}(S)$.

$\mathbf{P}_1(\Sigma)$ $\quad \{(X_1 \cup X_2 \cup \ldots \cup X_i \cup \ldots \cup X_N)\}$
$\mathbf{P}_2(\Sigma)$ $\quad \{X_1, (X_2 \cup \ldots \cup X_N)\} \ldots S_2 \ldots \{(X_1 \cup \ldots \cup X_{N-1}), X_N\}$
$\mathbf{P}_3(\Sigma)$ $\quad \{X_1, X_2, (X_3 \cup \ldots \cup X_N)\} \ldots S_3 \ldots \{(X_1 \cup \ldots \cup X_{N-2}), X_{N-1}, X_N\}$

$\mathbf{P}_j(\Sigma)$ $\quad \ldots S_j \ldots$
$\mathbf{P}_{j+1}(\Sigma)$ $\quad \ldots S_{j+1} \ldots$

$\mathbf{P}_{N-2}(\Sigma)$ $\quad \{(X_1 \cup X_2 \cup X_3), \ldots, X_N\} \ldots S_{N-2} \ldots \{X_1, \ldots, (X_{N-2} \cup X_{N-1} \cup X_N)\}$
$\mathbf{P}_{N-1}(\Sigma)$ $\quad \{(X_1 \cup X_2), \ldots, X_N\} \ldots S_{N-1} \ldots \{X_1, \ldots, (X_{N-1} \cup X_N)\}$
$\mathbf{P}_N(\Sigma)$ $\quad \{X_1, X_2, \ldots, X_i, \ldots, X_N\}$

FIG. 2: Le Treillis $\mathbb{P}(\Sigma)$

TAB. 2: Contingency Table

M_x \ M_y	Y_1	Y_2	Y_j	...	Y_K	
X_1			⋮			
X_2			⋮			
⋮			⋮			
X_i	...	...	p_{ij}	...	...	$p_{i.}$
⋮			⋮			
X_N			⋮			
			$p_{.j}$			

- $p_{i.} = \sum_j p_{ij}$, $p_{.j} = \sum_i p_{ij}$ are the marginal probabilities
- p_{ij} is the a posteriori probability

5 Conclusion and discution

In this paper, we considered the data fusion problem from Structural Analysis point of view. It has been established that the optimum fusion rule was obtained by researching optimal partition of the system minimizing an entropy criterion. We also shown that the estimation of the pdf based on the ME-ML distribution (prior probabilities) and nonparametric EM improve the optimal data fusion. Our objective is to apply these results to more complex systems, for example, to the industrial process supervision.

References

[1] Z. Abid. *Contribution à l'Analyse Structurale par la Théorie des Demi-treillis.* Thèse de doctorat 3ème cycle, Université Claude Bernard, Lyon 1, Octobre 1979. N^o d'ordre 892.

[2] J. Dufour. *Méthodes et méthodologies d'Analyse des Systèmes Complexes, Applications aux Procédés Industriels et aux Systèmes Macroéconomiques.* Thèse d'etat, Université Claude Bernard, Lyon 1, Mai 1979.

[3] M. Richetin. *Analyse Structurale de Systèmes complexes en vue d'une Commande Hiérarchisée.* Thèse de doctorat 3ème cycle, Université Paul Sabatier, Toulouse, 1982.

[4] V. M. Toro Cordoba. *Contribution à l'Analyse Structurale de Systèmes complexes à l'aide de l'Entropie et ses Généralisations.* Thèse de

doctorat 3ème cycle, USTLFA, Mars 1982. *N°* d'ordre 955.

[5] Ali Mohammad-Djafari. Estimation par Maximum de Vraisemblance des Lagrangiens des Distributions à Maximum d'Entropie. In *13ème Colloque sur le Traitement du Signal et des Images, GRETSI*, pages 9–12, 1991.

[6] Ali Mohammad-Djafari. A Matlab Program to calculate the Maximum Entropy Distributions. In C.R. Smith et al., editor, *Maximum Entropy and Bayesien Methods*, pages 221–233. 1992 Kluwer Academic Publishers, Printed in the Netherlands, 1991.

[7] Gilbert Saporta. *Probabilités, Analyse des données et Statistique.* Editions Technip, Paris, 1990.

[8] Stephan Chauvin. Evaluation des performances du modèle bayesien de fusion appliqué à l'imagerie satellitaire. In *15ème Colloque sur le Traitement du Signal et des Images, GRETSI*, 2, pages 949–952, 1995.

[9] Mourad Zribi. *Les Fonctions Spéciales et Les Représentations des Groupes pour La Reconnaissance de Formes: Application à l'Imagerie Médicale.* Thèse de doctorat 3ème cycle, Université de Rennes I, Octobre 1997.

[10] Mohammed Barboucha. *Modélisation Structurale des Systèmes complexes, Extraction et Validation des Règles d'un système Expert.* Thèse de doctorat d'état, USTLFA, Juin 1987.

[11] M. Barboucha and M. Staroswiecki. Building rules from contingency tables. In *5ème Journées Internationales Analyse des données et Informatique*, INRIA, Versailles - Paris, 29 Sept-2 Octobre 1987.

[12] Alan Julian Izenman. Recent Developments in Nonparametric Density Estimation. In *Journal of the American Statistical Association*, 86(413):205–224, March 1991.

[13] D. Potts and G. Steidl Preconditioners for ill-conditioned Toeplitz matrices. *Received January 1998. Revised Novembre 1998.*

[14] M. Staroswiecki and M. Barboucha. Apprentissage automatique par filtrage de tableaux de contingence. In Diday and Kodratoff, editors, *Induction Symbolique et Numérique*, Cepadues, Paris, 1991.

Towards the Fusion of Distributed Binary Decision Tree Classifiers

Qian Zhang
EECS Department, 121 Link Hall
Syracuse University, Syracuse, NY 13244, USA
Email: qizhang@syr.edu

Pramod K. Varshney
EECS Department, 121 Link Hall
Syracuse University, Syracuse, NY 13244, USA
Email: varshney@syr.edu

Abstract
Multiple sensor fusion and binary decision tree classifiers have been used to successfully solve many real world problems. These topics are usually studied separately. Fusion of binary decision tree classifiers in a multiple sensor environment has received very little attention. In this paper, we formulate the problem, investigate its scope, outline some issues associated with decision tree classifiers and multiple sensor fusion, and present some solution methodologies. The results are illustrated by means of an example.

Key words: sensor fusion, binary decision tree.

I. Introduction

Multiple sensor decision fusion is an important problem with many practical applications. This problem has been studied quite extensively and many significant results on this topic have been obtained [1-3]. In most studies, one-stage decision making procedures are employed at the sensor as well as at the fusion center. By a one-stage procedure, we mean that a single test is employed to distinguish between all the hypotheses. Such one-stage decision procedures may become too complex and impractical in situations where there are many variables and many object classes (hypotheses). These types of problems arise in areas such as automatic target recognition (ATR), team medical diagnosis and telemedicine, and large-scale surveillance systems. In these situations, it may be desirable to use multistage decision making procedures at the sensors and/or at the fusion center. There are several approaches for implementing multistage decision making structures. One popular approach is by means of a binary decision tree (BDT) [4]. The basic idea is to take a complex M-ary hypothesis testing problem, break it into several simpler binary hypothesis testing problems that are organized in a hierarchical tree structure to make decisions regarding the M hypotheses or object classes. Here, we investigate the use of BDTs in multisensor fusion problems. The benefits of using BDTs in multisensor decision fusion are multifold:

1. It is well known that the design of distributed detection systems that employ one-stage decision making for binary hypothesis testing problems is NP-hard. This design for M-ary hypothesis testing is even harder. BDTs make a sequence of binary decisions in a hierarchical manner that are easier to design, efficient and computationally simpler to implement with simpler decision regions. Thus, the use of a BDT may make the decision making procedure feasible for practical situations that have time or processing constraints. In addition, communication bandwidth efficiency may be achieved because transmission of binary decisions instead of M-ary decisions will be required.
2. Some of the available sensors may not be capable of distinguishing all the object classes in a pairwise manner. BDTs provide a framework for integrating the capability of all the sensors for multisensor decision fusion and for enhanced system performance.
3. Decision making via a BDT has an inherent flexibility to design tests at the internal nodes. This flexibility provides the ability to handle sensor defects, missing sensor observations/decisions, etc., thereby enhancing system robustness. Also, the flexibility may help in improving system performance.

The design of a BDT based multisensor fusion system involves the design of the BDT, design of the tests at the internal nodes of sensor BDTs, design of the fusion rule, and design of the system topology including communication structure of the multisensor fusion system. Goals of this design include enhanced overall system performance (recognition ability) and robustness, using least possible computation and communication. Some aspects of this problem have been addressed in the literature. Demirbas [5] proposed a non-parametric centralized object recognition scheme based on a BDT. Each sensor processes its data and extracts some features that are transmitted to the fusion center. Object recognition is carried out using a BDT generated from a training set. Dasarathy [6] concentrated on the architectural aspects of a system that fuses binary decisions into a single M-ary decision. The main goal was to design architectures that satisfy processing time constraints. Zhu et al. [7] also considered the problem of M-ary hypothesis testing using a parallel fusion topology where local detectors transmit binary decisions. They focussed on the design of decision rules and on system

performance. The main goal of this paper is to examine the overall problem of BDT based multisensor decision fusion problem, identify the issues that need to be addressed further and propose some solution methodologies. In Section II, we formulate the problem. In Section III, we consider a specific BDT based decision fusion system. In Section IV, we give an example to illustrate the results of Section III. In Section V, we make some concluding remarks.

II. Problem Overview

We focus our attention on the parallel architecture for BDT based multisensor fusion systems in this paper. The block diagram of a BDT based parallel decision fusion system is shown in Figure 1. The system consists of K sensors that observe a common phenomenon in parallel. The goal is to recognize a given unknown object that belongs to the set of objects $\{O_1, O_2, \ldots, O_M\}$. Let X_k be the observation vector of the kth sensor. Each sensor uses a BDT to make its decisions. Let T_k denote the BDT used by the kth sensor and U_k the decision made by the kth sensor. These decisions are transmitted to the fusion center that combines them to yield U_0, the global decision. The fusion rule $\Gamma_0(.)$ is a function that maps local decisions $u_1, \ldots, u_K$ into u_0 and makes a decision regarding the unknown object.

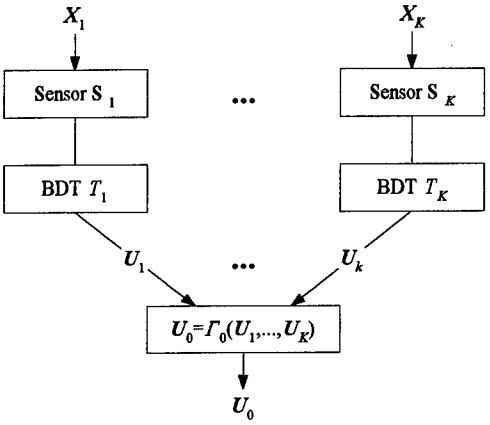


Figure 1: BDT based parallel decision fusion system

We assume that the reader is familiar with the notion of tree and associated terminology. For details, the reader may refer to [4]. A general BDT is shown in Figure 2. X denotes the feature vector. U denotes the decision made by the BDT at terminal nodes. Since each value u of U corresponds to a unique path from the root node to a terminal node, u can be encoded as the sequence of binary decisions made by all the nodes in the corresponding path. At node t, $\Phi(t)$ denotes the set of features used by the BDT, and $\Gamma(t)$ denotes the decision rule, which is a function that maps the space of features specified by $\Phi(t)$ into set $\{0,1\}$.

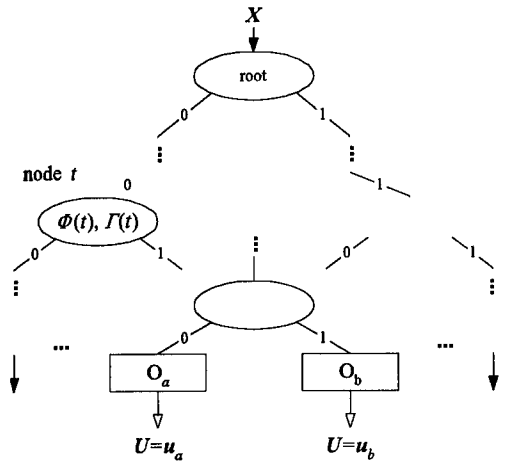


Figure 2: A general BDT

Using the basic system architecture for the decision fusion system shown in Figure 1, several design approaches and modes of operation can be envisaged. Here we categorize BDT based parallel decision fusion systems in four types:

1. Each sensor employs a BDT for decision making. These sensor BDTs are assumed available and are designed independently of each other. Sensor decisions are sent to the fusion center. The fusion center either uses a one-stage procedure or a BDT to determine the final decision. The main issue here is the design of an optimum fusion procedure. This problem is analogous to the design of an optimum fusion rule in distributed detection systems [9]. This problem is applicable to the fusion of BDT classifiers that may have been designed independently.
2. BDTs at the individual sensors are designed jointly by employing coupled cost functions. Decisions are available locally so that appropriate action can be taken at the sensor. Decision fusion is not employed to combine sensor decisions. This problem is analogous to the one considered by Tenney and Sandell [10] in a distributed detection context.
3. Decisions made by sensor BDTs are conveyed to the fusion center that combines them to yield the final decision. No other communication is allowed. Sensor BDTs and the fusion rule are designed. This system is a generalization of distributed detection systems where only one-stage decision procedures are allowed.
4. In this case, the system is the same as system 3 except that two-way communication between the sensors and the fusion center is allowed. The

sensors navigate through their BDTs under the supervision of the fusion center in a truly cooperative manner. This is a new system architecture that requires close coordination between system elements.

In the four types of systems described above, some or all of the following issues need to be addressed.

- Deign of BDTs for each sensor and/or the fusion center.
- Selection of sensor features for sensor BDT design and decision-making at internal nodes of sensor BDTs.
- Decision-making rules at internal nodes of sensor BDTs.
- Decision-making rules at the fusion center.
- Communication protocol used by the sensors and the fusion center for coordinated navigation of their BDTs.
- Evaluation of system performance - recognition ability, robustness etc.

The treatment of the above issues is guided by many factors, such as sensor characteristics, available processing resources, operating environment and the nature of the objects, etc. Due to the novelty of these problems, much effort is needed to develop a systematic solution to the overall BDT based multisensor fusion problem. In the following we make some observations on the differences between how the above issues are treated in the conventional sense and under the formulation of BDT based decision fusion. Conventional BDT design methodologies can be applied here, but with the goal of optimizing the overall system performance instead of just the performance of individual sensors. All BDTs should be designed in such a way that they collectively make the best classification of the unknown object. A single sensor BDT may not be optimal when used as a stand-alone classifier. Sensor features ought to be selected in such a way that objects are best separated across the sensor suite for enhanced performance of the overall system. The selected features may not be the best for individual sensors. The decision rules at internal nodes of sensor BDTs need to be designed such that best system performance is achieved. These rules may not be the best for individual BDTs.

III. Design of BDT Based Interactive Systems

In this section, we focus on BDT based decision fusion systems that involve two-way communication and where all system components are jointly designed. The quality of such a decision fusion system is given by the recognition rate of each object and the associated average number of steps in a recognition operation. The first quantity serves as a measure of effectiveness of the fusion system in carrying out detection/recognition, while the second quantity indicates the efficiency in terms of computational time/effort. We propose a method of designing BDTs, sensor rules and fusion rules. We assume that the sensor observations are conditionally independent given the object class. We also assume that each sensor observation is characterized by a probability distribution function given each object class and that it is known *a priori*. Finally, each sensor uses all the available features at each node of its BDT.

First, we propose a way to cooperatively use BDTs at the fusion center and the sensors. The fusion center BDT is used to carry out a sequential partition of the object space. At each internal node of its BDT, the fusion center tests one subset of objects against another subset of objects. It collects sensor local decisions and uses them to make a global decision on which subset to test further, i.e., it selects the path of the tree to follow. Based upon this global decision, it chooses an appropriate child node at which it tests two new subsets of objects. The fusion center repeats this procedure till it reaches a terminal node where exactly one object is left, and then declares this object to be the unknown object. We notice that at each node of the fusion center BDT, the local decision from a sensor reflects to what degree this sensor distinguishes the two subsets of objects that are under test. To optimally utilize the capability of this sensor, it is necessary that this sensor test the same two subsets of objects as the fusion center because adding/removing objects to these two subsets would make the local decision of this sensor less relevant to the recognition task at the fusion center. Based on this fact, we let every sensor BDT partition the object space in the same way as the fusion center BDT does. For this reason, all the BDTs may be considered identical with respect to the object space. However, since sensor local decisions may not always agree with the global decisions, some sensors may not choose the same path as the fusion center if the sensors use their own local decisions to select child nodes. If so, the local decisions from these sensors are less useful to the fusion center. To make sure that all system elements follow the same path of their BDTs, a mechanism of coordination is necessary. Such a mechanism is implemented via a simple two-way communication protocol. Suppose the sensors and the fusion center arrive at a node t, the sensors transmit their local decisions to the fusion center. Based upon these local decisions, the fusion center makes a global decision and sends it back to the sensors. Then both

the fusion center and the sensors use this global decision to choose the same child node. Although all sensor BDTs are the same in the object space, they are generally different in the feature space and sensor decision rule space. Furthermore, each sensor uses the past global decision to determine which features and what sensor decision rule to use at a node.

Now let us investigate how the structure of a BDT affects the recognition rate and the average number of steps. The ability of such a decision fusion system in achieving high recognition rates relies more on the actions at the higher level nodes of its BDTs. To see this, we note that rejection of an object at a node prohibits any further classification and ultimate recognition of this object. Therefore, an error that occurs at a higher level node is more costly than an error that occurs at a descendent node. From this viewpoint, at each node of the BDT one needs to distinguish objects that appear most dissimilar to the sensors. For this purpose, one would construct a BDT that tests and distinguishes the pair of most different objects at each node. However, the total number of internal nodes of such a BDT grows astronomically with the total number of objects because only two objects are distinguished at each internal node. Thus, the average number of steps may become prohibitively large. To alleviate this burden, one may choose to test more than two objects at each node, but this consequently decreases the dissimilarity between the objects being tested and hence the recognition rates. In the extreme case, one may choose to test all the remaining objects at a single node. This will use the least average number of steps but inevitably increase the probability of error. Therefore, we need to make a compromise between the dissimilarity of objects being tested and the number of objects to test. In the following, we propose a BDT construction method that makes such a compromise. Using this method, the more distinct the objects are, the earlier the stage at which they are tested. Thus, the quality of classification monotonically degrades along a path. This provides a natural way of encoding the global decisions into a sequence of binary bits with decreasing significance. Such a format becomes useful when it is desired to determine only the group to which an object belongs but not its exact identity.

Construction of BDTs

In our method, we start with the root node, and then repeat the following procedure for all new nodes as long as they contain more than one object. At each node we choose the two subsets of objects to distinguish, create its child nodes and associate with them the appropriate sets of objects for further classification. Suppose we are dealing with node t where a set $\Lambda(t)$ of objects remain to be further classified. Our goal is to select subsets Λ_l and Λ_r of $\Lambda(t)$ according to a criterion that balances the goal of object dissimilarity and the number of objects to test.

Let $|\Lambda_l \cup \Lambda_r|$ denote the cardinality of $\Lambda_l \cup \Lambda_r$, which is the number of objects to distinguish.

Let $D(\Lambda_l,\Lambda_r)$ be the dissimilarity between subsets Λ_l and Λ_r which is defined as

$$D(\Lambda_l,\Lambda_r)=\min_{\forall O_l,O_r: O_l\in\Lambda_l, O_r\in\Lambda_r} d(O_l,O_r)$$

where $d(O_l,O_r)$ is an information distance measure between objects O_l and O_r. An overview of such distance measures can be found in [8].

Let $D(n)=\max_{|\Lambda_l\cup\Lambda_r|=n} D(\Lambda_l,\Lambda_r)$ where $D(n)$ is the best possible dissimilarity that can be achieved when n objects are tested. It is not difficult to see that $D(n)$ is a monotone decreasing function of n.

During the construction of a BDT, it is desirable to maximize both $D(n)$ and n. But they are conflicting objectives and we need to balance these two objectives. Her we maximize n subject to

$$\alpha \ge 2\ln\frac{n}{2} - D(n) + D(2) \qquad (1)$$

where α is a prescribed constant.

This criterion is based upon the following observation. Suppose all objects occur with the same probability π. In a centralized recognition scheme, given $\{\Lambda_l, \Lambda_r\}$ and $|\Lambda_l \cup \Lambda_r| = n$, the probability of error $P_e(n)$ can be bounded by the following [8]

$$P_e(n) \le \sum_{O_l\in\Lambda_l}\sum_{O_r\in\Lambda_r} P_e(O_l,O_r)$$

where $P_e(O_l,O_r)$ is the probability of error when only O_l and O_r are tested.

The right hand side can be bounded by a class of information distance measures between O_l and O_r [8] and we have

$$P_e(n) \le \sum_{O_l\in\Lambda_l}\sum_{O_r\in\Lambda_r} ce^{-d(O_l,O_r)} \le \sum_{O_l\in\Lambda_l}\sum_{O_r\in\Lambda_r} ce^{-D(\Lambda_l,\Lambda_r)}$$

where c is a constant.

Using the best possible dissimilarity $D(n)$ we have

$$P_e(n) \le \sum_{O_l\in\Lambda_l}\sum_{O_r\in\Lambda_r} ce^{-D(n)} \le c\left(\frac{n}{2}\right)^2 e^{-D(n)}$$

Using the above bound as an approximation for $P_e(n)$, we have

$$\ln\left(\frac{P_e(n)}{P_e(2)}\right) = 2\ln\frac{n}{2} - D(n) + D(2)$$

The left-hand side reflects the performance loss due to an increase in the number of objects to test. Denote the maximum tolerable performance loss by α, we have the criterion (1).

Sometimes, there are multiple pairs of $\{\Lambda_l, \Lambda_r\}$ that correspond to the same value of $D(n)$, i.e., they yield the best possible dissimilarity. In such situations, one may compare two designs using the following method. For each design, find the shortest inter-object distance between Λ_l and Λ_r, and then choose the design with the smaller such distance. If there is a tie, count the number of all the object pairs between Λ_l and Λ_r that bear this distance, choose the design with the smaller count. If there is still a tie, repeat the previous procedure for the next shortest distance. If at the end, the two designs are still tied, randomly choose any one of them. This method is based on the fact that shorter distances contribute more to the error than longer distances.

Finally we present our algorithm for the construction of BDTs:

1. Set α. Let N denote the set of new nodes. Put root node t_0 into N. Set t_0 as the current node t.
2. If the current node t contains exactly one object, remove t from N, and go to step 3; otherwise use α to find the maximal n, $D(n)$ and the corresponding $\{\Lambda_l, \Lambda_r\}$. Create left child node t_l, let $\Lambda(t_l) = \Lambda(t) - \Lambda_r$. Create right child node t_r, let $\Lambda(t_r) = \Lambda(t) - \Lambda_l$. Put nodes t_l, t_r into N, remove t from N.
3. If N is empty, stop; otherwise find a member of N, set it as the current node t, and then go to step 2.

Design of the fusion rule and the sensor rules

At each node t, the sensors make local decisions regarding $\{\Lambda_l, \Lambda_r\}$. Based on these local decisions, the fusion center makes a global decision. Recall that a node affects the system performance more than any of its descendent nodes. Therefore at each node, we want to design the corresponding fusion rule and sensor decision rule in such a way that the probability of misclassification at that node is minimized regardless of what happens at its descendent nodes. This is basically a greedy algorithm. Such rules can be considered as the solution to a binary hypothesis-testing problem in which objects in Λ_l are tested against objects in Λ_r. Since BDTs are used at all the system elements, and the optimal fusion rule and the optimal sensor rules for the current node depend upon past sensor decisions and past global decisions. This adds to the complexity of the problem. It has been shown [11] that under a mild condition this problem reduces to a conventional binary decision fusion problem, and the optimal fusion rule and sensor rules can be designed based on results available in [1].

IV. An Example

In this section, we present an example to illustrate the results developed in the previous section. We use the communication protocol developed in the previous section to coordinate the sensors and the fusion center. We will discuss construction of the BDT, design of the sensor rules and the fusion rule, and performance evaluation. We also compare this design with the centralized scheme, an ad-hoc M-ary decision fusion scheme and the optimum single sensor scheme.

Let us consider a decision-fusion system consisting of three independent identical sensors and a fusion center. By identical sensors we mean that the sensor observations have the same characteristics and all the sensors use the same BDT. This system is used to identify four equally likely objects O_1, O_2, O_3 and O_4. Each sensor observation is assumed to be a scalar. The objects are represented by four evenly spaced points on the real line in the sensor observation space as shown in Figure 3. The distance between adjacent points is assumed to be a. A sensor observation is corrupted by additive white Gaussian noise of zero mean and unit variance.

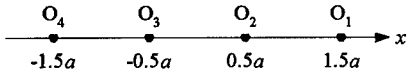


Figure 3: Object constellation

For the purpose of comparing our design to an ad-hoc M-ary decision fusion scheme that uses 2 bits for each sensor, we need a two level BDT that uses one bit at each level. Since the average number of steps is equal to 2 in this problem, we design a BDT and the corresponding sensor rules and the fusion rule that maximizes the average recognition rate.

Tree construction

We use the Kullback divergence to compute the dissimilarity between subsets of objects. Since the three sensors are identical and conditionally independent, the dissimilarity is additive and we have

$$d(O_i, O_j) = 3K(O_i, O_j)$$

where $K(O_i, O_j)$ is the Kullback divergence between O_i and O_j using a single sensor observation. Furthermore, since the noises are additive white Gaussian, the Kullback divergence is a quadratic function of the Euclidean distance between objects

$$K(O_i, O_j) = |O_i - O_j|^2$$

The inter-object distances for our problem are shown in Table 1.

Table 1: Inter-object distances

$d(.,.)$	O_1	O_2	O_3	O_4
O1	0	$3a^2$	$12a^2$	$27a^2$
O2	$3a^2$	0	$3a^2$	$12a^2$
O3	$12a^2$	$3a^2$	0	$3a^2$
O4	$27a^2$	$12a^2$	$3a^2$	0

Since the desired BDT has two levels, we need to evenly divide all objects into two subsets at the root node. Here n=4, and it is not difficult to find out that $D(4)=3a^2$. There are three candidate designs

Design 1: $\Lambda_r=\{O_1, O_2\}$, $\Lambda_l=\{O_3, O_4\}$.
Design 2: $\Lambda_r=\{O_1, O_4\}$, $\Lambda_l=\{O_2, O_3\}$.
Design 3: $\Lambda_r=\{O_1, O_3\}$, $\Lambda_l=\{O_2, O_4\}$.

Using the selection method that we developed in the previous section in case of ties, we find that the Design 1 is the best. Hence at the root node, we test $\Lambda_l=\{O_3, O_4\}$ against $\Lambda_r=\{O_1, O_2\}$. The design of the two second-level nodes is trivial, hence is omitted. The fusion center BDT is shown in Figure 4.

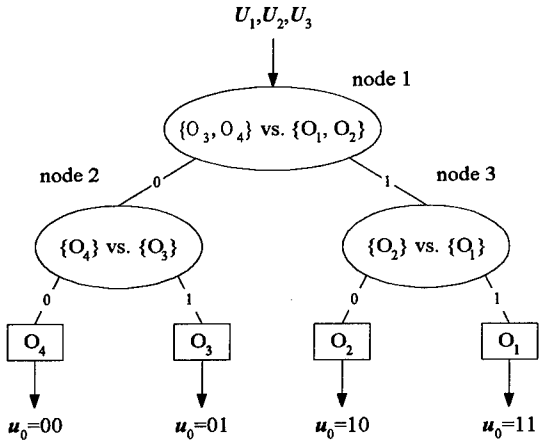


Figure 4: Fusion center BDT

The BDT used by the kth sensor is shown in Figure 5. In Figure 5, d is the binary decision made by the kth sensor at stage 1. We note that the kth sensor uses the global decision to navigate its BDT.

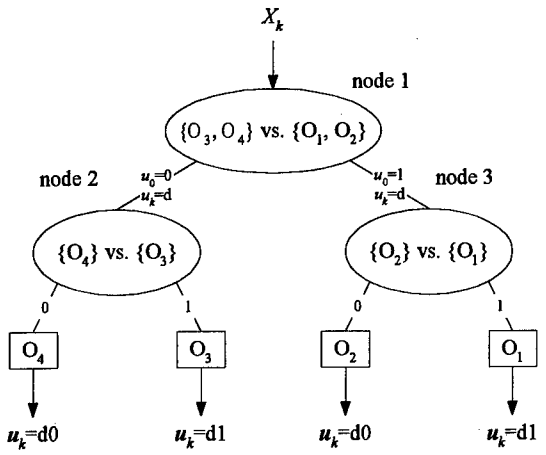


Figure 5: The kth Sensor BDT

Sensor decision rule and fusion rule

The system takes two stages to identify an unknown object. At stage1, it tests $\{O_1, O_2\}$ against $\{O_3, O_4\}$ and then chooses one of them say $\{O_1, O_2\}$. At stage 2, it tests O_1 against O_2. Thus we need to design sensor rules and the fusion rule for each stage. At each stage, we minimize the total probability of misclassification.

Stage 1

At this stage, we have a binary hypothesis-testing problem of $\{O_1,O_2\}$ vs. $\{O_3,O_4\}$. The sensor decisions u_1, u_2 and u_3 are single binary bits. Based on these sensor decisions, the fusion center makes a binary global decision. It is well known [9] that given fixed sensor rules, and given that the sensors are conditionally independent given the unknown object, the optimal fusion rule is the MAP fusion rule which can be expressed as

$$u_0 = \begin{cases} 1 & \text{if } \sum_{i=1}^{2}\prod_{k=1}^{3} p_i(u_k) \geq \sum_{j=3}^{4}\prod_{k=1}^{3} p_j(u_k) \\ 0 & \text{if } \sum_{i=1}^{2}\prod_{k=1}^{3} p_i(u_k) < \sum_{j=3}^{4}\prod_{k=1}^{3} p_j(u_k) \end{cases}$$

where $p_i(u_k)$ is the conditional probability of decision u_k of the kth sensor when the unknown object is O_i.

It is also well known [1] that the necessary condition for a sensor rule to be optimal under the conditional independence assumption is that it is a likelihood ratio test. Since each sensor observation X_k is a Gaussian random variable, a likelihood ratio test results in a threshold test and a binary partitioning of the real line x_k. Recall that identical sensor rules are used, and one can easily see that the decision boundary is at the origin x_k=0. With such sensor decision rules, the optimal fusion rule further simplifies to a majority rule

$$u_0 = \begin{cases} 1 & \text{if } u_1 + u_2 + u_3 \geq 2 \\ 0 & \text{if } u_1 + u_2 + u_3 \leq 1 \end{cases}$$

Stage 2

At this stage, the system either tests O_1 vs. O_2, or O_3 vs. O_4. Because of the symmetry of the object constellation, the sensor decision rules and the fusion rule at stage 1, the result on testing O_1 vs. O_2 is essentially the same as that for testing O_3 vs. O_4. So without loss of generality, we only consider testing O_1 vs. O_2.

Now the sensor decisions $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ and $\boldsymbol{u}_3$ are two-bits vectors with the first bit representing the sensor decision at stage1 and the second bit representing the sensor decision at stage 2. Based on these vectors, the fusion center makes a global decision. Similar to the result at stage 1, the optimal fusion rule is

$$\boldsymbol{u}_0 = \begin{cases} 11 & \text{if } \prod_{k=1}^{3} p_1(\boldsymbol{u}_k) \geq \prod_{k=1}^{3} p_2(\boldsymbol{u}_k) \\ 10 & \text{if } \prod_{k=1}^{3} p_1(\boldsymbol{u}_k) < \prod_{k=1}^{3} p_2(\boldsymbol{u}_k) \end{cases}$$

where $p_i(\boldsymbol{u}_k)$ is the conditional probability that the kth sensor decision is $\boldsymbol{u}_k$ when the unknown object is O_i.

Again by the same argument as used for stage 1 [1], the optimal sensor rules are likelihood ratio tests. Such a test corresponds to a binary partitioning by means of a threshold either in region $x_k>0$ if the first bit of $\boldsymbol{u}_k$ is 1, or in region $x_k<0$ if the first bit of $\boldsymbol{u}_k$ is 0. These decision boundaries are functions of the inter-object distance a.

We let a range from -15dB to +5dB with step size equal to 0.5dB. For each value of a, we compute the optimal sensor decision rules, and then express the MAP fusion rule as a Boolean function of the binary sensor decision vectors. This result is given in Table 2. Here $\boldsymbol{u}_0$ is the global decision that the fusion center makes at node 3. When $\boldsymbol{u}_0$=11, O_1 is declared, and when $\boldsymbol{u}_0$=10, O_2 is declared. $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ and $\boldsymbol{u}_3$ are the incoming sensor decision vectors. Since the fusion center chooses $\{O_1, O_2\}$ at the root node using a majority fusion rule, there are at least two 1s out of the first bit of $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ and $\boldsymbol{u}_3$. Different permutations of these vectors share the same entry in this table. One can see that the fusion rule changes with the inter-object distance a. Such phenomena are marked by adjacent 11 and 10 for $\boldsymbol{u}_0$ in Table 2. For small values of a (<-5dB), the fusion center prefers O_1(11) to O_2(10) unless the sensors have strong support for O_2. This is because O_2 is sandwiched between $\{O_3, O_4\}$ and O_1, and there is little room for O_2. For large values of a (>4dB), the fusion center chooses O_2(10) over O_1(11) unless the sensors strongly support O_1, In this situation, because the signal is sufficiently strong, most of the time the sensor decision vectors belong to the upper four entries in Table 2. In such cases, the fusion rule is essentially a majority rule. For intermediate values of a (from -5dB to 4dB), the fusion rule exhibits a change from the weak signal form to the strong signal form.

Table 2: Stage 2 fusion rule: $O_1(\boldsymbol{u}_0=11)$ vs. $O_2(\boldsymbol{u}_0=10)$

a (dB)			(-15, -5)	(-4.5, -3)	(-2.5, 3)	3.5	(4, 5)
u_1	u_2	u_3	u_0				
10	10	10	11	10	10	10	10
10	10	11	11	11	11	10	10
10	11	11	11	11	11	11	11
11	11	11	11	11	11	11	11
10	10	00	10	10	10	10	10
10	10	01	10	10	10	10	10
10	11	00	10	10	10	10	10
10	11	01	11	11	10	10	10
11	11	00	11	11	10	10	10
11	11	01	11	11	11	11	10

Performance evaluation

The system performance is given by the probability of misclassification P_{mc}. In Figure 6, the P_{mc} of the BDT based decision fusion system is plotted against the inter-object distance a as a solid curve. In the same figure, we also plot the performance of the optimal centralized classifier and the performance of the optimal single sensor classifier. The optimal centralized classifier uses all the raw sensor observations to make a one-stage classification of the unknown object. It has the smallest possible P_{mc} that serves as a lower bound to the P_{mc} of any other scheme. This P_{mc} is plotted as a dotted curve. The optimal single sensor classifier uses one sensor observation to classify the unknown object. The corresponding P_{mc} is plotted as squares in Figure 6. The BDT based decision fusion system outperforms the optimal single sensor classifier. Also it is slightly inferior to the optimal centralized classifier. This is further shown in Figure 7 where the increase in P_{mc} normalized by that of the centralized classifier is plotted. For instance, when a=-10dB, for the optimal single sensor classifier, P_{mc}=0.6558; for the optimal centralized classifier, P_{mc}=0.5881; for the BDT based system, P_{mc}=0.5998.

In Figure 7, the BDT based decision fusion system is compared with the ad-hoc decision fusion system in which the three sensors are designed as identical optimal single sensor classifiers and the MAP fusion rule is used to fuse their decisions. For each system, the increase in P_{mc} with respect to the optimal centralized classifier is plotted. It is shown that the BDT based decision fusion system is better than the ad-hoc decision fusion system. For instance, when a=-10dB, for the ad-hoc system, P_{mc}=0.6047. The increase of P_{mc} for the BDT based system is -17dB, for the ad-hoc system is -15.5dB.

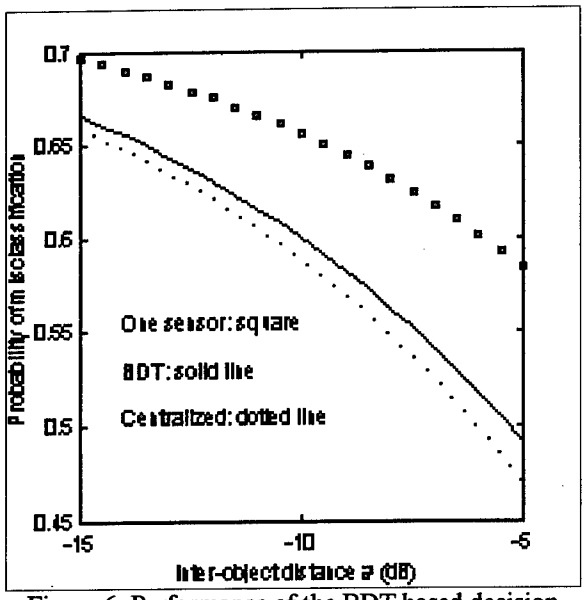


Figure 6: Performance of the BDT based decision fusion system

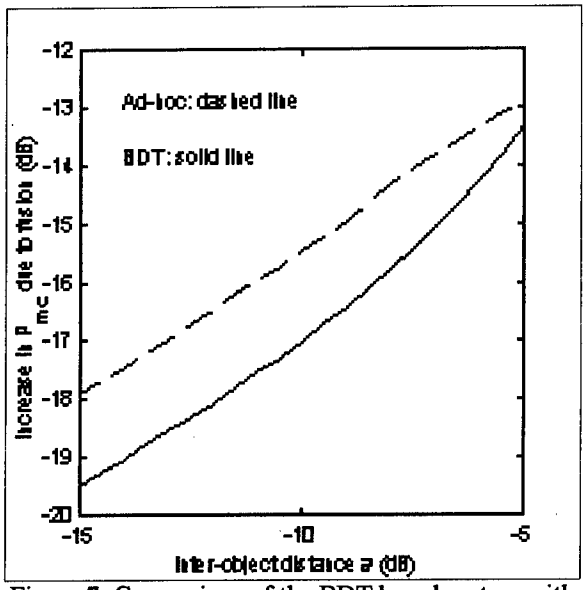


Figure 7: Comparison of the BDT based system with an ad-hoc system

V. Summary

We considered a new class of decision fusion problems that employs binary decision trees at the sensors and/or at the fusion center. Such problems were categorized into four types according to how the binary decision trees are designed and used by the fusion system. Various aspects of the design of such systems were discussed. We proposed a systematic design methodology for one such system. This methodology was illustrated by means of an example. Many aspects of this class of problems remain unsolved and provide a fruitful area for future research.

Acknowledgement

Research sponsored by Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. F49620-96-1-0185.

References

[1] Pramod K. Varshney, *Distributed Detection and Data Fusion*, Springer-Verlag New York, Inc. 1997.

[2] R. Viswanathan and P. K. Varshney, "Distributed Detection with Multiple Sensors-Part I: Fundamentals," *Proc. IEEE*, vol. 85, No.1, pp. 54-63, Jan. 1997.

[3] R. S. Blum, S.A. Kassam and H. V. Poor, "Distributed Detection with Multiple Sensors: Part II – Advanced Topics," *Proc. IEEE*, vol. 85, no. 1, pp. 64-79 Jan. 1997.

[4] S. Rasoul and D. Landgrebe, "A Survey of Decision Tree Classifier Methodology," *IEEE Trans. AES*, May 1991, pp. 660-674.

[5] K. Demirbas, " Distributed Sensor Data Fusion with Binary Decision Trees," *IEEE Trans. AES*, Vol. 33, No. 2, pp. 223-234, Aug. 1993.

[6] B. Dasarathy, "Operationally Efficient Architectures for Fusion of Binary-Decision Sensors in Multidecision Environment," *Opt. Eng.*, 36(3), pp. 632-641, March 1997.

[7] X. Zhu, M. Kam and C. Rorres,"M-ary Hypothesis Testing with Binary Local Decisions," *Proc. CISS*, Princeton, March 1998.

[8] M. Basseville, "Distance Measures for Signal Processing and Pattern Recognition," *Signal Processing*, 18 (1989) pp. 349-369.

[9] Z. Chair and P.K. Varshney, "Optimal Data Fusion in Multiple Sensor Detection Systems," *IEEE Trans. AES*, Vol. AES-22, No. 1, pp. 98-101, Jan., 1986.

[10] R.R. Tenney and N.R. Sandell, "Detection with Distributed Sensors," *IEEE Trans. AES*, Vol. AES-17, pp. 98-101, July 1981.

[11] Q. Zhang, Ph.D. dissertation in progress.

Session RC3
Sensor Fusion for Automatic Target Recognition
Chair: Erik Blasch
Air Force Research Laboratory, OH, USA

Fusion of HRR and SAR information for Automatic Target Recognition and Classification

Erik Blasch
Air Force Research Lab
2241 Avionics Cir, WPAFB, OH 45433 USA
Email: erik.blasch@sensors.wpafb.af.mil

Abstract: *This paper explores the fusion of moving High Range Resolution (HRR) and stationary Synthetic Aperture Radar (SAR) for automatic target recognition and classification. The tradeoffs of resolution and time-to-classify are investigated through simulation. By using a fusion approach, targets are effectively classified in a multitarget-multisensor scenario; however the Bayesian analysis does not account for measurement confidences.*

1. Introduction

Multisensor automatic target recognition (ATR) algorithms include target detection, classification, recognition, and identification [1]. One of the key issues in a moving and stationary ATR assessment is the decision when to use the Synthetic Aperture Radar (SAR) mode to get an two-dimensional image of the target, shown in Figure 1. A radar sensor manager must select sensors, select the detected moving and stationary targets to classify, and align the sensors to the target [2]. Thus, the sensor manager must control the measurement and ATR process from recognition to classification. A sensor classification policy is best described as a problem in sequential decision making under uncertainty. Prominent elements of the problem include competitors, a dynamic environment with uncertainties in target movement and measurement clutter, and complexity arising from many possible sensor actions and outcomes.

From an ATR point of view, geometric target information of movement is essential to selecting radar modes. Since you can't generally predict the geometric perspective of an object in the image it makes it difficult to determine whether the target is moving or stationary. You can simplify the target recognition and classification task by determining coarse information concerning the target type and movement from a one-dimensional HRR sensor. If the target is stationary, the HRR return will be cluttered, but this procedure saves time in determining whether to wait for the SAR update. Many algorithms have focused on the finer analysis of automatic target recognition [3,4], however, these algorithms may have a processing time constraint for real-time operations. In the cases of tracking scenarios, it might be beneficial to have a coarse measurement system to capture moving targets and a fine measurement system to classify the target. We seek dynamic target properties, as measured by spatial/spectral intensity of a 2D SAR output and 1D HRR range measurements. Thus, understanding the sensor management HRR and SAR tradeoffs can be useful for 1D and 2D radar fusion.

ATR is the ability of a system to detect and recognize a target. Fusing information obtained from other sensors, the ATR solution can be extended to include target classification and identification. One of the inherent limitations of radar processing for target classification is that the target dynamics need to be known *a priori* as in the case of HRR for moving targets, shown in Figure 2, or SAR for stationary targets, shown in Figure 1. We explore the use of a Bayesian metric to capture unknown target dynamics to determine whether a detected target is transitioning from moving to stationary or stationary to moving

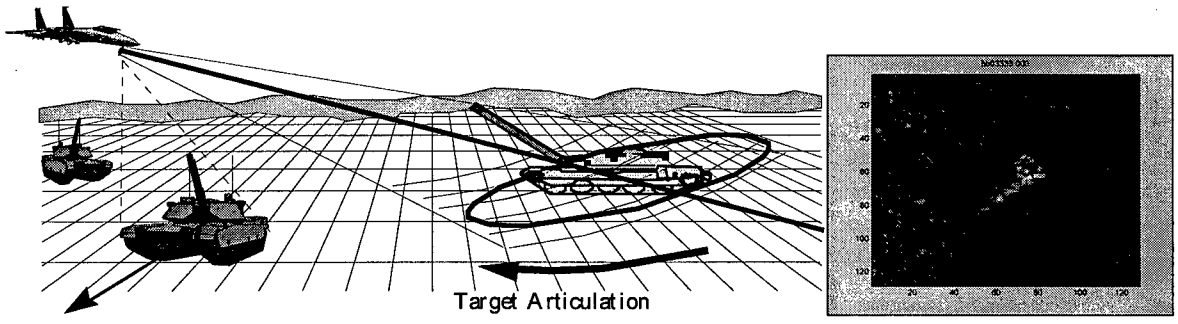


Figure 1. Synthetic Aperture Radar Collection for a Tank.

modes. In the case of a moving target, a one-dimensional HRR profile can be used to classify the target size, whereas in the case of a stationary target, a 2D SAR image can be used to determine the target type.

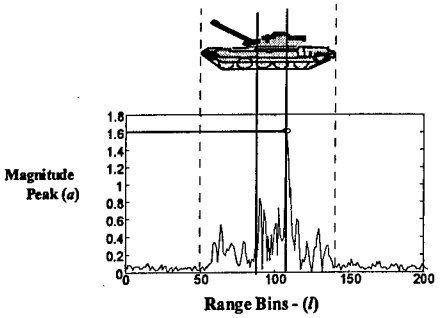


Figure 2. HRR Profile, showing target size.

This paper explores radar fusion, resolution comparisons, and confidence differences to determine whether a target is in a stationary or moving mode. The initial work is focused towards addressing the ATR problem of classifying targets transitioning from a stationary to moving state or vice versa. Section 2 formulates the problem and Section 3 describes the Bayesian approach. Section 4 presents results and Section 5 draws conclusions and suggests further work.

2. Problem Formulation

Feature extraction can be used for object tracking, identification, and classification. For tracking, image content and registration are important for time and location referencing [4]. The image content includes coarse information on the target type and movement. Additionally, ATR algorithms are subject to capacity constraints. For instance, if the image is thought to be traveling through an information channel, then the desired output is to maximize the information available, given bandwidth and time constraints.

Radar systems are effective for surveillance applications due to their distance range resolution invariance, all weather, and measurement capabilities. The radar antenna has a tradeoff between ground moving target indicator (GMTI), HRR, or SAR modes, shown in Figure 3. HRR radar offers a method for imaging *moving* targets by extracting energy returns from range profiles. If the target is *stationary*, a collection of HRR radar signatures can be processed to form a SAR image. In the HRR radar mode, the cross range resolution is the *radar beam width*, which is large at long range. However, the both HRR profiles and SAR images are formed from radar scans, but differ in number of scans used in coherent integration. SAR processing can be achieved in conjunction with GMTI for detection of the relative target location which limits the beam width. Once detected, SAR information values can enhance target classification confidence. HRR and GMTI information enhances the target classification by reducing the search area associated with the target and determining the target size, but is limited in confidence since it is a 1D scan. Likewise, target classification helps predict movement detection for each target. In addition to correctly classifying targets [5], a target classification system in military

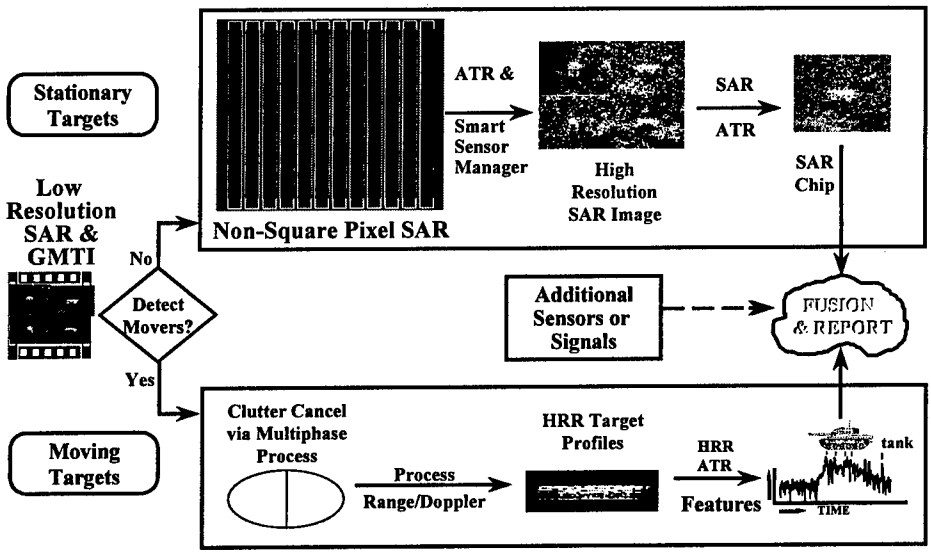


Figure 3. Detection and SAR Image Extraction for Stationary Targets versus HRR Data for Moving Targets.

scenarios must also robustly classify unknown targets.

2.3 Scenario

The scenario is an aircraft classifying a target using multi-sensor fusion [6,7]. There are two kinds of sensors: an HRR and SAR. Each sensor returns two elements: 1) a probability value and 2) a 10 bit measurement vector that represents the proposition to which the probability is attached. There are 10 types of targets that can be grouped into four SAR target classes: {truck, small tank, large tank, other}, and three HRR sizes: {short}, {medium}, and {long}. If a bit is set to 0, this reflects that the sensor declares that the target in question is not of the type associated with that bit, and if the bit is 1, the target may be of that type.

In the simulation, the aircraft is 100 nautical miles (nm) away from and approaches the stationary or moving target. The measurements are taken every 1 nm for the HRR sensor and every 7 or 10 nm for the SAR sensor, depending on whether the target is moving or stationary. Uncertainties vary with the type of sensor and are proportional to the distance between the aircraft and the target. The performance confidence(*a priori* sensor characteristic probability) is shown in Table 1. Each sensor is provided with an array of data. The data set includes the {{range (*nm*)}, {10 measurement bits}, {sensor performance confidence value}}. Two sets of data (for a truck and tank) are analyzed for the different methodologies.

Table 1. Sensor Characteristics

		T_1	T_2	T_3	T_4	T_5	T_N
	prior	0.10	0.10	0.10	0.10	0.10	0.50
Short	$P(H_S \mid T_k)$	0.50	0.50	0.50	0.25	0.25	0.33
Medium	$P(H_M \mid T_k)$	0.25	0.25	0.25	0.25	0.70	0.33
Long	$P(H_L \mid T_k)$	0.25	0.25	0.25	0.70	0.25	0.33
Truck	$P(S_{Tr} \mid T_k)$	0.70	0.10	0.10	0.70	0.10	0.25
Small Tank	$P(S_{ST} \mid T_k)$	0.10	0.70	0.10	0.10	0.10	0.25
Large Tank	$P(S_{LT} \mid T_k)$	0.10	0.10	0.70	0.10	0.10	0.25
Other	$P(S_O \mid T_k)$	0.10	0.10	0.10	0.10	0.70	0.25

3.0 Theoretical Background

3.1 Bayesian Probability Analysis

The joint, marginal, and conditional probabilities can be combined to form the mutually exclusive properties of *Bayes' Rule:*

$$P(B_j \mid A_i) = \frac{P(A_i \mid B_j) \cdot P(B_j)}{\sum_{i=1}^{N} P(A_i \mid B_j \cdot P(B_j))} \quad (1)$$

where $P(A_i|B_j)$ is the likelihood function, and $P(B_j)$ is the update from the *a priori* information. From Bayes' Rule, these *axioms* hold:

$$P(\phi) = 0 \quad (2)$$

$$P(\bar{A}_i) = 1 - P(A_i) \quad (3)$$

$$P(A_i \cup B_j) = P(A_i) + P(B_j) - P(A_i B_j) \quad (4)$$

$$P(A_i) = P(A_i|B_j) \cdot P(B_j) + P(A_i|\bar{B}_j) \cdot P(\bar{B}_j) \quad (5)$$

To update the uncertainty based on the new evidence, Bayes' Rule is formulated as:

$$P(C \mid A_i B_j) = \frac{P(B_j C \mid A_i)}{P(B_j \mid A_i)} \quad (6)$$

If C is an element of a mutually exclusive and collectively exhaustive set of potential outcomes, and B is a set of data that has been collected, then:

$$P(C_k \mid A_i B_j) = \frac{P(B_j C_k \mid A_i)}{\sum_{k=1}^{N} P(B_j \mid C_k A_i) \cdot P(C_k \mid A_i)} \quad (7)$$

from which it can be rewritten as:

$$P(C_k \mid A_i B_j) = \frac{P(B_j \mid C_k A_i) \cdot P(C_k \mid A_i)}{\sum_{k=1}^{N} P(B_j \mid C_k A_i) \cdot P(C_k \mid A_i)} \quad (8)$$

where:

1. $P(C_k|A_i)$ is an *a priori* (or prior) probability of C_k occurring, based on the state of information A_i;
2. $P(C_k|A_i \cdot B_j)$ is the *a posteriori* (or posterior) probability of C_k given data B_j observed and prior state information A_i;
3. $P(B_j|C_k \cdot A_i)$ is the *likelihood function*, the likelihood of observing data B_j conditioned on C_k and prior information state A_i;
4. $\sum_{k=1}^{N} P(B_j \mid C_k A_i) \cdot P(C_k \mid A_i)$ is the *preposterior* probability of the observing the occurring data, given the prior state information, but conditioned on all possible outcomes C_k.

It is possible to aggregate the probability statements from a lower level of abstraction to a higher one using the equations above, and the development can be derived for continuous as well as discrete events and scalar and vector and matrix notations. The likelihood expressions represent how confident (subject to change) a given probability statement is. The functions must be developed prior to collecting

the data by analysis. Note that the *preposterior* is simply the combination of all the likelihood functions and the prior distributions.

3.2 Bayesian Analysis of the HRR Sensor

The Bayesian method is based upon the basic probability axioms. The first step in the Bayesian approach is to determine, for each sensor, its likelihood function based upon target type. Thus, for the HRR sensor, the relationship is:

$$P_{HRR}(D|T_k) = \sum_{i=1}^{N} P_{HRR}(D|S_i) \cdot P(S_j|T_k) \qquad (9)$$

where S_i is $Size_i$ (S - short, M - medium, L - Long), D is Data, and T is the type. $P_{HRR}(D|S_i)$ must be determined from the data measured and the *prior* probabilities learned through experience as shown in Table 1. The probabilities are combined by the relationship:

$$P_H = 1 = P_H(D|S) + P_H(D|M) + P_H(D|L) \qquad (10)$$

The data given is in the form of binary values for detection form a single sensor, such as : {{ 1, 1, 1, 0, 0,0,0,0,0,0}, {0.5}} which is in the from {{Pr(Detect Short), Pr(Detect Medium), Pr(Detect Long)},{Pr(Prior)}}. Taking into account data confidence and bit values, the following probabilities are updated:

$$P_H(D|L) = [1+Pr(Prior)] \cdot P_H(D|L) \cdot P_H \qquad (11)$$

So, $P_H(D|L) = [1+ 0.5] \cdot (0.222) = 0.333$
and similarly, $P_H(D|L) = P_H(D|M) = 0.222$.

The above analysis satisfies the condition that if the confidence of the data is zero, then the maximum uncertainty entropy is achieved, while if the confidence of the data is one, mutual information is obtained from measurements in time. For the above data, if the confidence is zero, each probability would then equal 0.222, which satisfies the entropy condition.

3.3 Bayesian Analysis of the SAR Sensor

Following the same procedure, the SAR sensor (over its set) gives:

$$P_{SAR}(Data|T_k) = P_{SAR}(Data|T_1) \cdot P(T_1)|T_k) + \ldots + + P_{SAR}(Data|T_N) \cdot P(T_N)|T_k) \qquad (12)$$

and the relationship below holds:

$$P(T_i|T_k) = \begin{cases} 1 & \text{for } i = k \\ 0 & i \neq k \end{cases} \qquad (13)$$

So the probability of the type data can be simplified for the SAR sensor to be just equal to the measured data itself. The interpretation of the data for the SAR sensor is similar to that the HRR sensor. For example, if the SAR sensor returns the data string: {{0,0,0,1,1,0,0,0,0,0}},{0.7}} of the form {{$P(T_{1\text{-}10})$ }, P(Prior)}; then the probabilities would be determined as follows:

$$\sum_{i=1}^{N} P_{SAR}(Data|T_i) = P_{SAR} = 1 \qquad (14)$$

$$P_S(D|T_i)) = (1+Pr(Prior) \cdot Pr(Detect\ T_i) \cdot \frac{1}{P_S} \qquad (15)$$

$$1 = \sum_{i=1}^{N} [(1+Pr(Prior) \cdot Pr(Detect\ T_i)] \cdot \frac{1}{P_S} \qquad (16)$$

So with the data given:

$$\frac{1}{P_{SAR}} = 2 \cdot (1+1 \cdot 0.7) + 8 \cdot (1+ 0 \cdot 0.7) = 3.4$$

which gives $P_{SAR} = 0.294$

So with the data given $P_{SAR}(D|T_{(4,5)}) = 0.5$ and $P_{SAR}(D|T_{(1,2,3)}) = 0.294$.

3.4 SAR / HRR Fusion

The fusion of SAR and HRR information is a function of whether the target is moving or stationary; however fusion is necessary if a target is transitioning from stationary to moving. Additionally, a situation may result where a target is moving and stopping in which case there is fusion of information over time. The sensor data fusion is completed by the Bayesian update of information. Since the above information is assumed available, then *independent* likelihood functions can be integrated to get the joint likelihood function based upon target dynamics. The data fusion is performed by:

$$P_{Fused}(Data|T_k) = P_{HRR}(D|T_k) \cdot P_{SAR}(D|T_k) \qquad (17)$$

but the information desired is the likelihood of target type, given the data instead of likelihood of data given the type. Using Bayes' Rule, the relationship is:

$$P_{Fused}(T_k)|Data) = \frac{P_F(Data|T_k) \cdot P_{Prior}(T_k)}{P(Data)} \quad (18)$$

where the normalizing factor is:

$$P(Data) = \sum_{i=1}^{N} P_F(Data|T_k) \cdot P_{Prior}(T_k) \quad (19)$$

To determine the target SAR class and HRR target size, the axioms of probability are used, where the joint likelihood function and the prior information are used to obtain the data:

Target Size: {Short, Medium, Long}

$$P(N_i|D) = \sum_{i=1}^{N} P_F(T_k|D) \cdot P(N_i|T_k) \quad (20)$$

$$P(S|T_k) + P(M|T_k) + P(L|T_k) = 1 \quad (21)$$

Target Class: {Truck, Small Tank, Large Tank, Other}

$$P(C_i|D) = \sum_{i=1}^{N} P_F(T_k|D) \cdot P(C_i|T_k) \quad (22)$$

$$P(Tr|T_k) + P(ST|T_k) + P(LT|T_k) + P(O|T_k) = 1 \quad (23)$$

A concern in using Bayes' rule is the need for *a prior* distribution over the events of interest. In the real world, Bayes' rule necessitates a subjective interpretation of probability. By using the *principle of indifference*, one can arbitrarily set the probabilities equal for each outcome. The Bayesian approach to sensor integration recursively updates probability information at each measurement; however, measurement uncertainty is not captured with its implementation.

4.0 Results

Four test cases were run. The first case was the normal, control test for target 2 with one SAR update for every 7 seconds and a measured confidence. The second test was run with lower confidences. The third case was run for a different test target. The fourth case was run with SAR updates of 10 seconds.

4.1 Normal Case

For the normal test case, a short tank was run with a SAR update of every 7 seconds.

Note from Figures 4-6 that the HRR is faster to determine the length of the target and that SAR is slower but has a higher probability. The final Bayesian probability is the *a priori* probabilities associated with the sensor; however the fused result reaches a higher confidence.

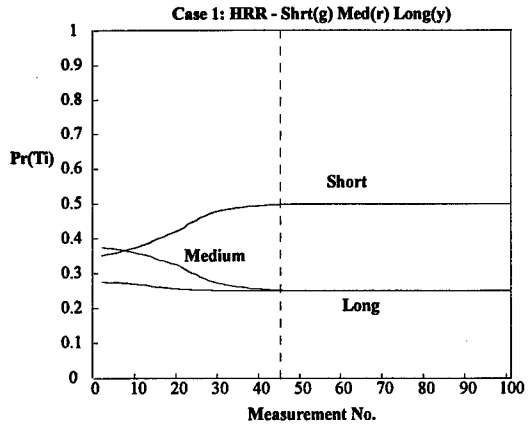


Figure 4. HRR Probability.

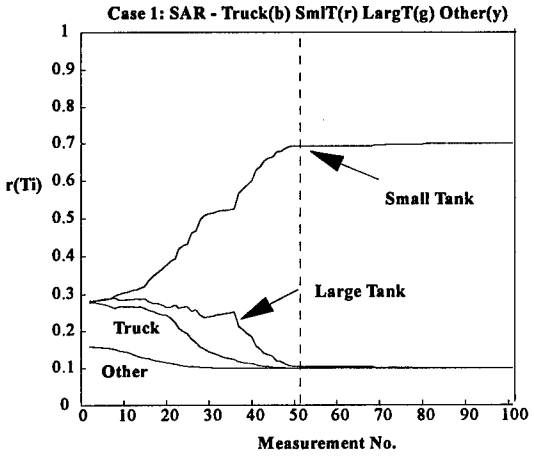


Figure 5. SAR Probability – Update 7 Seconds.

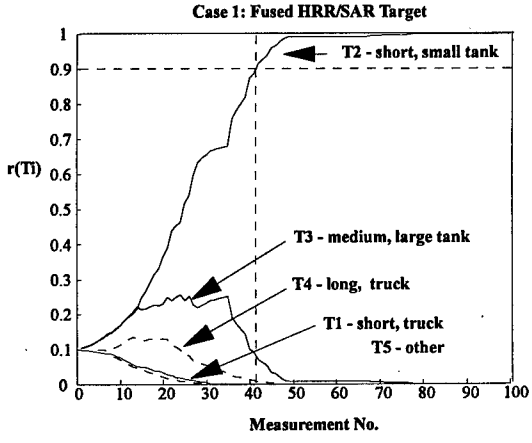


Figure 6. Fused HRR and SAR(7) Probability.

4.2 Lower Confidence Updates

For the second test case, the short tank was run with a SAR update of every 7 seconds and the probabilities were reduced by half. Note that the inherent normalization by the Bayes' rule results in the same

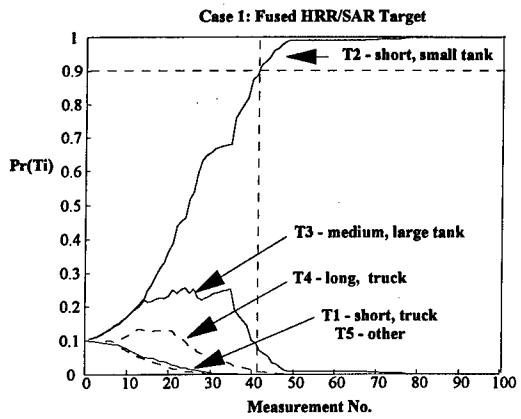


Figure 7. SAR Probability – Lower Confidence.

values. The same values are a function of the available probabilities. Hence, Bayes' rule has a limitation in that it does not capture incomplete sensor knowledge.

4.3 Another Test Target

For the third test case, the long truck was run with a SAR update of every 7 seconds with the measured probabilities.

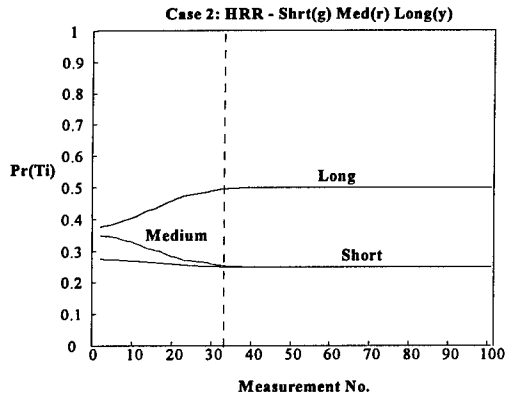


Figure 8. HRR Probability.

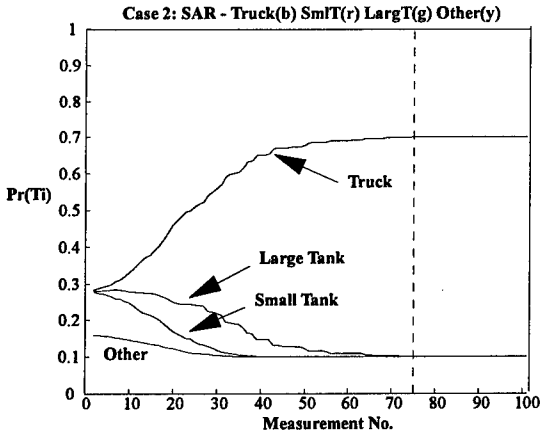


Figure 9. SAR Probability – Update 7 Seconds.

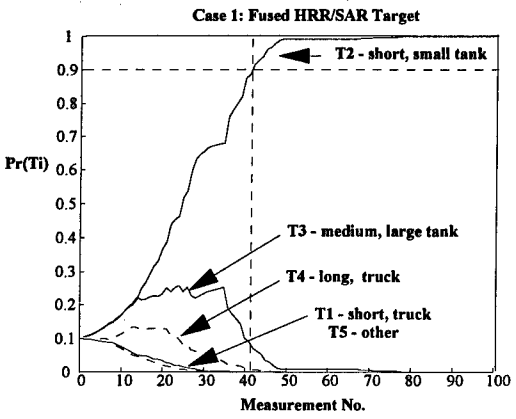


Figure 10. Fused HRR and SAR Probability.

In Figures 8-10, we see that the probability updates are similar to the first target case.

4.4 SAR Update Every 10 Seconds

For the fourth test case, the short tank was run with a SAR update of every 10 seconds with the measured probabilities. Note, HRR is given more confidence in the decision making since the target is moving.

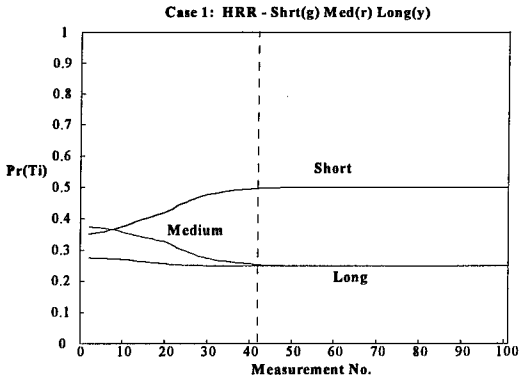


Figure 11. HRR Probability.

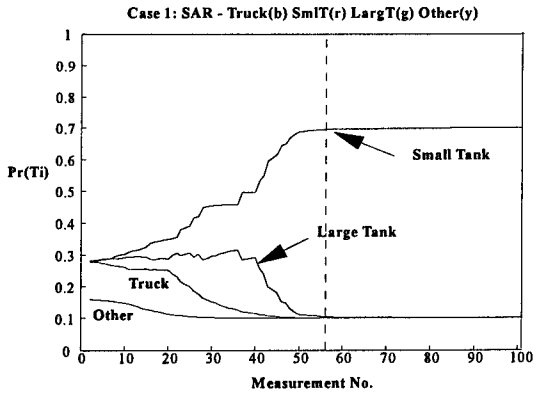


Figure 12. SAR Probability – Update 10 Seconds.

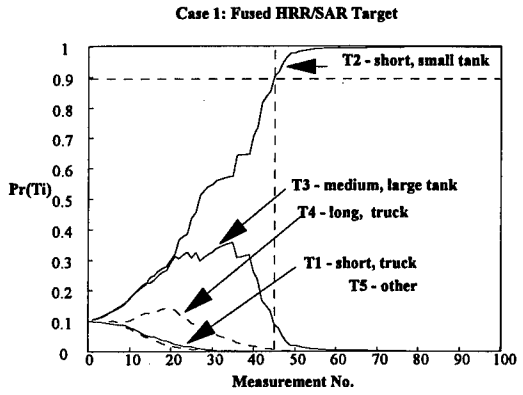


Figure 13. Fused HRR and SAR Probability.

4.5 Time Comparisons

The time comparisons are shown as a comparison for decision making where the fused estimate is shown for minimizing the time to classify the target.

Table 2. Time Comparisons

Time to Reach Decision	Fuse	HRR	SAR
Case 1 (Small Tank)			
Normal (SAR update 7)	41	45	51
Lower Confidence	41	45	51
SAR Update 10	44	41	57
Case 2 (Light) - L, Short			
Normal (SAR update 7)	42	31	71
Lower Confidence	42	31	71
SAR Update 10	48	25	82

5.0 Conclusions

The paper addressed a situation in which a target was recognized and classified when it was transitioning from a stationary to a moving scenario. The research included methods to detect and classify a lone dim target with imperfect HRR and SAR sensors. In a series of simulation experiments, the fused result obtained a desirable solution. Using a Bayesian metric in a recursive approach classifies the target; however, it does not account for sensor confidences and thus is not robust. Further research will focus on a combination of predicting target dynamic techniques for classification problems and exploration in problems involving multiple stationary and moving targets, multiple sensors, and inclusion of state preferences and obscured image features. We will further explore real data and develop an algorithm to overcome the limitations of non robust classification through confidence measures.

References

[1] E. Blasch, "Information assessment of SAR for ATR." *NAECON98*, Dayton, OH, July, 1998, pp. 414 – 419.

[2] S. Musick and K. Kastella, "Comparison of Sensor Management Strategies for Search and Classification", *9th National Symposium on Sensor Fusion*, Naval Postgraduate School, Monterey, CA, March 11 - 13, 1996.

[3] E. Zelnio and F. Garber, "Characterization of ATR performance Evaluation," *SPIE Signal Processing, Sensor Fusion, and Target Recognition V*, Orlando, FL, April 8 - 10, 1996.

[4] E. Blasch, "SAR Information Exploitation Using an Information Filter Metric," *7th ATRWG*, Monterey, CA, March 2-4, 1999.

[5] D. Castañon "Optimal Detection Strategies in Dynamic Hypothesis Testing," *IEEE Transactions on Systems, Man, And Cybernetics*, Vol. 25, No. 7, July 1995, pgs. 1130-1138.

[6] L. Hong, and A. Lynch, "Centralized/Distributed Recursive Algorithms for Information Fusion by Dempster-Shafer Techniques with Applications to Target Identification," *IEEE Trans. Aerospace and Electronic Sys*, V.28, 1992.

[7] E. Waltz and J. Llinas. *Multisensor and Data Fusion*. Artech House, Boston, MA, 1990.

A SAR-FLIR Fusion ATR System

Yang Chen *and* **Kurt Reiser**
HRL Laboratories, LLC, RL69
3011 Malibu Canyon Road
Malibu, CA 90265
{ychen, kreiser}@hrl.com

Abstract

We present an automatic target recognition (ATR) system which fuses information from synthetic aperture radar (SAR) and forward-looking infrared (FLIR) images in air-to-ground applications. The system is a hierarchical architecture, in that fusion of information takes place at detection, indexing and hypothesis levels. Combining complementary information from SAR and FLIR allows us to achieve higher detection and lower false alarm rates than are attainable with a comparable ATR system using FLIR images only. For hypothesis fusion, we use the Dempster-Shafer (D-S) belief function for its convenient representation of uncertainty and its ease in weighting information from multiple sources. Experimental results using real SAR and FLIR data are presented.

Key Words: Automatic Target Recognition, ATR, Information Fusion, Sensor Fusion, Dempster-Shafer

1. Introduction

Automatic Target Recognition systems relying on a single sensor can have serious performance limitations due to problems inherent in the specific sensor modality. For example, an ATR system using a forward-looking infrared sensor may have a good recognition capability at close range to the targets even in complete darkness. But the system's range capability and the area of coverage might be limited by the resolution of the FLIR sensor the system has. Also, FLIR sensors are susceptible to weather such as rain or snow, which lowers the thermal contrast of the scene. A synthetic aperture radar-based ATR system, on the other hand, has excellent long-range capability, and can operate in adverse weather conditions. However, SAR sensors have resolution-dependent aperture times, which restrict multiple map processing gain, and SAR target signatures are strongly dependent on viewing geometry.

In this paper, we present an ATR system which combines information from both SAR and FLIR images for air-to-ground targeting applications. Our SAR-FLIR fusion ATR system ("fusion ATR system" for short) tries to combine the best of both sensor modalities to achieve a target recognition performance level neither sensor alone can reach. Our system consists of a hierarchical fusion architecture with emphasis on fusion of information at detection, indexing and hypothesis levels. Detection level fusion associates target candidates detected from SAR with those detected from FLIR. This association reduces system false alarms significantly, thanks to the complementary responses to SAR and FLIR for natural clutter. Feature level fusion allows us to inject information extracted from one modality into the processing in the other modality restrict target candidate hypotheses and reduce the probability of object misclassification. The fusion of SAR and FLIR at the hypothesis (class) level is carried out using the Dempster-Shafer (D-S) belief function calculus. D-S belief function formalism has a more natural representation for the information at hand and also offers a good mechanism to weight different information sources. A system block diagram of our fusion ATR system is shown in Figure 1.

The rest of the paper is organized as follows. We first introduce the D-S representation and the Dempster's Rule of Combination in Section 2. In Section 3, we discuss detection-level fusion, i.e., the matching of SAR and FLIR detections. We present hypothesis generation methods for SAR and FLIR in Sections 4 and 5, respectively. In Section 6, we discuss the issue of modeling and handling non-target (clutter) objects. Hypothesis fusion method is presented in Section 7. Finally, we give experimental results in Section 8, and conclude this paper with a summary in Section 9.

2. Information Fusion Using Dempster-Shafer Belief Function Theory

The Dempster-Shafer belief function theory [1] is one of the calculi used by researchers for uncertainty reasoning [2] and information fusion [3][4]. The D-S theory can be considered as an extension of the traditional probability theory in that it assigns mass to *sets* of discrete outcomes of a variable, rather than only to singleton outcomes as in the case of probability. In D-S theory, the domain of a variable D is called the "frame of discernment," denoted as W_D, and is represented by a set of finite and exclusive outcomes that the variable can take:

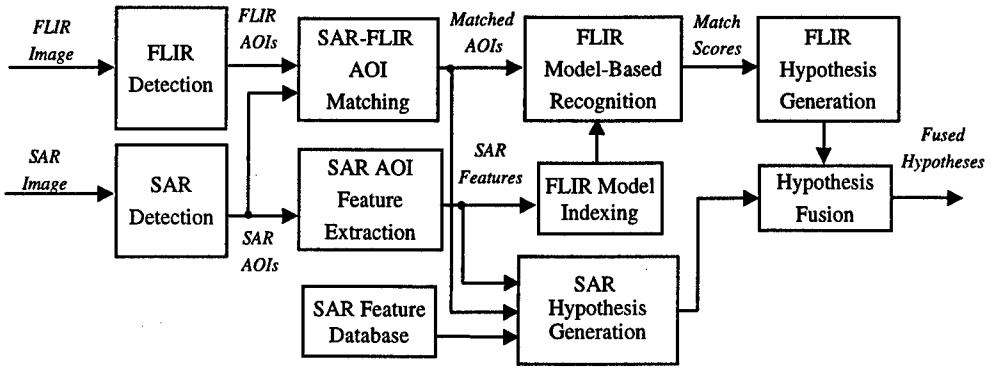


Figure 1. The Fusion ATR system block diagram.

$$W_D=\{d_1, d_2,..., d_L\} \tag{2.1}$$

A belief function representation of certain knowledge about D is often represented as a *basic probability assignment*, or BPA, over the frame of discernment. A BPA on W_D is a mapping from the subsets of W_D to [0, 1]:

$$m: 2^{W_D} \rightarrow [0.0, 1.0] \tag{2.2}$$

with the following constraints:

$$\sum_{A \subseteq W_D} m(A) = 1.0, \quad m(\varnothing) = 0.0 \tag{2.3}$$

where 2^{W_D} is the power-set (the set of all subsets) of W_D. The definition of the BPA requires that $m(A)$ represent the evidence of, or support for, the subset A and only A, not each individual elements or subsets of A. This is a way of representing ignorance about how the mass is distributed among the elements of the subset. Therefore information can be appropriately represented at either fine-grain (element by element) level, or at a coarser (subset) level, offering different level of ignorance and commitment for each piece of information. At one extreme, if all mass is allocated to singletons, the BPA becomes a familiar probability distribution, and we call the result a *Bayesian* BPA. On the other hand, when all the mass is assigned to a single subset:

$$m(A) = 1.0, \quad A \subseteq W_D \tag{2.4}$$

This is called a *logical* BPA. When the subset A in the above equals to the whole frame W_D, the BPA is called a *vacuous* BPA:

$$m(W_D)=1.0 \tag{2.5}$$

It is "vacuous" since it does not tell us anything about the variable D. It represents the state of complete ignorance.

Two alternative representations of the BPA function are the belief function Bel() and the plausibility function Pl(), which have the following relationships with the corresponding BPA:

$$\text{Bel}(A) = \sum_{B \subseteq A} m(B), \quad A \subseteq W_D \tag{2.6}$$

$$\text{Pl}(A) = \sum_{B \cap A \neq \varnothing} m(B), \quad A \subseteq W_D \tag{2.7}$$

While a BPA represents each individual pieces of support for the subsets of the frame of discernment, Bel(A) represents the sum of evidence in support of A and evidence which implies A, and Pl(A) represents all evidence that may potentially support A. Any representation among the three of BPA, Bel() and Pl() are entirely interchangeable. Hence they are often collectively referred to as the "belief function" representation.

To use D-S belief function to combine or fuse information from multiple sources, we use the so-called "Dempster's Rule of Combination." Suppose we have two BPA's, m_1 and m_2, on the same frame W_D, which represent distinct pieces of information. Then the combined belief function m_3 can be written as

$$m_3 = m_1 \oplus m_2 \tag{2.7}$$

where "$\oplus$" is the operator for the combination using Dempster's rule, and m_3 is defined by

$$m_3(A) = \frac{1}{K} \sum_{B \cap C = A} m_1(B) m_2(C), \quad A \subseteq W_D \tag{2.8}$$

where K is a normalization constant

$$K = 1 - \sum_{B \cap C = \varnothing} m_1(B) m_2(C) \tag{2.9}$$

When a piece of information comes from an unreliable source, we can use "discounting" [1] to discount its BPA before combining it with other BPA's. The discounting operation is carried out as follows:

$$m'(A) = \begin{cases} (1-\alpha) m(A) & A \neq W_D \\ m(W_D) + \alpha(1 - m(W_D)) & A = W_D \end{cases} \tag{2.10}$$

where $0 \leq \alpha \leq 1$ is a discount factor. The bigger the α is, the more the BPA m is discounted. If α is 1.0, then the resulting BPA m' becomes a vacuous BPA.

3. Detection Level Fusion

Our current system assumes that separate SAR and FLIR subsystems work on their own to detect potential target candidates, called "areas of interest" or AOI for short. In order for the fusion ATR system to combine information about the AOI's, each AOI from the FLIR needs to be matched with an AOI from the SAR, and vice versa. That is, we need to find some mapping R:

$$R(\mathbf{p}) = \begin{cases} \mathbf{q} & \text{if } \mathbf{p} \text{ and } \mathbf{q} \text{ are of the same object} \\ \varnothing & \text{if there is no correspondence} \end{cases} \quad (3.1)$$

where $\mathbf{p} \in \{\mathbf{p}_i\}$ and $\mathbf{q} \in \{\mathbf{q}_j\}$, $\{\mathbf{p}_i\}$ and $\{\mathbf{q}_j\}$, are the sets of AOI's in the FLIR and the SAR images, respectively. A secondary, but very important, goal of the matching is to remove clutter AOI's from the candidate list to reduce false alarms.

In general the matching problem can be very complicated due to the non-linear nature of the imaging transforms involved. In our case, we can assume that the sensor parameters, including the location, sensor viewing geometry and the (intrinsic) sensor model, are known. If the sensor parameters are accurate, we can easily achieve the mapping in Eq.(3.1) using a geometric transformation T_R derived from the sensor parameters:

$$\mathbf{q}_j = T_R(\mathbf{p}_i | \mathbf{q}_j = R(\mathbf{p}_i)) \quad (3.2)$$

In reality, there are always noise and errors associated with the sensor parameters. Therefore Eq.(3.2) is only correct in theory.

Our approach to the problem consists of two steps, an initial match step and a refinement step. In the initial match step, we first use the sensor parameters to construct an approximation of the transformation T_R. Let us call it T'_R. We then project the FLIR AOI's into the SAR image coordinates

$$\mathbf{q}_i' = T'_R(\mathbf{p}_i) \quad (3.3)$$

Now let

$$C = \{ (\mathbf{q}_k, \mathbf{q}'_k) \mid \mathbf{q}_k = R(\mathbf{p}_i), \mathbf{q}_k' = T'_R(\mathbf{p}_i) \} \quad (3.4)$$

be the set of pairs of SAR AOI's and the corresponding transformed FLIR AOI's for all the common objects detected in SAR and FLIR. Due to the noise and errors in the sensor parameters, $\mathbf{q}_k$ and $\mathbf{q}'_k$ usually do not align exactly. That is,

$$\mathbf{d}_k = \mathbf{q}_k - \mathbf{q}'_k \neq 0 \quad (3.5)$$

These residual errors are solved in the match refinement step. We use the fact that under certain assumptions, the difference in the locations between $\mathbf{q}_k$ and $\mathbf{q}'_k$ can be approximated by a global translation:

$$\mathbf{d}_1 = \mathbf{d}_2 = \ldots = \mathbf{d}_k = \mathbf{d} \qquad k = 1, \ldots, \|C\| \quad (3.6)$$

where $\|C\|$ is the cardinality of the set C. Therefore our goal is to find the global translation $\mathbf{d}$. We have developed a method based on a generalized Hough transform [9] to solve this 2-D matching problem involving only translation. Our method solves for $\mathbf{d}$ in Eq.(3.6), and also updates the estimated transformation T'_R. Due to the page limit of this paper, the details of this technique has been left out for a separate presentation elsewhere.

The result from the matching process is the set C (see Eq.(3.4)) which establishes the correspondence among the SAR AOI's and the FLIR AOI's. Note that C only represents the common set of objects *detected* in both SAR and FLIR. Since the target detection process in either SAR or FLIR could miss certain targets, we may not have all the targets we want in C. There are two ways we can reduce the risk of missing a target. One is to keep the remaining AOI's from both SAR and FLIR, knowing that we will not be able to fuse any information from the other modality for them, and the recognition result will be as good as with a single sensor. Another way, which is what we are using in this paper, is to lower the detection thresholds in both the SAR and FLIR processing modules to ensure that all targets are detected. This will inevitably increase the number objects detected as AOI's, which act as "noise" and may cause problems for the matching process. Fortunately, our matching method is very robust to this high clutter-to-target ratio situation. By keeping only the common object set in C, we can eliminate most of the clutter AOI's that do not appear in the other sensor's image. The primary benefit of this is to be able to achieve both a very low miss rate (due to the lowed detection thresholds) and a very low false alarm rate at the same time. A secondary benefit is the reduced system load and therefore increased throughput due to the reduced number of AOI's the system need to process from that point on (especially for FLIR).

4. SAR Hypothesis Generation

SAR target candidate detection is carried out using a standard constant false alarm rate (CFAR) detector. Once the CFAR processing is carried out, the points in the SAR image exceeding a set threshold are clustered and screened to rule out unlikely targets (e.g., either too small or too large). The image point clusters form the SAR AOI's, and their locations are computed from the center of gravity of the points in each cluster. Other properties of the clusters that may bear clues about the underlying objects' identities can be computed. These include target length, width and orientation, among others. We call these properties "target features." The list of SAR AOI's is used for matching with the FLIR AOI's as described in Section 3. The result is a reduced list of matched AOI's. We

then can use the features of each SAR AOI from this reduced list to generate target identity hypotheses.

But before we can generate SAR hypotheses, the system needs to know the relation between the objects we are interested in (the targets) and the features. Let $F=\{f_1, f_2,..., f_M\}$ be the set of features, $X=\{x_1, x_2,...,x_K\}$ be the set of values a feature $f_i \in F$ can take, and $O=\{t_1, t_2,...,t_N\}$ be the set of targets. In a traditional statistical pattern recognition system, target features are collected from a large collection of sample target images, and the conditional probability distributions for the features, $P(f_i \mid t_j)$, are estimated. (We have assumed that the feature values of a given target type are mutually independent). At run-time, a set of feature values, one for each feature, of an AOI is computed and the identity t of the AOI can be determined according to the maximum *a posteriori* (MAP) principle:

$$t = t_k \left| \max_k \left(\prod_i P(f_i \mid t_k) \right) \right. \qquad (4.1)$$

In Dempster-Shafer, there is an analogous way of computing posterior belief, which is called "conditional embedding" [5] or "generalized Bayesian Theorem" [6]. The conditional embedding scheme encodes the conditional probability distribution $P(f_i|t_j)$ for a given t_j into a BPA m_{ij} on the joint frame of the targets O and the feature measurements X. This belief function has the desirable properties that when marginalized to the frame of the feature values, it yields vacuous belief, and when conditioned on a feature observation $f_i=x$, it results in the conditional probability $P(f_i=x|t_j)$ itself. Combining all the m_{ij}'s for all target types t_j's using Dempster's Rule of Combination gives us a BPA for feature f_i,

$$m_i=m_{i1} \oplus m_{i2} \oplus ... \oplus m_{iN} \qquad (4.2)$$

When given the measurement of a feature, $f_i=x$, we can combine the BPA representing this feature measurement, with m_i in Eq.(4.2), and marginalize the result to the target frame O, which gives us the posterior belief given the feature measurement. It can be shown that the resulting posterior belief, denoted as m_{Si}, can be expressed in a closed formula [7]. We can then combine all the posterior beliefs from all features with Dempster's rule:

$$m_S = m_{S1} \oplus m_{S2} \oplus \cdots \oplus m_{SM} = \bigoplus_{i=1}^{M} m_{Si} \qquad (4.3)$$

This represents the combined hypothesis from the SAR for a single AOI.

5. FLIR Processing and FLIR Hypothesis Generation

5.1 FLIR Detection and Model Indexing

Like most model-based ATR system using FLIR images as input, ours works as follows. A target candidate detection process detects "hot spots" in a FLIR image where a cluster of image pixels contains higher (or lower) values than the surrounding background due to IR emission from the potential target. These "hot spots" can be further filtered to screen out clutter objects using the target size anticipated to be seen in the FLIR images. This results in a list of AOI's from the FLIR module, which is used for matching with the SAR AOI's as described in Section 3. The SAR-FLIR matching module returns a reduced list of AOI's to the FLIR module for further processing, carrying with each FLIR AOI the corresponding SAR AOI and its features.

The next step in the FLIR module is model matching. The system stores a model database consisting of a set of "templates" for each of the target types in the target set. A matching process evaluates the similarity of an AOI to each of the templates in the model database. A similarity measure is computed for each template thus visited. At the end, the AOI is given a score for each target type based on the best-scored template for that target type.

Obviously, this exhaustive search strategy for the best match is not very appealing since the model database can get quite large. The model templates for each target need to cover the entire view of the target at full 360-degree aspect and at different elevation angles. In addition, the models also have to cover targets as seen from different distance. To reduce the computation involved in the search, we use the information from the corresponding SAR AOI to trim the model database and reduce the number of model templates that need to be compared. We call this model "indexing." For example, each SAR AOI has an estimated orientation and size of the underlying object. These can be used to index a set of model templates that fit these features. With indexing we can usually reduce the search space by a factor of five without sacrificing the system performance.

5.2 FLIR Hypothesis Generation

Hypothesis generation in FLIR converts the output from the model-based ATR system into BPA's for hypothesis fusion. As mentioned in the last section, the FLIR model matching process gives scores to each AOI based on the similarity measures of the AOI to every target type in the model database. Let $G=\{g_1, g_2,...,g_N\}$ be the set of scores an AOI receives, where

$0 \le g_i \le g_{max}$ is the score for target type t_i, i=1,...,N, and g_{max} is a fixed positive number. In our case, the bigger the score g_i is, the better the match of the FLIR AOI to the target. Now we need to convert the set of scores G into a BPA representation so that we can combine it with the BPA's representing the SAR hypotheses for the corresponding SAR AOI.

There are many ways we can perform the above-mentioned conversion. In a simple-minded treatment, we can normalize the scores by the sum of the scores so they add up to 1.0, and use the result as a Bayesian BPA over the target type frame. The disadvantage of this approach is that there is no fixed scale in the resulting number (since the normalization factor is different for each AOI). An immediate consequence of this is that one will not be able to compare the converted scores of two FLIR AOI's for the same target type to see whether either AOI is more similar to the said target type. It can also result in false confidence in the result since if all g_i's are small (i.e., the AOI matches poorly to the models), the normalized scores can become large.

Our current method involves in three steps, equalization, normalization and encoding. The idea behind equalization is that we know the similarity score g_i has an uneven distribution (more densely distributed towards the lower end). An equalization process can bring the scores into a uniform distribution in the range [0, 1], which linearizes the scale of the similarity measure, and makes comparison with a variable threshold easier, if desired. Since we have limited training image for FLIR, we use an approximation to the equalization by a function as follows:

$$g_i' = 1 - \frac{1}{g_i^4 + 1}, \quad g_i \in G \tag{5.1}$$

The resulting g'_i has a range of [0, 1]. In the second step, we use a fixed normalization (by the maximum of the total score, N) because we need to preserve the scale across all AOI's. The disadvantage is that the mass (and therefore the belief) for any single target has a limit of $1/N$. Finally we can express the FLIR hypothesis in BPA form as:

$$m_F(A) = \begin{cases} \dfrac{g'_i}{N} & A = \{t_i\}, t_i \in O \\ 1 - \dfrac{1}{N}\sum_{j=1}^{N} g'_i & A = O = \{t_1, t_2, \cdots, t_N\} \end{cases} \tag{5.2}$$

where N is the total number of targets in the target frame. In other word, we first re-normalize the equalized score, and then assign the result as the mass for each target type. The remainder of the mass goes to the full target frame. This corresponds to the system's "ignorance" about the AOI based on the FLIR.

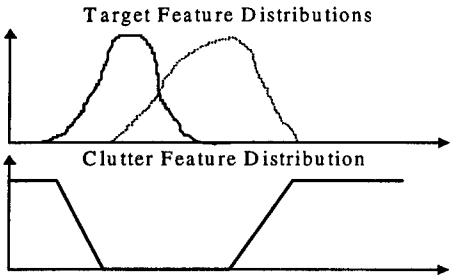


Figure 2. Modeling clutter feature distribution as the complement of the target feature distributions to help target discrimination

6. Modeling Clutters

Clutters are non-target objects detected in the SAR or FLIR images. Clutter objects are not explicitly modeled in our system so far. Part of the difficulties in modeling clutter objects is that there can be too many types of clutter, caused by both man-made and natural objects. Therefore it is not feasible to have a representative set of images of different types of clutter for training. In our system, these issues are dealt with separately in SAR and FLIR.

For SAR, since we can screen the AOI's based on the size, the type of clutter that the detection module will pick up in SAR is limited. We can use a uniform feature distribution to model the clutter objects, as is done in the traditional Bayesian approach. An alternative method is to model the clutter features for better discrimination than recognition in that it will provide target or non-target information rather than which target type. This is because the resolution of the SAR image in our system is not high enough for reliable target identification. The SAR information is used mostly to screen out non-target clutters. We can achieve this by using a distribution that contains 0 or a very small value in the range of all target feature distributions, and high values otherwise, as shown in Figure 2. This way, we can include "clutter" as another type of target (e.g., t_N) in our database for target hypothesis generation as described in Section 4.

For FLIR, since there is no "model" for clutter objects, we cannot perform model-based matching as we do with the known targets. Therefore we do not have "clutter" scores from the model-based ATR system. However, we still maintain a clutter type in the target frame O in our system. If t_N is assigned to the clutter type as in SAR, then the FLIR hypothesis BPA shown in Eq.(5.2) needs to be modified as follows:

$$m_F(A)=\begin{cases}\dfrac{g'_i}{N} & A=\{t_i\}, t_i\in O, i\neq N\\ 1-\dfrac{1}{N}\sum_{j=1}^{N} g'_i & A=O=\{t_1,t_2,\cdots,t_N\}\end{cases} \quad (6.1)$$

In other words, we do not attribute any positive mass toward clutter objects, which is reasonable since we do not have any measured evidence about it in our system.

7. Hypothesis Fusion

We have discussed how to generate SAR hypotheses from SAR feature distributions and FLIR hypotheses from FLIR model matching scores in Sections 4, 5, and 6. Now we are ready to perform the next step of fusion in the ATR system, the hypothesis fusion. In this section, we discuss how hypothesis fusion is done, special consideration and the treatment of unbalanced information contents between SAR and FLIR, and how we make decisions for target identification based on Dempster-Shafer belief functions.

7.1 Combining Hypothesis

Due to the way the SAR images are collected, we do not get updated SAR images as frequently as we do with the FLIR images. In fact we have only one SAR image for a sequence of FLIR images. This is often the case in a tactical air-to-ground ATR scenario, where a SAR image is acquired by the host aircraft or a third-party, and the FLIR is acquired by the host aircraft when it is near the target area. This means all information we have about the targets through the SAR comes from the same SAR image, while the information from the FLIR gets updated frequently (up to the FLIR image rate).

Since we only have one SAR image, the SAR detection and hypothesis generation need to be carried out only once. As each FLIR image comes in, we need to perform SAR-FLIR AOI matching for each FLIR image with the same SAR image, resulting in possibly a different set of matched AOI's for every FLIR frame. For hypothesis fusion, we combine the SAR hypothesis and the FLIR hypothesis for each matched AOI using Dempster's Rule of Combination, resulting in a separate set of fused hypotheses for each FLIR image. An alternative approach would be to track the FLIR detections across the sequence of images and use some integrated measure (based on the model matching scores) as the basis of FLIR hypotheses. This will potentially give better and more consistent results than if we consider each FLIR image separately. This is currently not done in this paper.

7.2 Hypothesis Discounting

Another issue is related to the relative weights of SAR and FLIR hypotheses. We have found that if we proceed hypothesis fusion as outline above, we are putting too much weight on the SAR hypotheses. This is because the low-resolution SAR images we use are not very informative as mentioned earlier. As a result, any error in the SAR hypotheses will show up in the fused results for the entire FLIR sequence.

Discounting the SAR hypotheses solves this problem. Referring to Section 4, m_{Si} is the posterior belief given feature f_i's value. According to Eq.(2.10), the discounted BPA is

$$m'_{Si}(A)=\begin{cases}(1-\alpha)m_{Si}(A), & A\subset O\\ m_{Si}(O)+\alpha(1-m_{Si}(O)), & A=O\end{cases} \quad (7.1)$$

where α is a discounting factor between 0 and 1. In our tests, α is set to 0.8 through experiment. As a result, Eq.(4.3) needs to be modified accordingly:

$$m_S=m'_{S1}\oplus m'_{S2}\oplus\cdots\oplus m'_{SM}=\bigoplus_{i=1}^{M} m'_{Si} \quad (7.2)$$

7.3 Making Decisions for Target Identification

Using the results from Equations (6.1) and (7.2), the combined hypothesis for each matched AOI can be written as

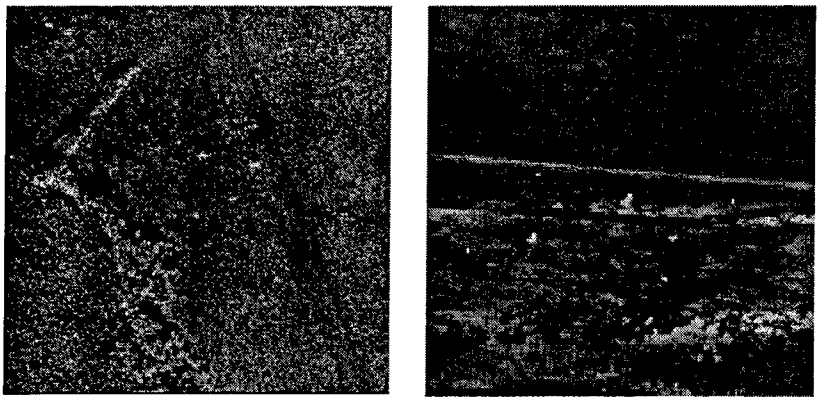
$$m_{fused}=m_S\oplus m_F \quad (7.3)$$

In order to make final decisions on the AOI's identity, we must have a scalar measure, such as a probability distribution, that represents the confidence level that an AOI belongs to certain target type. Since the Dempster-Shafer belief function (a BPA) does not support decision making directly, the Bel(), or Pl() values on the singletons have been used in the past for this purpose. We have used a method introduce by Smets [8], which converts the combined m_{fused} into a so-called "pignistic probability" distribution over the target frame O as follows:

$$P_{fused}(t_i)=\sum_{t_i\in A}\frac{m_{fused}(A)}{\|A\|}, \quad t_i\in O \quad (7.4)$$

The pignistic probability is optimal for decision making and it follows the so-called "generalized insufficient reason principle."[8] Once we obtain the pignistic probability, we can assign an AOI, say **p**, with target type t_i which has the largest pignistic probability:

$$TargetID(\mathbf{p})=t \mid P_{fused}(t)=\max_{t_i}(P_{fused}(t_i)) \quad (7.5)$$



(a) (b)

Figure 3. Sample SAR (a) and FLIR (b) images used in the tests

8. Experimental Results

We have tested our fusion ATR system on a number of test cases with real data, each of which includes a SAR image and a sequence of FLIR images. In these test cases, the SAR and the FLIR images contain three different targets, which, together with the "clutter" type, constitute the target frame O for the tests. Figure 3 shows a SAR and a FLIR image used in the tests. For each FLIR frame, we evaluate the pignistic probabilities (Eq. (7.4)) of the matched AOI's. For each matched AOI, the target type with the maximum pignistic probability is declared as the identity of the AOI (Eq.(7.5)). The result is then compared with the known truth.

To qualify the system's performance, we plotted the system ROC (receiver operating curve) using a scoring method called "top-n" as described below. From the matched AOI list we remove those AOI's that have been declared as clutter. Then we pick n AOI's with the largest (hence the name "top-n") pignistic probabilities as the potential targets and reject the rest. This gives us the performance shown in Figure 4.

Figure 4 shows the system performance for all test cases, with a total of 87 FLIR images. Figure 4(a) shows the system's recognition rates and (b) shows the detection rates. Three ROC curves showing the rates as functions of average false alarms per FLIR frame are plotted in each figure. The data points on the curves correspond to (from left to right) the n being set to 3, 4, 5 and so on in the top-n method. The first of these three curves (labeled as "HypothesisFusion") in each figure shows the performance of the full fusion ATR system. As a comparison, the third curve in each figure shows the system's performance without using any information from the SAR. This curve is labeled as "FLIR-only (50 AOIs)" because in each FLIR image, up to 50 most promising FLIR detections are kept and passed to the model-based FLIR ATR module. Since there are many more clutter objects than there are targets in the AOI's, the false alarm rates are very high as expected. This, however, is not a fair comparison between the fusion ATR and the FLIR-only ATR system, since if only FLIR images are used, we would not have retained 50 FLIR AOI's in each FLIR image. The reason we keep 50 of them for fusion ATR system is that our fusion system is able to reject clutters and reduce false alarms through detection-level fusion even if we use many more FLIR detections. This has the potential of increasing detection rates as has been shown in the performance curves. A more realistic operating level for a FLIR-only system would be to take only a few most promising FLIR detections for the model-matching process. On the one hand this reduces the system false alarm rate, but on the other hand it also increases the possibility of missing the real targets, causing the system detection rate to drop. The second curves (labeled as "FLIR-Only (5 AOIs)") in Figure 4(a) and (b) show what happens when we only use 5 AOI's for each FLIR frame in FLIR-only mode. Comparing this with the results for hypothesis fusion (the top curves), we can see the fusion ATR system is markedly superior to the comparable FLIR-only ATR system in both detection and recognition performance.

9. Conclusion

We have presented a fusion ATR system combining detection-level, indexing level and hypothesis-level fusion in a hierarchical architecture. Our experiments with real data show an increase of more than 15% in recognition and about 10% in detection rates compared with a FLIR-only ATR system at the same operating point in terms of false alarms per frame. The SAR images used in our experiments do not provide sufficient information for target identification purpose due to their relatively low resolution. Yet, our fusion ATR system is able to take advantage the complementary property of the SAR and FLIR

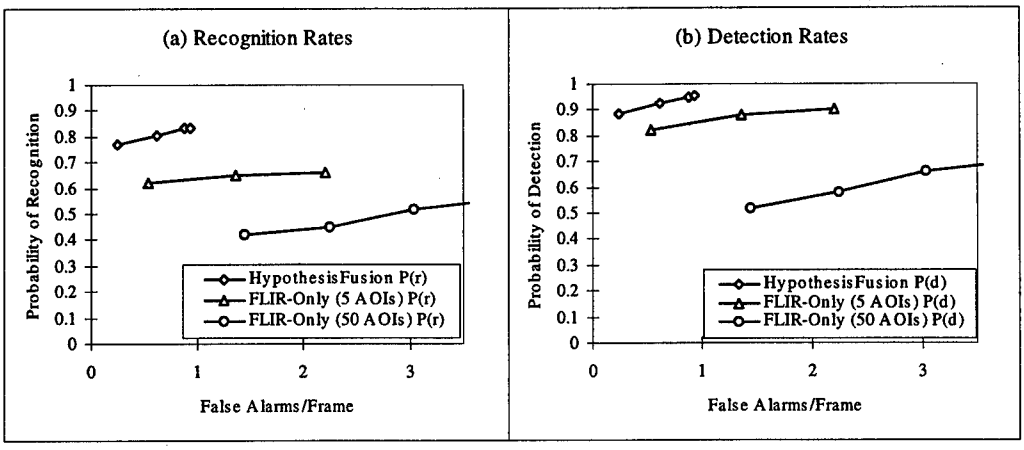


Figure 4. Fusion ATR system performance in comparison with that of a FLIR-only system

modalities and use information from both modalities to achieve a performance level neither SAR nor FLIR alone can achieve.

We use Dempster-Shafer belief function theory for the hypothesis fusion in our system. The belief function representation proved to be very useful in representing the type of information encountered in a typical ATR system. Statistical information (such as that for SAR features) is encoded into belief function through conditional embedding based on the Generalized Bayesian Theorem. FLIR model-matching scores are encoded into after equalization and fixed re-normalization. Furthermore, the D-S belief function provides a natural and powerful mechanism to weight multiple information sources before combining them through discounting, taking into account of the reliability and usefulness of the different information sources.

Acknowledgments

The authors would like to thank Dr. Charles A. McNary of Raytheon Systems Company (RCS) for providing support and guidance for the fusion ATR research at HRL Labs; Mr. David W. Webster of RCS for providing access to the FLIR ATR system and his support for the fusion activities; and Mr. Elbert Ezekiel of RCS for his support on SAR processing and feature extraction. We would also like to thank Mr. Neeraj Pujara of the Air Force Research Laboratory Sensors Directorate for his support of this work by providing test data used in this paper.

References

[1] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, 1976.

[2] P. P. Shenoy, "Using Dempster-Shafer's Belief-Function Theory in Expert Systems," in *Advances in the Dempster-Shafer Theory of Evidence*, R. Yager, M. Fedrizzi, and J. Kacprzyk, Eds., John Wiley & Sons, Inc., New York, 1994. pp.395-414.

[3] P. Bogler, "Shafer-Dempster Reasoning with Applications to Multisensor Target Identification System,". *IEEE Trans. on Sys., Man and Cyb.*, Vol.17(6), pp.968-977, 1987.

[4] Y. Lu, "Knowledge Integration in a Multiple Classifier System," in *Applied Intelligence*, Vol.6, pp.75-86, 1996.

[5] G. Shafer, "Belief Functions and Parametric Model,". *Journal of the Royal Statistical Society*, Vol.44(3), p. 322-352, 1982.

[6] P. Smets, "Belief Functions: the Disjuctive Rule of Combination and the Generalized Bayesian Theorem", in *Internationl Journal of Approximate Reasoning*, Vol.9, No.1, pp.1-35, 1993.

[7] Y. Chen, K. Reiser and Todd Baker, "Constructing Belief Functions on Propositions Using Conditional Embedding," Internal Report, Hughes Research Laboratories, Malibu, CA, July 1995.

[8] P. Smets, "Constructing the Pignistic Probability Function in a Context of Uncertainty", In *Uncertainty in Artificial Intelligence 5*, M. Henrion et al. Eds., Elsevier Science Publishers B.V. (North-Holland), 1990. pp.29-39.

[9] R.-C. Lo and W.-H. Tsai, "Perspective-Transformation-Invariant Generalized Hough Transform for Perspective Planar Shape Detection and Matching," *Pattern Recognition*, Vol.30, No.3, pp.383-396, 1997.

Some Considerations in Fusion, Tracking, Command and Control

Pailon Shar* **and X. Rong Li†**

Department of Electrical Engineering
University of New Orleans
New Orleans, LA 70148, USA
Phone: (504)280-7416, Fax: (504)280-3950, xli@uno.edu

Abstract-A bird's view of modern command and control (C^2) systems is presented. The concept, functions and structures of the C^2 systems are discussed, with an attempt to clarify some confusions existing in the community. Special emphases are on two fundamental subsystems, data fusion and target tracking, and their interrelationships with the modern C^2 systems.

Key Words: fusion, tracking, sensor, command and control

1. Introduction

For a modern command and control (C^2) system, its two basic responsibilities are combat command (or auxiliary command decision-making, or decision support) and fire control (or weapon control). The quality of command and control, however, relies directly on data fusion and target tracking, two information refineries and/or converters in a C^2 system. Fusion, tracking, command and control are so important and so closely interrelated that they constitute the backbone of the whole combat system.

Of these four, data fusion is, relatively speaking, a newcomer, but it is one of the systems under most active development over the past few years. Although older, target tracking also has made great advances in resent years. In fact, it is the combination of fusion and tracking that makes most of these advances possible. Fusion and tracking are so tightly coupled that it is sometimes difficult to separate them. The juncture of fusion and tracking has formed a very active research arena and their development enforces each other. This combination has also strengthened the capability of modern C^2 systems remarkably.

Apart from fusion and tracking, advances in many other aspects, such as computer, network and information processing technologies, have also contributed to the fast development of C^2 systems. As a result, C^2 systems have greatly changed over the years. These changes are reflected in almost every facet of a C^2 system, including its basic features such as the concept, functions and structures. These developments and changes, however, have not been systematically examined and studied. Confusion and chaos have emerged in many aspects. They have hindered academic exchanges in these areas. In the end they will in turn impede the development of C^2 systems per se.

This paper studies the aforementioned changes and their impact on the development of C^2 systems. Special emphases will be on those changes brought about by the advances in data fusion and target tracking.

2. C^2 System—Its Concept

In recent years, some technical terms of military systems are so widely abused that they often cause confusion and chaos, and may even mislead people. Great diversity and the fast development pace of military systems should take more blame for this chaos than those less careful users. For the sake of the system development itself as well as for convenience of academic exchange, clarification should be in order now.

It is impossible to describe a C^2 system without mentioning combat systems first because they are so closely connected. For a modern military system with the responsibility of a combat, three basic components are necessary. First, it must have necessary sensors and other intelligence channels that can provide information about the enemy force, own force and the combat environment. Secondly, it should have necessary weapons, both hard and soft, to attack or anti-attack the enemy force. Thirdly, for the purpose of converting the sensor information into weapon launching control information, a processing unit is also necessary. Roughly, such a military system can be called a combat system. Some may argue that corresponding military personnel to operate and com-

* Also named Peilun Xia, visiting scholar, on leave from Ocean University of Qingdao, P. R. China.

† Supported by ONR via Grant N00014-97-1-0570, NSF via Grant ECS-9734285, and LEQSF via Grant (1996-99)-RD-A-32.

mand the military system should also be part of the combat system. But more often by a combat system we mean the "machine" part.

Modern combat systems are evolved from older fire control systems. Before and during World War II, the concept of a combat system had not formed yet. At that time, there was no direct information channel between sensors and weapons. Weapon firing was controlled by a system that later was called fire control system or weapon control system. Fire control systems at that time were very simple. They usually can control only one piece of weapon to attack one enemy target. The computational devices in these systems were mechanical, or mechanical and electrical, and therefore the computational ability was quite limited. Information from sensors was orally reported and then entered manually into the computational device of the fire control system.

Great changes took place only after computers were introduced into military systems. On one hand, computers facilitated the processing and conversion of sensor information for the fire control system. Consequently, real combat systems began to emerge. On the other hand, with the ever-growing power of computers, more and more functions had become necessary functions of the system. For example, sensor information processing, tactical situation evaluation and display, threat evaluation, attack or evade decision-making are some of the functions that were impossible in those early systems. Now they are very typical functions of combat command or auxiliary command decision-making. Of course, fire control functions were maintained and enforced in these newer systems. Therefore, for the central processing part between sensors and weapons, its responsibilities lie in roughly two aspects: combat command and fire control. That is why it is widely called command and control (C^2) system. According to this definition, a combat system consists of three parts: sensors, C^2 system and weapons. So a C^2 system is a subsystem of a combat system and obviously it is the core subsystem.

However, in reality there exist many other names for C^2 systems. Some of them are even quite popular. C^2I (command, control and intelligence) system, C^3 (command, control and communication) system, C^3I system, C^4 (with the last C for computers) system, and C^4I system are some of them. These names may be defined clearly in military encyclopaedias, though different countries may have different definitions. Different names for the C^2 systems are used with an emphasis on certain part of the systems. For example, the names with an "I" usually emphasize the importance of information or intelligence gathering and processing. C^3 emphasizes the importance of communications and C^4 emphasizes computers. Other less popular names - e.g., tactical data processing and fire control system, combat information and weapon control system - can also be found occasionally. There are various reasons for this chaos. System diversity is one reason. Different countries and different system developers have their own naming systems. Technology progress is another reason. System developers keep on upgrading their products, and they often tend to emphasize their innovation by changing their names. This may be justifiable in some senses but so many names for a basically the same system is really annoying.

Nowadays, the combat system has been highly developed in almost every aspect. The number of sensors and weapons has greatly increased. The computers and the network are much more powerful and effective. The man-machine-interface (MMI) is friendlier. They can handle more and more targets and weapons simultaneously. New technologies such as data fusion and advanced target tracking techniques have great impact on almost every aspect of the entire system. The basic functions of the command and control system, however, remain virtually unchanged. They are still combat command and fire control. So the name of command and control system is not out of date yet.

3. C^2 System—Its Functions

It is important in many ways to understand the functions of the C^2 system. One of the reasons resulting in the naming chaos is the difference in system function designation. To define the functions exactly, however, is not so easy. The difficulty is that the functions of C^2 systems keep changing, and they may be quite different for different systems. Nevertheless, the basic lines can be drawn.

As stated before, the fundamental functions of a C^2 system can be roughly divided into two major parts: combat command and fire (or weapon) control. Combat command includes functions involving commanding information processing and display. On the other hand, fire control includes functions dealing with weapon launching information processing and display.

Combat command provides necessary information and means to assist associated commanders in decision-making. That is why it is sometimes called auxiliary command decision-making or decision support. Listed below are some basic command functions of a typical C^2 system.

1) Information gathering and processing: collect all possible information and convert it into a battle field situation picture as complete, accurate, and reliable as possible within system constraints.
2) Threat evaluation: evaluate the potential threat of every enemy target to own forces according to the situation picture and other necessary information like data base information.
3) Attack feasibility analysis: evaluate possible outcomes of attacking each target.
4) Choice making: attack or defense: the results of threat evaluation and attack feasibility analysis, among many other considerations, are used to make this choice.
5) Target indication and fire channel creation and management: once the decision to attack or defense is made, select targets to be attacked or evaded from; notify other related systems, e.g., the fire control channels; and choose weapons and corresponding launching devices.
6) Attack aid: if attack is chosen, provide necessary decision-making information for attack needed by the commander and the operators.
7) Defense aid: if defense is chosen, provide necessary decision-making information for defense.

Basically, fire control subsystem accepts instructions and necessary information from the command subsystem to fulfill the weapon launching and other related tasks. The following are some important fire control functions of a C^2 system.

1) Command and target acceptance: accept the target indication from the combat command subsystem and the related measurement sequences for each target.
2) Target motion analysis (TMA) or target tracking: estimate the kinematic states or parameters of the targets, such as position, velocity, heading and acceleration, with a significantly better accuracy than what is done in the command system.
3) Setting firing parameters: with target state estimates and property parameters of the selected weapon, it is possible to calculate the firing parameters like lead angle, turn angle, firing order and timing, etc. These parameters then should be preset into related weapons and launching devices.
4) Weapon launching control: the system can control the weapon launching procedure according to the preset time chain.
5) Weapon guidance and control: sometimes multiple firing waves are needed. The system should evaluate former waves to adjust upcoming waves. In other cases, system can control the weapons even after their launch. Wired torpedoes and wired anti-tank rockets are such examples. In these cases, the system should finish the guidance and control of the weapons at their targets.

It can be seen that the basic functions of both command and control remain almost unchanged. The contents of each function, however, have been remarkably enriched or strengthened. For example, information gathering and processing in early days may simply mean getting the measurement information from a single sensor and passing it over to the target tracking unit. It is definitely not comparable to the modern multisensor system with powerful information fusion abilities. Similarly for TMA - those primitive approaches with deterministic parameters are in no way comparable to advanced filters now widely employed in modern C^2 systems.

Apart from these basic functions, many systems have their own special capabilities. For example, navigation is so important to strategic ballistic missile submarines (SSBN) that accurate navigation ability is considered one of the necessary functions of their C^2 systems. This is not necessary the case for surface warships and attack submarines, although navigation is also very important for them. In addition, some other more technological than military functions are also very important. For example, modern systems are usually featured with user friendly interfaces in order to be more flexible and effective.

The automation of control functions is much earlier than that of command functions. Before their automation, command functions are human responsibilities - the commander and his subordinators made all assessments and decisions with the aid of primitive tools e.g., sand table and plot board.

4. C^2 System—Its Structures

The structure of C^2 systems is another complicated topic, primarily due to the great diversity of the systems. Roughly, three basic structures have been widely adopted. They are centralized, separated and distributed structures, respectively. The structure evolves naturally out of the advances in technology.

Computers were very expensive and clumsy when they were in their babyhood. One computer for one system was a natural choice. With all command and control functions centralized on such a primitive computer, they can not be expected to be powerful. That is why centralized C^2 systems usually can handle only single-target single-weapon situations. When better and less expensive computers became available, what

made the elimination of such systems inevitable are the more fatal defects, such as poor survivability, inconvenience of maintenance and lack of flexibility for system extension or redesign.

Separated C^2 systems are the consequence of less expensive yet more reliable computers, as well as the great advance in computer communication techniques. A few computers are used in a separated C^2 system. A typical example is a two-computer system, with one for command and the other for control. One notable feature of such separated systems is the overlapping in the responsibilities of the computers. The command computer can support some basic fire control functions as a back up. In case of failure of the fire control computer, the back up fire control functions in the command computer will be activated. So the system will not collapse, though some functions may be lost. For the same reason, the fire control computer has the potential of part of the command capability. This is the key to its superiority in survivability to the centralized systems. Of course, many other functions are also significantly enhanced, because of the more powerful computers and other innovations. For example, multitarget processing is a common task for the separated C^2 systems.

Some separated systems evolved naturally from their predecessors: fire control systems. During the transition from the fire control systems to the C^2 systems, some fire control systems had been inherited by their C^2 systems with little modification. Only a command subsystem was designed and added on the top of an already existing fire control system. A typical separated C^2 system then is formed.

In spite of its advantages over the centralized systems, a separated system is still relatively too "centralized." Its functions are "centralized" on a few computers. Although it is better than centralized on one computer, the improvement is limited. It is not difficult to imagine having a much better system if the system functions are more finely divided and implemented by more computers, and if the communication among these computers is sufficiently effective. This is the idea underlying a distributed C^2 system. The advances of microprocessors and network techniques in the late 80s provided a wonderful basis for the development of distributed C^2 systems. Now the newly developed C^2 systems are dominantly distributed in structure.

Although the computers used in the centralized and separated systems in the early days were called mini-sized or medium-sized, they were not as powerful as today's microcomputers. In modern distributed C^2 systems, more and more such powerful microprocessors are used, along with faster and faster local area networks. For the submarine C^2 systems example, some newly developed systems contain more than 100 powerful 32-bit microprocessors like Intel 80486 for general-purpose processing. Parallel computers are also used for special tasks, such as sonar and radar signal processing. In addition, optical fiber local area networks with transmission rate above 100M bps are widely used. With such a tremendous processing capability, many new devices, ideas and functions can be added into the system. Multisensor fusion and advanced tracking techniques are two important examples. Furthermore, this also makes it possible to have a high-degree functional redundancy, which is a prominent merit.

Another important feature of such distributed systems is that the traditional lines in a combat system to separate sensor, weapon and C^2 system blur now. With the high-speed data bus or information network as the center of a combat system, its component units are equally connected and treated. To the common data bus or network, each device is simply a node whether it is a sensor, a piece of weapon or a command and control module. A C^2 system is becoming more and more inseparable from a combat system and a combat system of this kind tends to be called a comprehensive combat system. In the meantime, the information flow within the system has also changed significantly with the structure development. Information in a centralized system flows predominately in one direction: from the sensor end to the weapon end. Information feedback from the weapon end to the sensor end is greatly enhanced in a separated system but the main stream is still sensor to weapon. Information in a distributed system flows in both directions with virtually equal opportunities. From this point of view, the line between command and control has also blurred.

5. Fusion and C^2 System

Obviously the realization of the command and control functions is based on information available, including information about the battlefield environment, own forces and enemy forces, etc. Information gathered during military operations is not only inaccurate, ambiguous and incomplete, but also with high false alarm rate and is very possibly deceptive [1-3]. That is why more and more sophisticated sensors are developed and introduced into the combat systems. Multisensor system is expected to draw a clearer and more accurate battlefield situation picture because at least sensors can cross check with each other.

However, multisensor information processing is not so easy. With more and more modern processors in each sensor, the processing ability of each sensor grows tremendously. Almost every modern target detection sensor can detect and process multiple targets simultaneously. This multisensor multitarget information will explode if it is not coped with effectively [4]. In addition, information from different sensors often conflicts with each other in a military environment. To face these challenges, there comes the important technique, fusion.

Basically, fusion answers a number of questions, including:

1) How many targets exist in the battlefield environment?
2) What are they?
3) What are their identities? Are they friend, foe or neutral?
4) What are the most probable measurement sequences (or if possible, the motion states) of each target?

Such basic information is needed to have a clear picture of the battlefield situation and for command and control. While each sensor may have their own answers to these questions, fusion provides answers that should be more comprehensive and reliable to some degree, because they synthesizes single-sensor answers. The unit that performs fusion in a C^2 system is often known as the fusion center.

The aforementioned questions include two types of target information, positional (or kinematic) and characteristic [3]. The former is about target position and motion, such as bearing, distance, course and velocity. Characteristic information includes target type and identity. To split the target information in this way is mainly for the convenience of processing because the techniques to handle these two types of information are usually quite different [3]. Positional information fusion is often conducted along with tracking, a topic to be discussed later, where the main approaches are based on estimation and filtering [5-7]. Techniques used to deal with characteristic information include reasoning with uncertain information [4,8,9] and artificial intelligence [10].

Fusion can greatly enhance command and control ability of a C^2 system. If a multisensor system could provide the potential of such enhancement, it would be fusion that makes this potential reality. Traditionally, a C^2 system uses single-sensor information for command and control directly, even if there are many sensors serving as information providers. The valuable potential of information enhancement among the sensors is pitifully wasted. Furthermore, with sensor information not well refined and condensed, the system, not to mention its operator and commander, can be easily flooded by redundant information.

Fusion has also physically changed the C^2 systems and combat systems. There is no special processing unit between sensors and the C^2 system. Data fusion center serves as a bridge and adapter between the multisensor system and the traditional C^2 system. This physical change raises the question of where to best locate the fusion center. Should it be considered as a newly added part of the C^2 system to maintain the traditional definition that a combat system is composed of sensors, C^2 system and weapons? Or, should it be treated as an entity outside the C^2 system so that a combat system is redefined as a combination of a sensor system, fusion center, C^2 system and weapons? The former seems a more reasonable choice. First, the responsibility of the fusion center virtually is information processing, which is a basic function of a C^2 system. Secondly, this choice reserves the traditional definitions of the combat systems and C^2 systems.

Since fusion center is such a key junction, it is very possible to become an information "bottleneck." That is why its design is so important. It should be effective, reliable and flexible. It should also be well-coordinated with sensors, the rest of the C^2 system and other related units.

6. Tracking and C^2 System

Tracking is a much older concept than fusion because it is not confined to multisensor multitarget problems. Simply put, target tracking tries to find out where the target is and how it moves, by using measurements from sensors. More technically, tracking is estimating the target motion states by using estimators or filters, which really are algorithms. Typical target motion parameters or states include bearing, distance, speed and acceleration. In the old C^2 systems with single-target ability, the measurements are usually directly from sensors. Fusion is not necessary. The only possible incoming information processing is measurement preprocessing such as smoothing and outlier removal. The corresponding tracking techniques used are also primitive, based mostly on approaches with deterministic parameters.

In modern multisensor multitarget cases, fusion is a necessity. Sensor measurements are processed at first in the fusion center. The condensed and refined information then is used as input for tracking algorithms. This relationship between fusion and tracking, however, does not necessarily mean that tracking is proce-

durally behind fusion. This was the case in the early days of developing multisensor multitarget techniques. Nowadays, fusion and tracking are more and more integrated [7]. They are processed simultaneously. That is why in many cases fusion and tracking are inseparable.

Tracking sometimes may be time consuming. This surely depends on many factors, such as the measurements available, the measurement error type and size, and the algorithm used. Another important factor that is often overlooked is the requirement for the tracking solution. Obviously, it will take a longer time to get a more accurate solution. In practice, the accuracy requirement changes greatly for different tactical considerations. For example, for decision-making purposes, quite often, it cannot be afforded to wait until the solution is very accurate to make a decision. The solution may be needed at any time, no matter what accuracy the solution reaches. For fire control purposes, the accuracy requirement is relatively more stringent. But different weapons may still have different requirements. For example, a guided weapon usually does not need as stringent an accuracy requirement as straight run weapons. For this reason, there are usually several tracking algorithms in a C^2 system for the same tracking problem. For example, one is for decision making purposes and the other for fire control purposes. The former may be fast and computationally efficient. The latter should be accurate because accuracy is important in this case. Sometimes in both cases, there are more than one algorithm implemented for different considerations. Tracking is not necessarily simply a problem of algorithm. For example, it should assume some responsibilities of sensor management in some cases. In some other cases, the tracking result has a very close relationship with the maneuver pattern of the platform [11,12] and thus maneuver strategies should be recommended by the tracking algorithm.

As stated earlier, target tracking depends on the information from sensors as well as the fusion center. This means that tracking is interrelated closely with sensors and the fusion center. At the same time, the result of tracking is the basis for command and control. So tracking also has a tight connection with command and control. From these relationships tracking can be seen as the center of a C^2 system. This shows partly why advanced tracking techniques are key to a superior C^2 system. As a matter of fact, the primary goals of tracking are higher tracking accuracy and shorter tracking time. Tactically these goals are very critical to the result of a battle. From the viewpoint of system design, to reach these goals is not easy. Apart from the requirement of having advanced tracking techniques, the coordination of tracking with fusion, command and control should also be treated well.

Listed below are examples of research hot spots in target tracking in recent years:

1) Maneuvering target tracking
2) Tracking with multiple sensors
3) Tracking with uncertain measurements
4) Tracking with passive sensors.

Most tracking problems have not yet been solved well. For this reason, some alternatives for command and control in real applications have been successfully developed. For example, the emergence of smart weapons with self-homing and anti-jamming abilities has profoundly lowered the requirements for target tracking. This trend will continue in the future.

7. Conclusion

Command and control are two basic functions of a modern military combat system. Fusion and tracking are fundamental components of a command and control system. Their interrelationships are studied in this paper. Some considerations for system development are given. It should be recognized that their research and development involves not only theoretical studies, but also engineering practices. Further work is needed to keep pace with new developments.

References

1. E. L. Waltz and D. M. Buede, "Data Fusion and Decision Support for Command and Control," *IEEE T-SMC-16*, No. 6, 1986.
2. E. L. Waltz and J. Llinas, *Multisensor Data Fusion*, Artech House, 1990.
3. P. Shar and X. R. Li, "Some Considerations of Submarine Sensor Fusion," *Proc. of 1998 Int. Conf. on Information Fusion (FUSION'98)*, Vol. II, Las Vegas, Nevada, July 1998.
4. G. B. Wilson, "Some Aspects of Data Fusion," in C. J. Harris and I. White (eds.), *Advances in Command Control and Communication Systems*, Peter Peregrinus Ltd., 1987.
5. Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*, Academic Press, 1988.
6. Y. Bar-Shalom and X. R. Li, *Estimation and Tracking*, Artech House, 1993.
7. Y. Bar-Shalom and X. R. Li, *Multitarget-Multisensor Tracking*, YBS Publishing, 1995.
8. G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, 1976.
9. K. A. Klein, "A Boolean Algebra Approach to Multiple Sensor Voting Fusion," *IEEE-T-AES-29*, No. 2, 1993.

10. C. J. Harris (ed.), *Application of Artificial Intelligence to Command and Control Systems*, Peter Peregrinus Ltd., 1988.
11. P. Shar and X. R. Li, "Passive Sonar Fusion for Submarine C^2 Systems," *Proc. of 1999 Int. Conf. on Information Fusion (FUSION'99)*, Silicon Valley, CA, July 1999.
12. P. Shar and X. R. Li, "Support Systems and Techniques for Submarine Sensor Fusion," *Ibid.*

Target Recognition and Tracking based on Data Fusion of Radar and Infrared Image Sensors

Jie YANG Zheng-Gang LU Ying-Kai GUO
Institute of Image Processing & Recognition, Shanghai Jiao-Tong University, China jyang@ippr.sjtu.edu.cn

Abstract - Target recognition and tracking is a very important research area in pattern recognition. Systems for target recognition and tracking based on single sensor (radar or infrared image sensor) have their limitations. We present the approaches of target recognition and tracking based on data fusion of radar/infrared image sensors, which can make use of the complement and redundancy of data from different sensors. Data fusion at characteristic level can combine characteristics from different sensors to improve the ability of object recognition. Approaches of target recognition based on inference of rules and a neural classifier are presented to deal with the recognition of dot targets and surface targets. Data fusion at decision level can improve the reliability and anti-interference of object tracking, an approach of object tracking by on decision certainty is presented.

Keywords: Target Recognition and Tracking, Data Fusion, Pattern Recognition, Neural Networks.

1. Introduction

Target recognition and tracking is a important research area in pattern recognition. Systems with single sensor (radar or infrared image sensor) have their limitations in target recognition and tracking. For the system with radar sensor, its precision of target recognition and tracking is relatively low. For the system with a IR image sensor, its sphere of action is relatively short, it is affected by weather environment (cloud, rain, fog). A system with multi-sensors can fuse data from different sensors to overcome the limitations in the system with single sensor, it can make use of the complementary and redundancy of data from different sensors to improve the precision of target recognition and tracking. A system with multi-sensors can improve the robustness and reliability because failure of signals from a sensor will not cause failure of the whole system. So data fusion of multi-sensors become very important research direction in target recognition and tracking[3,6,7]. Different kinds of fusion models (for example, Radar-IR[5], SAR-IR, Laser radar-FLIR[1], Shipboard radar-IR) are used to realize target recognition and tracking. According to the levels of information described, the approaches of data fusion are usually divided into three classes: fusion at data level, fusion at characteristic level, fusion at decision level. Fusion at data level is usually used for fusion of images obtained from different sensors. Fusion at characteristic level is usually used for target recognition according to the characteristics derived by data from different sensors. Fusion at decision level is usually used for target tracking by jointly inferences of tracking decisions derived by data from different sensors.

In our system for target recognition and tracking, radar and infrared image sensors are used. As radar sensor in our system can provide the information of the distance and direction of the target (not the image of the target), data fusion is realized only at characteristic level and fusion at decision level. For data fusion at characteristic level, characteristics of a target obtained from radar can be used in the subsystem based on IR Image to improve the ability of object recognition; characteristics of a target obtained from IR image can be used in the subsystem based on Radar to improve the ability of object recognition. The approaches of object recognition based on inference of rules and a neural classifier are presented in Section 2 to deal with the recognition of dot targets and surface targets. For data fusion at decision level, the subsystem based on

IR Image and the subsystem based on Radar infer decisions of target tracking respectively, the decision of target tracking in the system is determined by jointly inference based on decisions of target tracking of made from the subsystem based on IR Image and the subsystem based on Radar. An approach of target tracking based on the decision certainty is presented in Section 3 to improve the reliability and anti-interference of target tracking. Following is the structure of target recognition and tracking system based on data fusion of Radar and IR image sensors.

Radar Target Recognition and Tracking By Radar

tracking decision characteristics
based on Radar detonator
from IR image
from Radar
Controller for fusion

IR Image Target Recognition and tracking decision of
tracking decision
Tracking By IR Image
based on IR image
servo-control mechanism for target tracking
fusion at characteristic level
fusion at decision level

Figure1: Target tracking system based on data fusion of Radar and IR image sensors

2. Target recognition based on the data fusion at characteristic level

For data fusion at characteristic level, characteristics of a target obtained from radar can be used in the subsystem based on IR Image to improve the ability of object recognition; characteristics of a target obtained from IR Image can be used in the subsystem based on Radar to improve the ability of object recognition. In this section, we only discuss the former situation. The process of target recognition based on IR image analysis are composed of signal pretreatment (signal detection, noise elimination), image segmentation, recognition of objects segmented. Signal pretreatment based on FFT and other technique is not discussed in this paper. For image segmentation, a IR image is transformed into binary image according to the threshold of grayness, objects in the IR image are segmented by searching the edges of objects based on the algorithm of worm tracking (see figure 2, figure 3). According to area (number of pixels) of objects segmented, the recognition of objects segmented is divided into two classes: recognition of dot targets, face targets. When the area of a object segmented is less than 3×3, the object is seen as dot target; When the area of a object segmented is equal to or greater than 3×3, the object is seen as face target. Rule-based reasoning is used to deal with the recognition of dot targets; classifier based on neural network is used to deal with the recognition of surface targets.

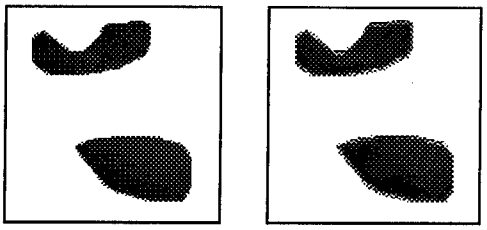


Figure 2 (left) an IR image, Figure 3 (right) Segmentation of the IR image based on worm tracking

2.1 Recognition of dot targets based on inference of rules

For a dot target, its characteristics obtained from a IR image is limited, the recognition of a dot target is mainly based on intelligent models. Intelligent models in our system are: the experimental relations between the distance of a dot target (obtained from the subsystem based on Radar) and the area of the target in the IR image; the prediction of motion direction of a

dot target; the continuity of motion path of a dot target.

For a specific IR image sensor, its waveband and resolution is defined, the possible biggest area of a target under its known distance can be estimated. Especially for a dot target (its distance is long), the experimental relations between the distance of a dot target and its area in the IR image is relatively stable. For simplification of mathematical model of the relations, only the thresholds of maximum distance R_1,R_2,R_3 need to be estimated under different areas (1, 2×2, 3×3) of a dot target. R1 means that a dot target has 1 area only if its distance less than R_1; R_2 means that a dot target has 2×2 area only if its distance less than R2; R3 means that a dot target has 3×3 area only if its distance less than R_3. So under known distance and area of a dot target, if the experimental relation is not satisfied, then the dot target is false target; if the experimental relation is satisfied, then the dot target will be recognized further.

For a true target, the direction of target motion predicated by the subsystem with Radar should be consistent with the direction of target motion predicated by the subsystem based on IR Image. Considering the complexity of the transform of space coordinate, the predication of the direction of target motion in a IR image is simplified by cross division (left up, left down, right up, right down).

According to the variation of the central position of a dot target in the sequence of two IR images, the direction of target motion can be predicated. Assume the coordinate of the IR Image is : 0 X

The central position of a target in the sequence of two IR images are: (x_1,y_1), (x_2,y_2). Y

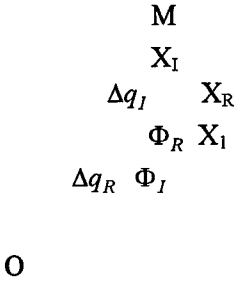
If $x_1< x_2$, $y_1< y_2$, then the prediction of target motion in the IR image is right up.

If $x_1< x_2$, $y_1> y_2$, then the prediction of target motion in the IR image is right down.

If $x_1> x_2$, $y_1< y_2$, then the prediction of target motion in the IR image is left up.

If $x_1> x_2$, $y_1> y_2$, then the prediction of target motion in the IR image is left down.

Meanwhile according to the direction of target motion obtained by the subsystem based on Radar, and the relations of angles among axes of the missile, radar and IR image sensor, the direction of target motion in the IR image can be predicated. The following is the figure about the relations of angles among axes of the missile, radar and IR image sensor, where OX_0 is the coordinate of the earth; OX_1 is the axis of the missile; OX_R is the axis of Radar; OX_I is the axis of IR Image; OM is the line of vision to the target (M is the target); Φ_R is the angle between axes of Radar and the missile; Φ_I is the angle between axes of IR Image and the missile; Δq_R is the angle between the axis of Radar and the line of vision to the target.



Assume Φ_{Rx}, Φ_{Ix}, Δq_{Rx} are respective projection of Φ_R, Φ_I, Δq_R in the horizontal direction; Φ_{Ry}, Φ_{Iy}, Δq_{Ry} are respective projection of Φ_R, Φ_I, Δq_R in the vertical direction.

If $\Phi_{Rx}+\Delta q_{Rx}<\Phi_{Ix}$, $\Phi_{Ry}+\Delta q_{Ry}>\Phi_{Iy}$, then the prediction of target motion in the IR image is right up.

If $\Phi_{Rx}+\Delta q_{Rx}<\Phi_{Ix}$, $\Phi_{Ry}+\Delta q_{Ry}<\Phi_{Iy}$, then the prediction of target motion in the IR image is right down.

If $\Phi_{Rx}+\Delta q_{Rx}>\Phi_{Ix}$, $\Phi_{Ry}+\Delta q_{Ry}>\Phi_{Iy}$, then the prediction of target motion in the IR image is left up.

If $\Phi_{Rx}+\Delta q_{Rx}>\Phi_{Ix}$, $\Phi_{Ry}+\Delta q_{Ry}<\Phi_{Iy}$, then the prediction of target motion in the IR image is left down.

If two predictions are not consistent with each other, then the dot target is a false target; If two predictions are consistent with each other, then the dot target will be recognized further.

2.2 Recognition of face targets based on a neural classifier

When the distance of a target is short, the topological shape of the target in the IR Image, the variation of position and motion direction of the target between a sequence of two IR images are strongly affected by the distance and motion direction between the target and the missile. Their mathematical relations are complicated and difficult for modeling. Because of the characteristics of self-learning, self-adaptation and fault-tolerance, neural network has been widely researched and applied[2]. Multi-layer perceptron network is used to realize a fault-tolerance classifier for recognition of face targets. IR images of face targets under different distances and directions are used to train the neural network.

Following characteristics of target are used as inputs of the neural classifier:

- distance of the target obtained from the subsystem with Radar.
- area of the target in the IR image, the variation of areas of the target in the sequence of two IR images.
- the mean grayness of pixels of the target.
- the variation of centrum positions of the target in the sequence of two IR images.
- the topological shape of the target (e.g. ratio of length to width, the number of forks in the frame extracted. 〖see figure 4, figure 5〗)
- the direction of the target motion predicated by radar and the relation of angles among the axes of the missile, Radar and IR Image.

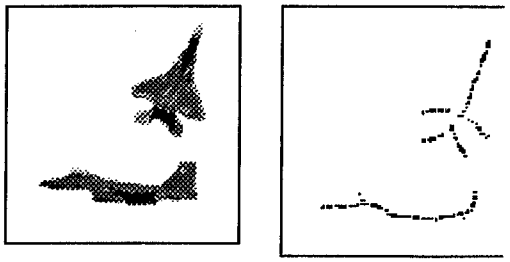


Figure 5 (left): targets in a IR image ,
Figure 6 (right): The extracted frame of the targets

Classification model for target recognition can be learned automatically by a multi-layer perceptron networks (figure 6) according to a pair of training examples. Nodes $Z_1...Z_M$ in the input layer represent the descriptors of characteristics of a target to be recognized. $O_1...O_K$ in the output layer represent the result of recognition of the target. $d_1...d_K$ represent the desired output of $O_1...O_K$; For example, d_1=1, d_2=0 represents that the target is recognized as true target; d_1=0, d_2=1 represents that the target is recognized as false target. According to Error Back-Propagation Algorithm and differences between the desired and actual neuron's response, weights of the output layer and the hidden layer $W \leftarrow W + \eta\delta_o y^t$, $\leftarrow V + \eta\delta_y z^t$ are adjusted until cumulative cycle error is less than E_{max}.

For example, the characteristics of a target obtained from the subsystem based on Radar and the subsystem based on IR Image are inputted into the neural classifier. The outputs of the classifier are: O_1=0.9, O_2=0.1, then the face target is recognized as true target and will be tracked according to the variation of its centrum positions in the IR image.

3. Target tracking based on decision fusion

After data fusion at characteristic level, a true target is recognized by the subsystem based on Radar and the subsystem based on IR Image. The subsystem based on Radar gives the decision of target tracking $\dot{q}_{Radar}$ (tracking angular velocity of the axis of Radar); the subsystem based on IR Image gives the decision of target tracking $\dot{q}_{IR}$ (tracking angular velocity of the axis of IR Image); according to these two respective decisions data fusion at decision level is to make joint decision of target tracking $\dot{q}_{fusion}$ (tracking angular velocity of the missile). According to the flight stage of the missile, data fusion at decision level is divided into three stages (initial stage, middle stage, end stage).

1) At initial stage (the distance of the target is long, IR Image sensor can not detect the target), the decision of target tracking given by the subsystem based on Radar is used to control the servo mechanism of the missile to track the target; that is, $\dot{q}_{fusion} = \dot{q}_{Radar}$, and is used to guide the servo mechanism of IR Image to track the target so that the target will be in the visual angle of IR Image sensor.

2) At end stage (the distance of the target is short, the subsystem based on IR Image can recognize and track a target independently and reliably), the decision of target tracking given by the subsystem based on IR Image is used to control the servo mechanism of the missile to track the target; that is, $\dot{q}_{fusion} = \dot{q}_{IR}$, because at end stage the decision of target tracking given by the subsystem based on IR Image is more reliable than that given by the subsystem based on Radar.

3) At middle stage (the subsystem based on IR Image can detect the target but can not recognize and track a target independently), factor "decision certainty" is introduced to realize data fusion at decision level, which represents the relatively certainty of decisions of target tracking, $0 \le CF_R \le 1, 0 \le CF_{IR} \le 1$.

Decision certainty CF_R of the subsystem based on Radar is defined as following:

$$CF_R = \alpha_R \times \Delta R \times \beta_{Radar} \times \frac{P_{R-capture}}{P_{R-falarm}}$$

where α_R is the normalizing factor, ΔR is the distance of the target, $\beta_{Radar} \in [0,1]$ is the ratio of signal to noise, $P_{R-capture}$ is the probability of capturing the target, $P_{R-falarm}$ is the probability of false alarm.

Decision certainty CF_{IR} of the subsystem based on IR Image is defined as following:

$$CF_{IR} = \alpha_{IR} \times \tau_{match} \times N_{pixel} \times \beta_{IR} \times \frac{P_{IR-capture}}{P_{IR-falarm}}$$

where α_{IR} is the normalizing factor, τ_{match} is the ratio of match obtained by the neural classifier. N_{pixel} is the number of pixels of the target in the IR image, $\beta_{IR} \in [0,1]$ is the ratio of signal to noise, $P_{IR-capture}$ is the probability of capturing the target, $P_{IR-falarm}$ is the probability of false alarm.

The joint decision of target tracking (that is, tracking angular velocity of the missile) is:

$$\dot{q}_{fusion} = \frac{CF_{IR}}{CF_{IR} + C_R} \times \dot{q}_{IR} + \frac{CF_R}{CF_{IR} + CF_R} \times \dot{q}_{Radar}$$ From the definition of CF_R, CF_{IR}, $\dot{q}_{fusion}$, following conclusions can be derived:

-- CF_R declines along with the decrease of the distance of the target; CF_{IR} ascends along with the decrease of the distance of the target. At the beginning of middle stage, q_{fusion} is mainly determined by q_{Radar}; along with the decrease of the distance of the target, the proportion of q_{Radar} in q_{fusion} decreases gradually while the proportion of q_{IR} in q_{fusion} increases gradually. So the joint decision of target tracking can realize the smooth transition from initial stage ($q_{fusion} = q_{Radar}$) to end stage ($q_{fusion} = q_{IR}$).

When the subsystem based on Radar is interfered, the joint decision of target tracking q_{fusion} is mainly determined by the decision of target tracking q_{IR} obtained by the subsystem based on IR Image; when the subsystem based on IR Image is interfered, the joint decision of target tracking q_{fusion} is mainly determined by the decision of target tracking q_{Radar} obtained by the subsystem based on Radar.

4. Conclusions

Data fusion is very important and useful for target recognition and tracking. A system with multi-sensors can fuse data from different sensors to overcome the limitations in the system with single sensor, it can make use of the complementary and redundancy of data from different sensors to improve the precision and robustness of target recognition and tracking. Data fusion at characteristic level can combine characteristics from different sensors to improve the ability of target recognition. Recognition of dot targets based on inference of rules and recognition of face targets based on a neural classifier have been presented, which simplify the modeling of target recognition and can deal with target recognition effectively. Data fusion at decision level can improve the reliability and anti-interference of object tracking, target tracking under three stages and based on decision certainty have been presented, which combines the advantages of Radar (e.g. big sphere of action) and the advantages of IR Image (e.g. high precision of target recognition and tracking when the distance of the target is short) and can realize the smooth transition of three stages.

Hardware realization [4] of our system target recognition and tracking will be discussed elsewhere.

Acknowledgment

This research is partly supported by National Science Foundation and National Defense Research Foundation of China.

Reference

[1] Beveridge J.R. and Hanson A. etc. Model-based Fusion of FLIR, Color and LADAR. SPIE Vol. 2589.

[2] Jing Z. etc. Information fusion and rracking of maneuvering targets with ANN. IEEE Intern. Conf. on Neural Networks, 1994.

[3] Libby E. W. Sequence Comparison Techniques for Multisensor Data Fusion and Target Recognition. IEEE Transactions on Aerospace and Electronic Systems. Vol. 32, No.1, Jan. 1996, pp52-64.

[4] Mattis W. Multisensor SIMD Architecture for Data Fusion Target using 21020 DSP Family.

[5] Parmar N.C, Kokar M.M.Target Detection in Fused X-band Radar and IR Images Using the Functional Minimization Approach to Data Association. 1994 IEEE InternationalSymposium on intelligent Control. 16-18, August, 1994, Columbus, Ohio, USA, pp 51-56.

[6] Shetty Shreenath and Alouani A.T. A Multisensor Tracking System With an Image-Based Maneuver Detector. IEEE Trans. on Aerospace and Electronic Systems. Vol. 32, No.1, Jan. 1996, pp167-181.

[7] Wadsworth J. Recent Advances in Track Fusion Techniques. IEE Computing & Control Divis. Colloquium on Algo for Target tracking.

Late Papers

Fuzzy Reasoning System for State Estimation and Information Fusion

P. Korpisaari and J. Saarinen
Signal Processing Laboratory
Tampere University of Technology
P.O. Box 553
33101 Tampere
Finland
Email: pkorpi@cs.tut.fi
Tel: +358 3 3653392
Fax: +358 3 3653095

Abstract ***In this paper we introduce an approach to fuzzy reasoning where negative rule weights are taken into account. Bipolar weights are applied to standard additive system (SAM). It is shown that such bipolarly weighted SAM is capable to perform reasonig over time which corresponds to Kalman filtering. Moreover, we consider a possibility to apply random fuzzy sets in such a reasoning. This enables a fusion of fuzzy sets and random variables in the context of Kalman filtering.***

Keywords: Extended fuzzy reasoning, Information fusion, State estimation

1 Introduction

Standard additive system (SAM) is a well-known fuzzy reasoning model. In such a reasoning system product is used as a T-norm, combination of multiple conclusion sets is carried out by summation and centroid is used as a defuzzification method [1]. It can be shown that SAM determines the defuzzified output as a weighted average where the averaging weights are convex coefficients. The averaging is performed over the centroids of rule patches and the averaging weights are normalized firing strengths of the set of rules in a given rulebase. In this paper we extend the conventional convex SAM to non-convex SAM where rules are weighted with a possible negative weight. Traditionally, the rule weights are positive but we apply also negative weights and consider the consequencies into reasoning system.

Ellipsoidal rule patches are commonly used in SAM-modelling. Rule patches are defined by the corresponding antecedent set, conclusion set and an ellipsoid associated to these sets. In this paper we concentrate on applying non-convex SAM as an information filter which is algebraically identical to Kalman filter. It will be shown that the estimation scheme can be represented as two non-convex SAMs. The ellipsoidal rule patches of these SAMs depend on each other. The rule patches are defined from the covariance matrices of the apriori estimate and the observation. Together these vectors form an aposteriori estimate which is a sum of two output vectors of the two above mentioned SAMs. As the covariance matrices are updated by the formula defined in the Kalman filter the two parallel reasoning systems act identically to the Kalman filter.

The prediction-correction estimation scheme is essentially a fusion process where two vectors, apriori estimate and observation, are fused together so that they produce aposteriori estimate. Thus, the above mentioned procedure can be generalized into more general fusion schemes where two pieces of information will be fused together. This kind of fusion scheme is a

non-convex adaptive SAM whose adaptation procedure is adopted from the Kalman filter. Additionally, the system is a fuzzy reasoning system which allows a use of fuzzy sets as inputs. The adaptation procedure assumes that uncertainty information related to covariance matrix is associated to each fused variable. Thus, we use random fuzzy sets for including fuzzy sets into the fusion process. This fusion process enables integrating fuzzy linquistic information with stochastic observations. Stochastic observations are modelled as conventional random sets which are a singleton-case of random fuzzy sets.

2 Standard Additive Systems

Standard additive systems (SAMs) are a special case of fuzzy reasoniong systems. SAMs are described by the following properties:

- product is used as a T-norm
- addition is used for rule combination
- centroid is used as a defuzzification scheme

The above properties define the following equation [1]:

$$y = \mathbf{F}(x) = \frac{\sum_{j=1}^{m} w_j a_j(x) V_j c_j}{\sum_{j=1}^{m} w_j a_j(x) V_j} \quad , \qquad (1)$$

where y is a defuzzified output of the reasoning system, $\mathbf{F}$ represents the reasoning system itself, m is a number of rules, w_j is a weight of the jth rule, $a_j(x)$ is the firing strength of the jth rule, V_j is a volume of the jth rule patch and c_j is a corresponding centroid of the jth rule's conclusion set.

The Eq. 1 defines the output as a weighted average of the rules' conclusion centroids:

$$\begin{aligned} y &= \sum_{j=1}^{m} p_j(x) c_j \\ p_j(x) &= \frac{w_j a_j(x) V_j}{\sum_{j=1}^{m} w_j a_j(x) V_j} \quad , \qquad (2) \end{aligned}$$

where $p_j(x)$ is a convex coefficient for the jth rule.

3 Bipolarly weighted SAM

In the SAM equation (Eq. 1) a weight w_j is assigned for each rule. The weight w_j determine rule's relative impact to the output. Rules with larger weights have bigger impact to the concluded output value y than rules with smaller weights. Thus, the weights are used for describing each rule's nature with respect to other rules. All fuzzy reasoning systems with rule weighting assume that the rules are weighted with positive weights. We introduce a *bipolarly weighted SAM* which uses both positive and negative weights. Bipolarity has been used , for example, to describe inhibitory-exhibitory natures of neurons. In [2] it is shown that it would be beneficial to apply bipolarity in fuzzy reasoning in the context of function representation abilities.

The reasoning equation of bipolarly weighted SAM is the following:

$$y = \frac{\sum_{j=1}^{m} w_j^{\pm} a_j(x) V_j c_j}{\sum_{j=1}^{m} w_j^{\pm} a_j(x) V_j} \quad , \qquad (3)$$

where $w_j^{\pm}$ denotes possibly negative rule weight for the jth rule.

4 Information granularity

Information granularity is a concept that deals with the information's representation problem [3, 4]. The same information can be represented in different levels of specifity. Information granules can be understood as pieces which are used to build up an overall understanding on the universe of discourse [5]. In particular, information granularity takes place among human beings who tend to process concepts with granules that are rougher than infinitely accurate scalar numbers [6]. We apply the concept of information granularity to illustrate the basic differencies between fuzzy sets and random sets. Fuzzy

sets concern directly the vagueness caused by the observer's way of describing the concept under consideration [7]. In contrast, random sets model the uncertainties related to stochasticity included in the process.

Fuzzy sets can be used to describe concepts. One fuzzy set is defined with a membership function whose shape and position in the universe of discourse illustrates one possible way to describe that one specific concept. Moreover, a set of fuzzy sets define alphabets with which the whole universe of discourse is handled. A key feature in the information description is a number of fuzzy sets used to fill the whole universe of discourse. An extreme case is to use singleton sets defined on the continuous basis. Thus, the number of sets is infinite and one single set describes exactly one value. It does not overlap with other fuzzy sets. On the other hand human beings are the other extreme as we may use very few fuzzy sets for dealing with such concepts as age and temperature. These fuzzy sets usually overlap each other which is a key feature of fuzziness.

As fuzziness describes the alphabets used for describing a variable, randomness describes uncertainties associated to alphabets. Such uncertainties are caused, for example, by an imperfect detection mechanism. These two different kind of uncertainties can be described with random fuzzy sets which are essentially fuzzy sets coupled with random sets [8]. Thus, each fuzzy set has its own probability distribution function. Randomness can be attached to fuzzy sets in several ways. We treat randomness as uncertainty in the position of the corresponding fuzzy set. A singleton random fuzzy set is the conventional random variable which is defined by its probability distribution. For example a singleton set $\delta(x_0)$ may be normally distributed denoted as $N(\delta_0, \Sigma)$, where Σ is the covariance matrix of the gaussian distribution function. Similarly, a fuzzy set $\mathcal{A}_i$ may be normally distributed $N(\mathcal{A}_i, \Sigma)$. Thus, each fuzzy set $\mathcal{A}_i$ has its own covariance matrix describing stochastic uncertainty associated to it.

5 Matrix inverse with bipolarly weighted SAM

A matrix inverse may be interpreted as a tool for resolving linear dependencies. Thus, it is a kind of reasoning mechanism that defines causes based on the given detections. A simple problem is to resolve values $x_1, x_2, \ldots, x_n$ in $\underline{\mathbf{x}}$ from detected $\underline{\mathbf{y}}$ based on the following relation:

$$\underline{\mathbf{y}} = A\underline{\mathbf{x}} \tag{4}$$

If solution exists, it is

$$\underline{\mathbf{x}} = A^{-1}\underline{\mathbf{y}} \tag{5}$$

The Cramer's rule [9] defines the inverse of a square matrix as follows:

$$A^{-1} = \frac{adj(A)}{det(A)}\,, \tag{6}$$

where $adj(A)$ is an adjoint matrix of A and $det(A)$ is a determinant of A. Both the adjoint matrix and the determinant are defined by A^{ij}:s which are the minors of the matrix A. ijth minor of matrix A is a determinant of the submatrix , denoted as A_{ij} which is obtained from A when ith row and jth column of A are removed. The definitions of $adj(A)$ and $det(A)$ yield the following formula for x_i, the ith component of $\underline{\mathbf{x}}$:

$$x_i = \frac{\sum_{j=1}^{n}(-1)^{i+j}A^{ji}y_j}{\sum_{j=1}^{n}(-1)^{i+j}A^{ji}a_{ji}} \tag{7}$$

Assume a n-dimensional set whose base is $(n\text{-}1)$-dimensional polygon and height in the nth dimension is h. Assume further that the base polygon is defined by $n\text{-}1$ row vectors of the submatrix A_{ij} and the height is equal to element a_{ij}. By definition, the $(n\text{-}1)$-dimensional volume of the base polygon is equal to absolute value of A^{ij} which is the determinant of the matrix A_{ij}. Thus, the volume of this hyperpolygon denoted as $V_{ij}^{\diamond}$ is

$$V_{ij}^{\diamond} = |A^{ij}a_{ij}| \tag{8}$$

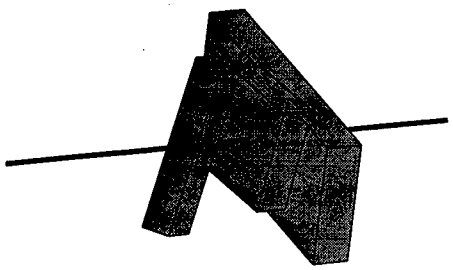


Figure 1: Polygons as THEN-part sets.

Eq.(7) can be formulated by using the volume $V_{ij}^{\diamond}$ as follows:

$$x_i = \frac{\sum_{j=1}^{m} w_{ij}^{\pm} V_{ji}^{\diamond} \frac{y_j}{a_{ij}}}{\sum_{j=1}^{m} w_{ij}^{\pm} V_{ji}^{\diamond}} \quad . \tag{9}$$

This is equal to bipolarly weighted SAM (Eq. (3)) with the following settings:

$$\begin{aligned} w_j^{\pm} &\simeq w_{ij}^{\pm} = (-1)^{i+j} sign(V_{ij^\circ}) \\ V_j &\simeq V_{ji}^{\diamond} \\ a_j(\underline{\mathbf{y}}) &\simeq 1 \\ c_j &\simeq \frac{y_j}{a_{ij}} \end{aligned}$$

$w_{ij}^{\pm} = (-1)^{i+j} sign(V_{ij^\circ})$ defines the impact of the jth rule on the resulting x_i in a supporting/not-supporting or exhibitory/inhibitory sense. $V_{ij}^{\diamond}$ is a volume of the THEN-part sets, which are graphically illustrated in Fig. 1 for one dimension. $a_j(\underline{\mathbf{y}})$ is a membership value of the IF-part fuzzy set. It is natural that this value is equal to one since nothing fuzzy is considered so far. The inverse equation is formulated for crisp numbers only. Finally, c_j is a linear function of y_j : $c_j = \frac{y_j}{a_{ij}}$. This corresponds to Sugeno model which is one form of generalized SAMs described in [1].

6 Kalman filter with fuzzy observations

An information filter [10, 11] is an alternative representation of Kalman filter algorithm. These two algorithms are algebraically equivalent although the formula and representations are very different. The fusion of two vectors into one vector is carried out as follows:

$$\underline{\mathbf{x}} = (Y_{\underline{\mathbf{u}}} + Y_{\underline{\mathbf{v}}})^{-1}(Y_{\underline{\mathbf{u}}}\underline{\mathbf{u}} + Y_{\underline{\mathbf{v}}}\underline{\mathbf{v}}) \quad , \tag{10}$$

where Y is an information matrix of the variable denoted in the subscript. The information matrix is an inverse of the variable's covariance matrix. An intuitive interpretation of the above equation is that it is an weighted average of the vectors where the weights are information ellipsoids. The more certain variable is the bigger is the weighting ellipsoid since the inverse of the covariance matrix comes larger. This principle is very similar to basic idea of the SAM reasoning method. However, this kind of interpretation is not strictly valid for matrices.

The matrix calculations in the above equation can be manipulated in the following way:

$$\begin{aligned} (Y_{\underline{\mathbf{u}}} + Y_{\underline{\mathbf{v}}})^{-1} Y_{\underline{\mathbf{u}}} &= (Y_{\underline{\mathbf{u}}} + Y_{\underline{\mathbf{v}}})^{-1}(Y_{\underline{\mathbf{u}}}^{-1})^{-1} \\ &= (Y_{\underline{\mathbf{u}}}^{-1}(Y_{\underline{\mathbf{u}}} + Y_{\underline{\mathbf{v}}}))^{-1} \\ &= (I + Y_{\underline{\mathbf{u}}}^{-1} Y_{\underline{\mathbf{v}}})^{-1} \end{aligned} \tag{11}$$

Using the following matrices:

$$\begin{aligned} U &= I + Y_{\underline{\mathbf{u}}}^{-1} Y_{\underline{\mathbf{v}}} \\ V &= I + Y_{\underline{\mathbf{v}}}^{-1} Y_{\underline{\mathbf{u}}} \end{aligned}$$

and applying Cramer's rule yields:

$$\begin{aligned} \underline{\mathbf{x}} &= U^{-1}\underline{\mathbf{u}} + V^{-1}\underline{\mathbf{v}} \\ &= \frac{1}{|U|}\mathtt{adj}(U)\underline{\mathbf{u}} + \frac{1}{|V|}\mathtt{adj}(V)\underline{\mathbf{v}} \end{aligned} \tag{12}$$

where $|U|$ denotes a determinant of U and $\mathtt{adj}(U)$ is U's adjoint matrix. Definition of the adjoint matrix [9] yields the following form for ith component (i is odd) of $\underline{\mathbf{x}}$:

$$\begin{aligned} x_i &= \frac{1}{|U|}[U_{1i}, -U_{2i}, \ldots, (-1)^n U_{ni}]\,[u_1, u_2, \ldots, u_n]^T \\ &+ \frac{1}{|V|}[V_{1i}, -V_{2i}, \ldots, (-1)^n V_{ni}]\,[v_1, v_2, \ldots, v_n]^T \end{aligned}$$

and for even i:

$$\begin{aligned} x_i &= \frac{1}{|U|}[-U_{1i}, U_{2i}, \ldots, (-1)^{n+1} U_{ni}]\,[u_1, u_2, \ldots, u_n]^T \\ &+ \frac{1}{|V|}[-V_{1i}, V_{2i}, \ldots, (-1)^{n+1} V_{ni}]\,[v_1, v_2, \ldots, v_n]^T \end{aligned}$$

Thus the general formula for x_i is the following:

$$x_i = \frac{1}{|U|}\sum_{j=1}^{n}(-1)^{i+j}U_{ji}u_j + \frac{1}{|V|}\sum_{j=1}^{n}(-1)^{i+j}V_{ji}v_j \tag{13}$$

This formula is of the same form as the bipolarly weighted SAM. Thus, summarizing the above formula: Information filter, and thus Kalman filter, can be represented as a sum of two bipolarly weighted SAMs as it is illustrated in Fig. 2.

With fuzzy logic concepts the above system is a combination of two parallel adaptive fuzzy reasoning systems. The adaptivity means in this case that the covariance matrices that directly defines the rulebase of these systems are tuned by the means of information filter's updating equations.

The above reasoning formula can be understood as a bipolarly weighted SAM whose IF-part sets are singletons. This implies the fact that the rule firing strength is always one since the input fuzzy sets are also singletons. However, this is natural since only crisp values are considered in the above reasoning system that performs the state estimation task.

Given a random fuzzy set $\mathcal{A}$ with covariance matrix Σ it is straightforward to apply the bipolarly weighted SAM formula which allows rule firing strengths between 0 and 1. Such an equation is given in Eq. 3.

7 Conclusions

We present a bipolarly weighted SAM which is a fuzzy reasoning system with possibly negative rule weights representing the inhibitory role of the rule. It is shown that such a system can be used for fusing fuzzy information to stochastic estimation processes. This approach enables adding human expert knowledge or human observations into estimation process of stochastic random variables.

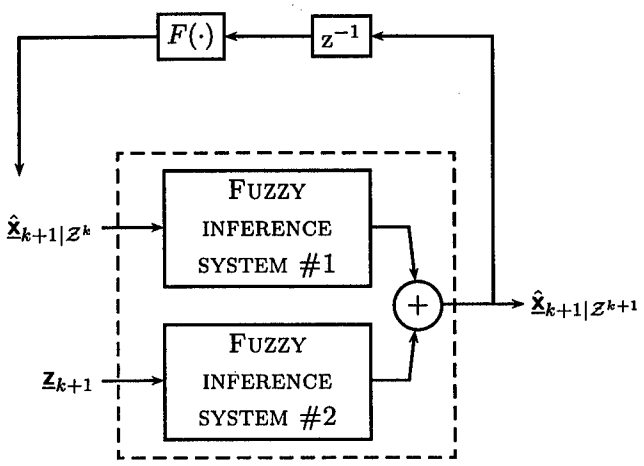


Figure 2: The state estimation process with two parallel fuzzy reasoning systems.

References

[1] B. Kosko, *Fuzzy Engineering.* New Jersey: Prentice-Hall, 1997.

[2] F. Watkins, "The representation problem for additive fuzzy systems," in *Proceedings of the IEEE FUZZ-95*, vol. 1, pp. 117–122, 1995.

[3] M. Laviolette and J. Seaman, "The efficacy of fuzzy reprersentations of uncertainty," *IEEE Transactions on fuzzy systems*, vol. 2, pp. 4–15, February 1994.

[4] H. Nguyen, "Fuzzy sets and probability," *Fuzzy Sets and Systems*, vol. 90, no. 2, pp. 129–132, 1997.

[5] L. Zadeh, "Fuzzy logic = computing with words," *IEEE Transactions on Fuzzy Systems*, vol. 4, pp. 103–111, May 1996.

[6] L. Zadeh, "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic," *Fuzzy Sets and Systems*, vol. 90, no. 2, pp. 111–127, 1997.

[7] D. Dubois and H. Prade, "The three semantics of fuzzy sets," *Fuzzy Sets and Systems*, vol. 90, no. 2, pp. 141–150, 1997.

[8] R.-J. Chao and B. Ayyub, "Distributions with fuzziness and randomness," in *Proceedings of ISUMA-NAFIPS '95*, pp. 668–673, 1995.

[9] E. Kreyszig, *Advanced Engineering Mathematics.* John Wiley & Sons, Inc., 1993.

[10] C.-B. Chang and J. Tabaczynski, "Application of state estimation to target tracking," *IEEE Transactions on Automatic Control*, vol. 29, no. 2, pp. 98–109, 1984.

[11] R. Mehra, "Approaches to adaptive filtering," *IEEE Transactions on Automatic Control*, vol. 17, pp. 693–698, October 1972.

Uncertain Reasoning for Adaptive Object Recognition

Sung Wook Baik and Peter Pachowicz
Department of Electrical and Computer Engineering
George Mason University
Fairfax, VA, U.S.A.

Abstract *This paper introduces an uncertain reasoning approach for adaptive object recognition under changing perceptual conditions. The uncertain reasoning is carried out for the fusion of model-based segmentation and data-driven segmentation of an image obtained under a new condition. The model-based segmentation is achieved by the RBF-based classifier and the data-driven segmentation is based on a boundary melting algorithm. The paper presents examples of segmentation results of two segmentations (i.e. model-based and data-driven), and an uncertain reasoning approach applied to the fusion of the results.*

Keywords: fusion, model-based segmentation, data-driven segmentation, uncertain reasoning, Dampster-Shaper theory

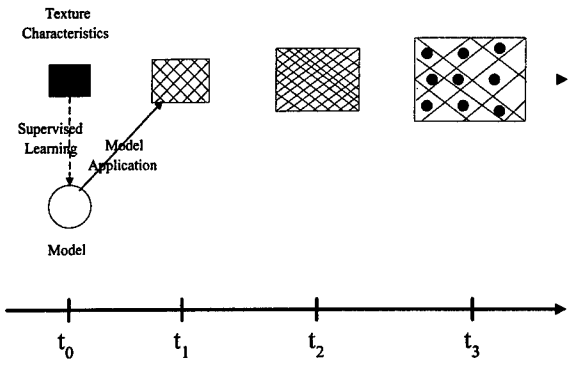


Figure 1: Problem Definition.

1 Introduction

We focus on a specific problem, which the resolutions of object surfaces change when an observer approaches object scenes gradually. We limit the experimental work to the texture recognition problem where the texture characteristics change significantly, but smoothly, with the change in perceptual conditions. It means that some detailed and visible information is brought over the increasing resolution. There are uniformed textures under low resolution whereas detailed and visible textures are coming up under high resolution.

Object models learn from the first image in a supervised way. In the typical research, the object models are applied to recognize the same image under the given condition. However, a challenging problem is how to recognize objects on forthcoming images of sequence through gradual model adaptation (model modification) to these new images under changing perceptual conditions. Figure 1 illustrates the sequence of texture images with different resolutions under changing perceptual conditions. When these models are applied to recognize the next image, it is not expected that they be well matched with the next image. It is so because direct matching results are just feedback obtained from the recognition of next image. The feedback is very useful to update the model, only if it is analyzed appropriately in an unsupervised way (without any human help). This paper presents an approach for analysis of the feedback information by using a fusion technique to integrate data-driven and model-based segmentation paradigms.

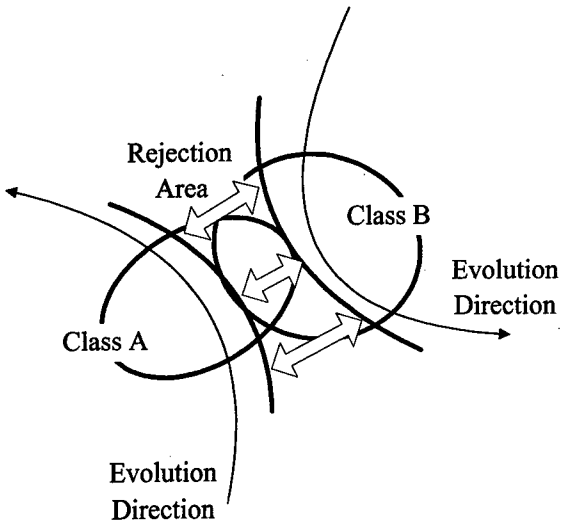


Figure 2: Overlapping problem and rejection concept.

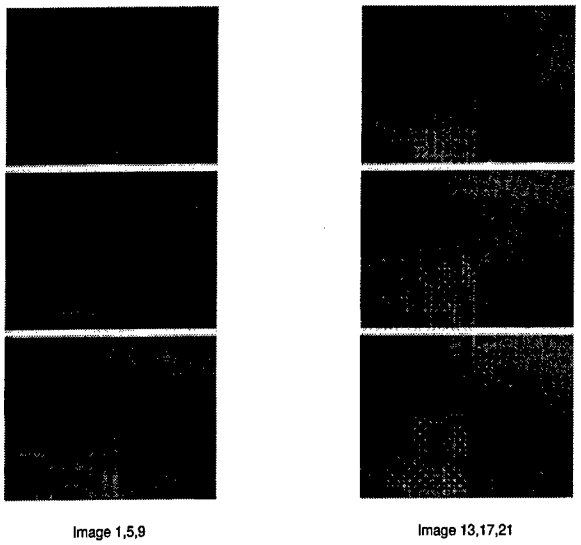


Figure 3: Six selected images from a sequence of 22 images.

2 Problems in Model Based Segmentation

This section presents a model-based segmentation. It also supports the necessity of fusion technique to integrate data-driven and model-based segmentation paradigms by presenting some problems happening when only a model-based segmentation is applied to analyze feedback information for efficient model modification.

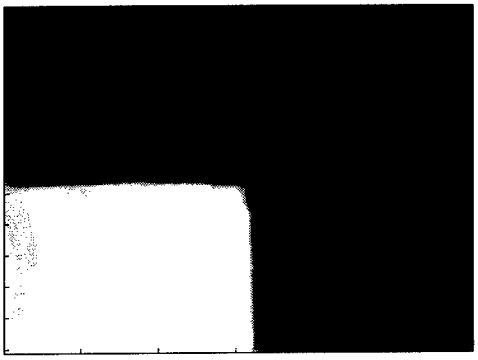
Object models are applied to a new incoming image using an RBF based classifier [1]. Recognition and segmentation processes are applied to separate the image into homogeneous areas. An image is segmented by classifying and grouping all pixels within the image to one of several classes according to the current model. The result is an annotated image with confidences supporting these annotations. The image contains class labels. Associated confidences are measurements reflecting classification strength. When classification confidences of a group of pixels are low, these do not pertain to any real pattern class. Instead, they are in an imaginary (background) class that indicates a rejection.

The results of image recognition and segmentation may be confused when an image is recognized by models associated with the previous one due to overlapping problem of object models for two or more classes. The results of image recognition and segmentation may not be satisfactorily good when an image is recognized by models associated with the previous one. It is so because these two consecutive images are always slightly different or objects change their appearance. Therefore, if only a model based segmentation technique is applied to feedback analysis, model modification may be performed incorrectly.

To resolve the difficulty, the basic and essential approach [2] is to use the rejection concept of object classification. The usage of the rejection concept can reduce the risk of using incorrect feedback information by avoiding the analysis of image areas of uncertain situation. Figure 2 presents an overlapping problem (uncertain situation) and a rejection concept for the overlapping problem. However, when the overlapping problem is more serious, model modification is impossible since most of classification results are rejected. Therefore, this paper presents an application of hybrid methods for image segmentation, which integrate supervised and unsupervised paradigms.

Figure 3 illustrates 6 selected images from a sequence of 22 images acquired by a digital b&w camera. Images were registered with the change in distance between the camera and the scene. The distance was gradually decreased. The next image was registered under the distance reduced by 4%. Total of 22 b&w images were obtained from the same scene. The size of each image is 240×320 pixels with a single pixel coded on 256 gray levels. Each image contains four classes of texture areas corresponding to different fabrics. Figure 4 shows an example of image segmentation results when these models are applied to the next image. It is seen that applying models from previous images to the next images, without modifying them to reflect the changes, causes significant quality degradation of image segmentation. Figure 4(a) shows class annotations and Figure 4(b) shows confidence values. In Figure 4(a), dark gray, medium gray, light gray, and white colors represent class A, class B, class C and class D, respectively. Black color represents a rejection class. In Figure 4(b), white areas represent high confidence whereas dark areas represent low confidence. In particular, pixels belonging to the rejection class have very low confidences.

(a) Class membership image



(b) Confidence level image

Figure 4: Image segmentation results when models learned from previous image are applied to the next image.

3 Unsupervised Segmentation

Texture images are not appropriate for data-driven segmentation. This is so because pixels with similar intensity values within a texture image are scattered in regular patterns rather than they form homogeneous areas by getting together within some boundaries. To resolve this problem, appropriate filters (eg. Gabor filter set) are applied to texture images. Such filtering helps to convert similar texture patterns of an original image to similar feature values of a feature image, which is obtained through filtering with one of filters. Pixels with similar feature values are grouped through data-driven segmentation.

Gabor filters [3] are useful to deal with texture images characterized by local spatial frequency and orientation information present in an image. Gabor filters are obtained through a systematic mathematical approach. They are normally used for image decomposition and are frequency-related. A Gabor function consists of a sinusoidal plane of particular frequency and orientation modulated by a two-dimensional Gaussian envelope.

Boundary melting algorithm [5] is employed for data-driven segmentation. It helps to derive accurate separation lines between regions and to merge less significant regions. Boundary melting is more suitable for our application because it is using boundary finding techniques based on a local gradient operation rather than region statistics. Gradient operation simply guarantees that area boundaries are preserved

Figure 5: Feature image of image 8 obtained by using a Gabor filter.

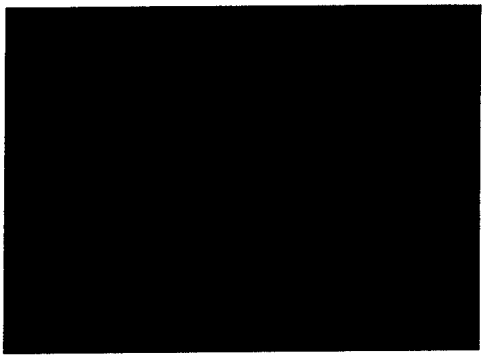


Figure 6: The segmentation result of image 8 obtained by using a data-driven segmentation technique.

rather than smoothed. Figure 5 and 6 presents a feature image and a result of data-driven segmentation for it.

4 Uncertain Reasoning for Fusion

An uncertain reasoning approach is applied to the fusion of the results from two segmentations (i.e. model-based and data-driven). The architecture of the hybrid approach for the fusion is presented in Figure 7. The results from the two segmentations are aligned on top of each other. The region boundaries of the "fused" image are the same as the ones in the data-driven segmentation, since this type of segmentation preserves precise boundaries. Within precise boundaries, the region annotations are determined by consideration (uncertain reasoning) of class memberships and classification confidences, of the pixels in those regions, which are determined from the model-based segmented image. The region within the boundary is regarded as a homogeneous region, which is annotated by a same class. However, some of such regions may be annotated by two or more different classes. Therefore, it is said that the identities of pixels in the region are still unknown, and class memberships and confidences of pixels in such regions are called uncertain data. The goal of the fusion is to explore a representative class membership of a homogeneous region. For class exploration, it is first required that the uncertain data should be formalized into evidential forms (evidence measurement). The class exploration is achieved by an inference of the results obtained by combining the formalized evidences (Evidence Combination/Inference). The two processes are called *uncertain reasoning.*

- Evidence measurement - To formalize a classification confidence into a piece of evidence. The evidence is a degree to support a certain hypothesis that a pixel (or an example) within a certain region should be annotated by a specific class. Pascalian gradation of the force of evidence is employed to measure the formulized evidence [4].

- Evidence Combination/Inference - To combine efficiently a couple of partial and auxiliary evidences and to verify the hypothesis according to the results through evidence combination. Dampster-Shaper theory is applied for the evidence combination and inference [4].

5 Conclusions

In this paper, we introduced a fusion approach by using an uncertain reasoning for adaptive

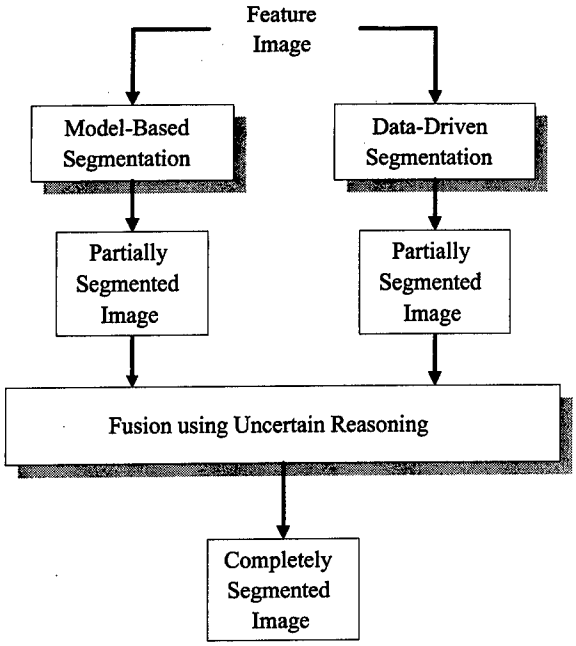


Figure 7: The architecture of the hybrid approach for the fusion.

object recognition. The fusion is hybrid approach to integrate two segmentation techniques with probability reasoning techniques. The hybrid approach complements the feedback analysis for the model modification by adding the data-driven segmentation technique to cover weaknesses of the model-based segmentation technique. It is promised that recognition results of objects under changing perceptual conditions will be better by the object models modified after feedback analysis with the cooperation of the hybrid fusion approach.

References

[1] Baik, S. and Pachowicz, P. Adaptive Object Recognition based on the Radial Basis Function Paradigm. *Proceedings of 1999 International Joint Conference on Neural Networks (IJCNN '99).*, Washington, DC, July 10-16, 1999.

[2] Fairhurst, Michael C. *Computer Vision for Robotic Systems.* Prentice Hall, 1988.

[3] Farrokhnia, Farshid. *Multi-channel filtering techniques for texture segmentation and surface quality inspection.* Ph.D. thesis, Dept. of Electrical Eng., Michigan State University, 1990.

[4] Schum, David A. *Evidential Foundations of Probabilistic Reasoning.* Wiley Inter-Science, 1994.

[5] Sonka, M., Hlavac, V. and Boyle, R. *Image Processing, Analysis and Machine Vision.* Chapman Hall, 1993.

Knowledge Fusion in the Large – taking a cue from the brain

Lokendra Shastri and Carter Wendelken
International Computer Science Institute
1947 Center Street, Suite 600
Berkeley, CA 94704

Abstract *Even the most commonplace cognitive behaviors such as vision and language understanding involve large-scale fusion of disparate pieces of evidence. Therefore, our capacity to rapidly and effortlessly produce coherent interpretations of visual and verbal inputs points to the remarkable ability of the human mind/brain to fuse evidence. We will discuss a neurally motivated computational model that attempts to replicate some of this remarkable ability, and illustrate the functioning of the model with the help of a few examples.*

Keywords: Neural Networks, Inference, Evidence combination, Coherence

1 Introduction

Even commonplace cognitive behaviors such as vision and language understanding involve large-scale fusion of evidence from disparate sources. Consider the task of understanding language wherein squiggles on a surface or fluctuations in air-pressure are mapped by the reader or hearer into coherent mental descriptions of events and situations. The process underlying this mapping appears to be remarkably complex. It involves, among other things, recognizing words, disambiguating word senses, incorporating grammatical constraints, and carrying out inferences based on fuzzy and partial knowledge to establish causal and referential coherence.[1]

Any system that attempts to explain our ability to establish causal and referential coherence during language understanding must possess a number of properties. First, such a system must be representationally adequate. It must be capable of encoding specific facts as well as general regularities (aka rules) that capture the causal structure of the environment. In particular, the system should be capable of encoding context-dependent and evidential cause-effect relationships. Second, the system should be inferentially adequate, that is, it should be capable of drawing a wide range of explanatory and predictive inferences. In doing so, the system should be able to *combine evidence* provided by disparate sources and arriving at *coherent* and mutually reinforcing interpretations. Third, the system should be capable of establishing referential coherence. In particular, it should be able to posit the existence of entities that may be only implicit in the "input" ("John bought a book" implies the existence of an entity that sold the book to John) and unify (or merge) entities by recognizing that two entities referred to in a discourse may be one and the same. Fourth, the system should be capable of learning and fine-tuning its causal model based on experience, instruction, and exploration. Finally, the system should be scalable and computationally effective.

In this paper we describe a neurally motivated system that exhibits — at least to a certain extent — the properties enumerated above. This system is an extension of SHRUTI [11]. It can express causal knowledge involving n-place relations, limited quantification, and type restrictions. It can encode specific events as well as context-sensitive priors over events. It expresses dynamic bindings via the synchronous firing of appropriate node clusters and performs inferences via the propagation of rhythmic activity over node clusters. This propagation amounts to a parallel breadth first activation of the underlying causal graph, and hence, the reasoning in SHRUTI is extremely fast. The use of weighted links and activation combination functions at nodes allow SHRUTI to encode soft rules and perform evidential inference. SHRUTI supports supervised learning which allows it to fine-tune its causal model in a data-driven manner. Moreover, SHRUTI supports short-term associative learning which allows it to dynamically favor stable coalitions of activity. The latter plays a critical role in establishing coherence. In this paper we focus on the ability of SHRUTI to (i) rapidly establish causal and referential coherence and (ii) combine evidence in a flexible and context-dependent manner using a family

[1] Causal coherence refers to the establishment of causal relationships among various events mentioned in a discourse. Referential coherence involves keeping track of entities referenced in a discourse and determining which entities are the same. It is well known that inferences required to establish causal and referential coherence occur rapidly and automatically during text understanding (see e.g., [6, 7, 5]). The evidence for the automatic occurrence of predictive inferences is mixed, but their occurrence cannot be ruled out [9].

of evidence combination functions. Other details may be found in [11, 10].

2 Representational Overview

To motivate and concretize the description of SHRUTI's behavior consider the following narrative: "(S1) John fell in the hallway. (S2) Tom had cleaned it. (S3) He got hurt." Upon being presented with the above narrative (we will see how below) SHRUTI rapidly infers the following: (i) Tom had cleaned the hallway, (ii) The hallway floor was wet, (iii) John was walking in the hallway, (iv) John slipped and fell because the floor was wet, and (v) John got hurt because he fell. Notice that SHRUTI draws several inferences required to establish referential and causal coherence. It explains John's fall by making the plausible inference that John was walking in the hallway and he slipped because the floor was wet. It also infers that John got hurt because of the fall. Moreover, it determines that "it" in (S2) refers to the hallway, and that "He" in (S3) refers to John, and not to Tom. SHRUTI draws these inferences based on commonsense knowledge such as that shown in Figure 1, as well as several additional commonsense rules and facts about cleaning, wet floors, and being hurt.

2.1 Interplay of structure and dynamics

A description of SHRUTI requires the specification of its *structure* as well as a description of its *dynamic* behavior. All long-term (persistent) knowledge is encoded in SHRUTI via structured networks of nodes and links. The dynamic aspects of SHRUTI involve the encoding and propagation of dynamic bindings via synchronous activity, the activation of long-term facts in response to dynamic bindings, evidence combination, the dynamic instantiation and unification of entities, the short-term increase (potentiation) of weights due to convergent activity, and the emergence of coherence in the form of reverberant activity along closed loops.

2.1.1 Encoding Relations Using Focal Clusters

Each relation is represented by a focal cluster depicted by a dotted ellipse in Figure 1. Consider the focal cluster for *slip*. This cluster includes an enabler node labeled *?:slip*, two collector nodes labeled *+:slip* and *-:slip*, and two role nodes labeled *slip-pat* and *slip-loc* for its two roles *patient* and *location*. In general, the cluster for an n-place relation will contain n role nodes, with the synchronized activity of each indicating a particular role binding (as described below).

Activity in *?:slip* indicates the strength with which information about the particular instance of the *slip* relation is sought. The activation levels of the collectors *+:slip* and *-:slip* encode a graded belief ranging continuously from *no* on the one extreme (only *-:slip* is active), to *yes* on the other (only *+:slip* is active), and *don't know* in between (neither collector is very active). If both the collectors receive comparable and strong activation then both collectors can be active, despite mutual inhibition. This signals a contradiction.

Links from the collector nodes to the enabler node of a relation convert a dynamic assertion of a relational instance into a query about the assertion. Thus the system continually seeks an explanation for active assertions. The weight on the link from *+:slip* (*-:slip*) to *?:slip* is proportional to the system's propensity for seeking explanations and inversely proportional to the probability of occurrence of a positive (negative) instance of *slip*.

Nodes are computational abstractions and correspond to *small ensembles of cells*, and a connection between nodes corresponds to several connections from cells in one ensemble to cells in the other. Phasic nodes, of which role nodes are an example, respond only to synchronous activity and fire only in synchrony with their inputs. Enabler and collector nodes, however, can integrate activity over a broader time window (see [10]).

The *dynamic* encoding of a relational instance corresponds to a *rhythmic* pattern of activity wherein bindings between roles and entities are represented by the *synchronous* firing of appropriate role and entity nodes [12, 11] With reference to Figure 1, the dynamic representation of the relational instance *(fall: ⟨fall-pat=John⟩, ⟨fall-loc=Hallway⟩)* (i.e., "John fell in the Hallway") will involve the synchronous firing of *+:John* and *fall-pat*, and the synchronous firing of *+:Hallway* and *fall-loc*. The entities *+:John* and *+:Hallway* will fire in distinct phases.

SHRUTI encodes two types of facts in its long-term memory: episodic facts (E-Facts) and taxon facts (T-facts). These facts provide closure between the enabler node and the collector nodes. While an E-fact corresponds to a specific instance of a relation, a T-fact corresponds to a distillation or statistical summary of various instances of a relation and can be viewed as coding *prior probabilities*. T-facts can be conditioned on the type of role-fillers (e.g., the T-fact *buy(Person,Car)* encodes how likely it is that a person would buy a car).

2.1.2 Encoding of Types and Instances

This is illustrated at the right of Figure 1. The focal cluster of each entity, A consists of a *?:A* and a *+:A* node. In contrast, the focal cluster of each type, T consists of a pair of *?* (*?e:T* and *?v:T*) and a pair of + nodes (*+e:T* and *+v:T*). The pair of v nodes and the pair of e nodes signify universal and existential quantification,

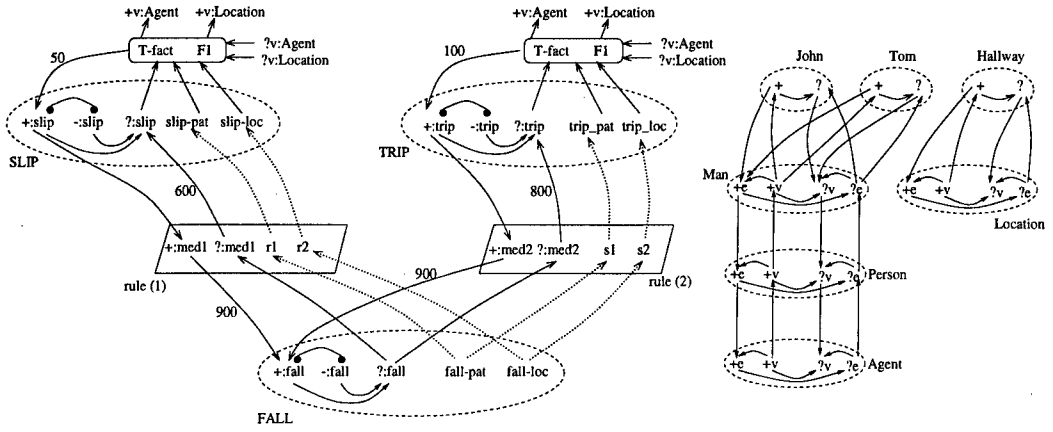


Figure 1: An example SHRUTI network encoding a subset of the knowledge base for the "John fell" example.

respectively. The activation *levels* of *?:A, ?v:T, and ?e:T* nodes signify the strength with which information about entity A, the type T, and an instance of type T, respectively, is being sought. Similarly, the activation *levels* of $+:A$, $+v:T$, and $+e:T$ signify the degree of belief that A, T, and an instance of type T, respectively, play appropriate roles in the current situation.

2.1.3 Encoding of Rules

A rule is encoded via a mediator focal cluster (shown as a parallelogram) that mediates the flow of activity between the antecedent and the consequent clusters. The mediator consists of a collector and an enabler node and as many role-instantiation nodes as there are distinct variables in the rule. The enablers of the consequent relations are connected to the enablers of the antecedent relations via the enabler of the mediator. The appropriate (+/–) collectors of the antecedent relations are linked to the appropriate (+/–) collectors of the consequent relations via the collector of the mediator. Each of these enabler and collector links for a rule has a weight which can be specified by a knowledge engineer and/or learned via supervised learning. The roles of the consequent relations are linked to the roles of the antecedent relations via appropriate role-instantiation nodes in the mediator. This linking reflects the correspondence between antecedent and consequent roles specified by the rule.

If a role-instantiation node receives activation from one or more consequent role nodes, it simply propagates the activity onward to the connected antecedent role nodes. If on the other hand, it receives activity only from the mediator enabler, it sends activity to the *?:e* node of the type specified in the rule as the type restriction for this role. This causes the *?:e* node of this type to become active in an unoccupied phase. The *?:e* node of the type conveys activity in this phase to the role-instantiation node which in turn propagates this activity to connected antecedent role nodes. This interaction between the mediator and the type hierarchy, in effect, creates activity corresponding to "Does there exist some role filler of the specified type?"

2.1.4 Mutual Exclusion and Collapsing of Phases

Entities in the type hierarchy can be part of a *phase-level* mutual exclusion cluster (ρ-mex cluster). The + node of every entity in a ρ-mex cluster has inhibitory links to and from the + node of all other entities in the cluster. As a result of the mutual inhibition, only the most active entity within a ρ-mex cluster can remain active in any given phase. A similar ρ-mex cluster can be formed by mutually exclusive types. Mutual exclusion also occurs in the type hierarchy as a result of inhibitory connections from the + nodes of a type (or an entity) to the ? nodes of all its siblings. This inhibition leads to an "explaining away" phenomenon. If for example, the type query "Is it a Person?" (i.e., activation of *?e:person*) leads to the queries "Is it a Man?" and "Is it a Woman?", then strong support received by *+e:Woman* reduces the strength of the query *?e:Man*. In essence, the query "Is it a Man?" is no longer considered important by the system since it was seeking a person and it has already found a woman.

SHRUTI allows separate phases to coalesce into a single phase, or new phases to emerge, as a result of inference. The latter is realized by the allocation of new phases resulting from the interaction between role-instantiation nodes in mediators and the type hierarchy. The unification of phases is realized in the current implementation by collapsing of phases based on activity within an entity cluster or within a focal cluster. In the first case, phase collapsing occurs whenever a single entity dominates multiple phases (for example if the same entity comes to be the answer to multiple queries). In

the second case, phase collapse occurs if two unifiable instantiations of a relation arise within a focal cluster. For example, an assertion *+:fall(John,Hallway)* alongside the query ∃ *x:man ?:fall(x,Hallway)* (Did a man fall in the Hallway) will result in a merging of the two phases for "a man" and "John".

> SHRUTI's ability to flexibly instantiate entities and collapse them into a single entity during inference is due to its use of temporal synchrony to represent dynamic bindings.

2.1.5 Short-term Potentiation

If *?:P*, the enabler node of *P* receives activity from a mediator enabler node and concurrent activity from one of *P*'s collector nodes, then the weight of the link from the mediator enabler to *?:P* increases for a short-duration.[2] With reference to the "John fell" example, this increase in weight has the following functional significance (refer to Figure 1): The activation arriving at *?:slip* from *?:med1* means that "John slipped" is being sought as a possible explanation of "John fell". The concurrent arrival of activity from *+:slip* would mean that at this very time "John slipped" is also being asserted. Under these circumstances, it is highly likely that "John slipped" may indeed be the explanation of "John fell". The increase in weight of the link from *?:med1* to *?:slip* marks *slip* as a more likely explanation for *fall* under the existing circumstances.

If *+:P* (*-:P*) receives activity from one of its T-facts and concurrent activity from a mediator collector node, then the weights of the links from the mediator collector to *+:P* (*-:P*) and from the active T-facts to *+:P* (*-:P*) increase for a short-duration. With reference to the "John fell" example, this increase in weights has the following functional significance (refer to Figure 1): The activation arriving from *+:med1* at *+:fall* means that "John fell" is being predicted as a possible consequence of "John slipped". The concurrent arrival of activity at *+:fall* from *?:fall* (via a T-fact) would mean that at this very time "John fell" is also being sought as a possible explanation of some event (this is why *?:fall* is active). Under these circumstances, it is highly likely that the event "John fell" actually occurred and is an effect of "John slipped". The increase in weight of the link from *?:med1* to *+:fall* marks *fall* as a more likely effect of *slip* under the existing circumstances.

2.1.6 "Explaining away" in the Causal Model

A "explaining away" phenomena also occurs in the causal model as a result of inhibitory connections between rules which share the same consequent. For the structure shown in Figure 1, there is for example an inhibitory link (not shown) from *+:med1* to the link from *?:fall* to *?:med2*. As a result, a strong activation of *+:slip* reduces the activation flowing from *?:fall* into *?:trip*. In essence, if the system is seeking an explanation for fall, then a strong belief in slipping is taken to be a sufficient explanation of falling, and hence, the search for tripping acquires lesser significance.[3]

> Taken together, the short-term associative increase in weights and the inhibitory interactions leading to the explaining away phenomena, provide a powerful and neurally plausible mechanism that enable SHRUTI to prefer coherent explanations over non-coherent ones.

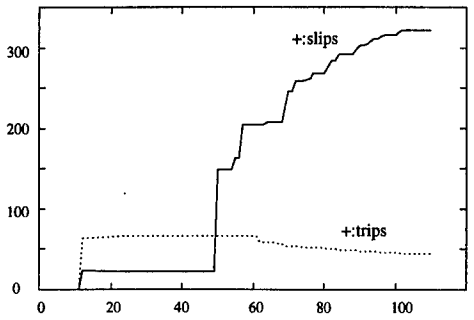


Figure 2: The activation trace of collector nodes *+:slip* and *+:trip* during the processing of the "John fell" story. X-axis records the number of cycles. Each cycle may correspond to ~50-100 msecs.

3 Simulation Result

The activation trace resulting from the processing of the "John fell" story is shown in Figures 2 and 3. Figure 2 shows the actual activation levels of the *+:slip* and *+:trip* nodes as the story is processed by SHRUTI. Figure 3 depicts the activation trace of a larger *subset* of nodes. The depiction in this figure, however, has been simplified to highlight key aspects of the network behavior. In particular, several nodes have been omitted, some intermediate cycles have been omitted and the activation levels of collector and enabler nodes have been discretized to four levels. Please note that due to simplifications made to Figure 3, the time scales along the x-axis in Figures 2 and 3 are not the same. To minimize confusion, we will refer to the times in Figure 2 as cycles and in Figure 3 as steps.

A sentence is conveyed to SHRUTI by activating the

[2] This is modeled after the biological phenomena of short-term potentiation (STP) [2].

[3] The use of inhibitory connections for explaining away is motivated in part by [1].

+ node of the appropriate relation and establishing role-entity bindings by the synchronous activation of the appropriate role and entity nodes. The sentences are presented in sequence and after each sentence presentation, the network is allowed to propagate activity for a fixed number of cycles. For example, the first sentence (S1) is communicated to SHRUTI in step 1 (cycle 0) by activating the node *+:fall*, the nodes *fall-pat* and *+:John* in synchrony, and the nodes *fall-loc* and *+:Hallway* in synchrony. The firing of nodes *+John* and *+:Hallway* occupy distinct phases — ρ_1 and ρ_2, respectively.

Activation from the focal cluster for *fall* reaches the mediator structures of rules (1) and (2) shown in Figure 1. Consequently, nodes *r1* and *r2* in the mediator for rule (1) become active in phases ρ_1 and ρ_2, respectively. Similarly, nodes *s1* and *s2* in the mediator of rule (2) become active in phases ρ_1 and ρ_2, respectively. At the same time, the activation from *+:fall* activates *?:fall* which in turn activates the enablers *?:med1* and *?:med2* (the activity of mediator nodes, and role nodes of *slip* and *trip* is not depicted in Figure 3). The activation from nodes *r1* and *r2* reaches the roles *slip-pat* and *slip-loc* in the *slip* focal cluster, respectively. Activation also reaches *trip-pat* and *trip-loc*. In essence, the system has created new bindings for the *slip* and *trip* relations. These bindings together with the activation of the nodes *?:slip* and *?:trip* encode two queries: "Did John slip in the hallway?", and "Did John trip in the hallway?". At the same time, activation travels in the type hierarchy and activates the nodes *?v:Man*, then *?v:Person*, and then *?v:Agent* in phase ρ_1, and the *?v:Location* node in phase ρ_2. The coincident activity of *slip-pat* and *?v:Agent* node, and the coincident activity of the *slip-loc* and *?v:Location* nodes leads to the firing of the T-fact F1 associated with slip. The activation of F1 causes activation from *?:slip* to flow to *+:slip*. The T-fact F2 associated with trip also becomes active in an analogous manner and conveys activation from *?:trip* to *+:trip*. The level of these activations is a measure of the prior probabilities that a person may slip or trip. At this time, "John tripped" is believed to be a more likely explanation of "John fell" than "John slipped."

While the activation spreads "backwards" from the *fall* focal cluster as described above, activation also travels "forwards" to the *hurt* focal-cluster (not shown in Figure 1) and leads to the prediction that John got hurt.

The introduction of sentence S2 in step 6 (Figure 3) (cycle 40 Figure 2) results in the instantiation of *clean* with the bindings (⟨*clean-agt=+:Tom*⟩, *and* ⟨*clean-loc=+e:Location*⟩). As a result, Tom gets active in phase ρ_3 and *+e:Location* in phase ρ_4. Note that now we have two instantiations of a location. The second instantiation gets merged with the first (Hallway) as a result of phase merging described in Section 2.1.4. This happens in step 8 (see activity of *+e:location* in Figure 3). At this time, *+:wetFloor* also becomes active as a result of activity arriving from *+:clean* via the mediator of the rule "cleaning leads to a wet floor" (not shown in Figure 1). By step 10 (Figure 3) *+:slip* becomes more active as a result of the high activation of *+:wetFloor*. The effect of "explaining away" kicks in and causes the activation of *+:trip* to go down by step 12. The strength of *+:slip* increases even further due to (i) the potentiation of links from the mediator for the rule "walking on a wet floor may cause slipping" (not shown in Figure 1), (ii) the potentiation of the link from *?:med1* to *?:slip*, and (iii) the effect of explaining away. The effect of these changes on the activation levels of *+:slip* and *+:trip* may be seen more vividly in the detailed trace shown in Figure 2.[4]

S3 is introduced in step 14 (cycle 80) with the binding (⟨ *hurt-pat=+e:Man*⟩). This leads to *+e:Man* becoming active in phase ρ_4 and a second dynamic instantiation of *hurt* (in addition to the earlier instantiation resulting from the inference *hurt(John)*). These two instantiations get merged immediately, and phase ρ_4 gets merged with ρ_1 (John) in step 15 as a result of the phase merging described in Section 2.1.

4 Evidence combination

The problem of evidence combination arises even in the limited example discussed above. This probolem, however, can become far more complex in real-world situations. It becomes apparent as the system is used to model increasingly complicated domains that there is a need for a significant degree of flexibility in the manner in which evidential values are combined.

There are many places in SHRUTI where activity converging on a node from different sources must be combined to determine an output value for the node. The combination of collector activity from multiple antecedents and also across multiple rules, and of enabler activity from multiple consequents and multiple rules, are prime examples of this. At the locations where evidence from facts is incorporated, in the influence of collector activity on an enabler node, and in propagation of activity through the type hierarchy, as well as in a number of other situations, evidence combination is also being performed.

Evidence combination in SHRUTI takes the form of a set of evidence combination functions, or ECFs. At each point in the network where evidence must be combined, a particular ECF is chosen. In selecting the range

[4] If sentence S2 were delayed, the activity in *slip* would lead to the instantiation of an instance of *clean* with an entity of type *agent* being instantiated as a potential filler of the role *clean-agt*. This entity, however, would get unified with *Tom* upon the introduction of S2.

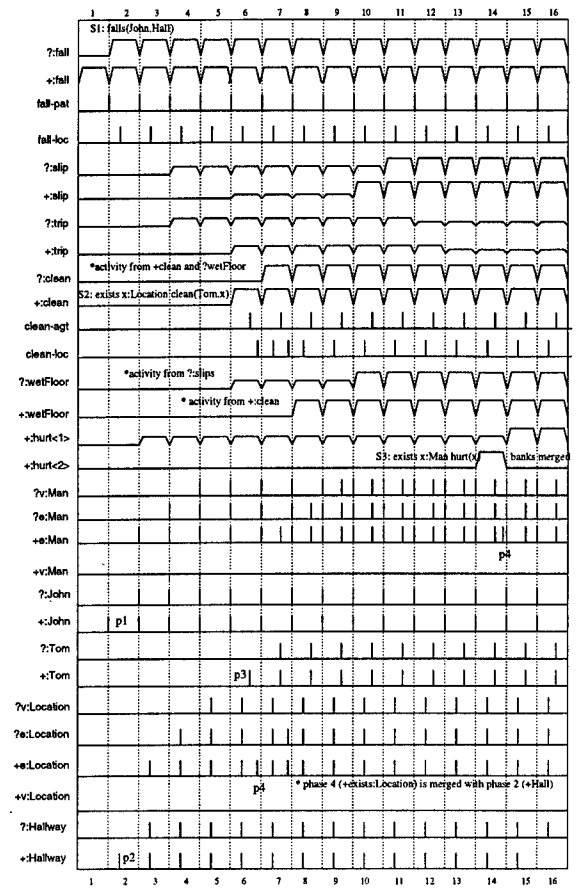


Figure 3: Schematized activation trace of selected nodes.

of functions, the goal was to have a set large enough to adequately model real-world data, but small enough such that the choice between functions for a particular situation is relatively simple. Moreover, these functions should be computationally simple, or at least decomposable into very simple parts, such that the biological plausibility of the system is not sacrificed. The set of functions developed is not intended to cover all possible relations, but instead to be sufficiently flexible so as to capture the vast majority of practical situations.

4.1 Background

An obvious source of inspiration for this undertaking is fuzzy logic, where a multitude of functions have been developed to combine fuzzy membership values in different ways [13]. Of particular note are the classes of binary operators known as T-norms and S-norms. These represent, respectively, general forms of fuzzy set intersection and union. The extension of T-norm and S-norm operators to handle combination of multiple evidence values is generally straightforward, and these two categories are prime candidates for evidence combination functions in SHRUTI.

In neural networks, greater representational complexity is achieved by adding more and more nodes and interconnections, keeping the combination function at the nodes very simple. The commonest form of evidence combination in this context is the sigmoid-sum; this function has a number of properties which make it appealing for use in neural nets. Although other functions can certainly be used as well, it would not make sense with a standard neural net to pick and choose particular functions for particular nodes, since the nodes have no special meaning. In a structured connectionist system like SHRUTI, however, nodes are meaningful and the network structure is relatively fixed, so it is useful to push more flexibility into the combination functions than is either necessary or possible in standard neural nets. Belief nets also utilize a form of evidence combination, found in the conditional probability tables associated with each node. In the case of a full CPT, the flexibility of combination is high but so are the storage and computational demands. The often used noisy-OR function [8] reduces these demands but when used universally, as is often the case, is overly restrictive. Other means of reducing complexity of the CPT, such as encoding it with a tree structure, demonstrate a different approach to evidence combination than that envisioned here [4].

4.2 Combination Functions

The combination functions developed for SHRUTI can be thought of as forming a continuum, with *and* at one end and *or* at the other. In between these two extremes are four basic categories of functions: *soft-and* (with values up to min), *soft-min* (ranging from min to average), *soft-max* (ranging from average to max), and *soft-or* (ranging from max and up). Although specific functions have been chosen to represent each of these categories, many of the functions developed for fuzzy logic could be used here. As a general rule, antecedents or consequents which are correlated will be combined into a single multiple-antecedent or multiple-consequent rule in SHRUTI, whereas uncorrelated factors will reside in separate rules. This means that for the former case, evidence combination functions should allow for this correlation, while in the latter assumptions of independence are usually justified. It is proposed that most meaningful combinations of evidence can be characterized as belonging to one of these four basic categories, on the basis of the necessity or sufficiency and also degree of correlation of their inputs.

Link weights can play an important and context-dependent role in many of these functions. The standard

use of link weights is to multiply them with the input values prior to doing evidence combination. However, instead of simply affecting values before they are combined, weights can also be used as additional function parameters, with different interpretation for different functions. The use of link weights in this manner, elaborated for each of the ECFs below, provides a significant degree of flexibility in the kinds of relations that can be represented. While this use of link weights appears to run counter to biological intuitions, it is possible to replace each of the so-weighted combination functions with an expanded network which involves only very simple combinations and which employs link weights in a more standard manner.

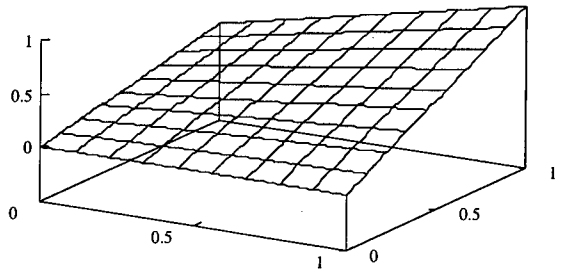


Figure 4: Graph of a weighted *and* function with antecedent weights of 1.0 and 0.6.

4.2.1 Soft-And and Soft-Or

At one end of the spectrum of functions are the *and*-like functions, corresponding to the T-norms of fuzzy logic, which are appropriate for combining evidential values which are deemed necessary. The basic *weighted and* is calculated as $\prod_i(1-(1-X_i)W_i)$ (where X_i is the evidential value and W_i is the weight for the ith incoming link). This *and* function is most appropriate for combining independent sources of evidence, such as in the following example rule: $\forall$*w,x:Person, y:Object*[AND(*canSell(x,y) 1000,wants(w,y) 800)* $\Rightarrow$ *sells(x,w,y) 500*] where both a potential seller's ability to sell and a potential buyer's desire to buy are necessary and independent prerequisites of a sale actually taking place. The collector node of the mediator for this rule, which combines activity from the antecedents *canSell* and *wants*, will utilize the *and* function. The weight of 500 specified for the consequent means this can only be concluded with half of the maximum possible strength. With the independence assumption relaxed, assuming instead that combined values are positively correlated, a *soft-and* function is appropriate which has a value greater than the product-based *and*. The function chosen for this purpose is the *and(X)/or(X)* function which is similar to the Hamacher product T-norm $H(x,y) = xy/(w+(1-w)(x+y-xy))$ (where w is a weight in [0,1]) generalized to n variables [3]. The utilization of link weights brings another dimension to the standard *and* function. Since the main characteristic of the and is that combined values are regarded as necessary, an obvious interpretation for the link weights is that they reflect the degree of necessity. In probabilistic terms, this would be the probability that the consequent is false given that the antecedent is false. This means that lower link weights on the *and* function generally result in higher output values. While this fact may be counterintuitive when considering the network behavior, assignment of degrees of necessity seems quite practical from a knowledge engineering standpoint, and makes the simple and function remarkably flexible. In the above example, the interpretation is that while *canSell* is absolutely necessary in order to draw any conclusion about a sale taking place, *wants* is not.

Shown above is a graph of a weighted *and* function with two antecedents of weights 1.0 and 0.6 (see Figure 2). The relative importance of the first value is seen in that the function value changes slowly while traveling along the near axis, but rapidly when traveling along the far axis.

At the other end of the spectrum is the weighted *or*, given as $(1-\prod_i(1-X_i*W_i))$. *Or*-like functions, which can be thought of as those having output values at least equal to the maximum input, are used when there is a notion of sufficiency of individual antecedents to affect the consequent. These correspond roughly to the S-norms of fuzzy logic, and many of these fuzzy operators might be adapted to the task. *Or* is the most commonly used function for combining activity from different rules that converge on a particular concept. As *or* assumes that antecedents are independent, a *soft-or* (the complement of the *soft-and*) is provided to handle cases where antecedents are correlated. This is in particular the function of choice for combining enabler activity from multiple consequents, which are most certainly correlated. The general requirement for *soft-or* is that its value be less than the *or* but still greater than the *max*. The natural interpretation of link weights for *or*-like functions is that they represent the degree of sufficiency of the source concept - the probability of the consequent given only the particular antecedent.

4.2.2 Weighted Averages: Soft-min and Soft-max

Covering the range between *min* and *max* are the weighted averages. Weighted averages are appropriate when individual antecedents are neither necessary nor sufficient. For all of these functions the link weights represent degrees of influence, giving the relative ef-

fect of an antecedent value on the output. There are two main functions in this class: the *soft-min* function $((\sum_i X_i^k W_i)/(\sum_i W_i))^{1/k}$ for $k \in (0,1)$, and the *soft-max* function with $k \in (1,\infty)$. It should be noted that *min* and *max* are the limits of the given *soft-min* and *soft-max* functions, as $k \to 0$ and $k \to \infty$, respectively, so this whole range from *min* to *max* is really only one functional form with a varying parameter. *Soft-min* is used when it is necessary that most of the evidence for the antecedents be available in order to conclude the consequent, but unlike *and* no single piece is required. Combining evidence about the symptoms associated with a particular syndrome is a place where *soft-min* can be appropriate. A syndrome is a specific set of co-occurring symptoms and so in deciding whether a particular syndrome is present, lack of evidence for one of the particular symptoms should weigh heavily against a positive conclusion. But it should still be possible to conclude that a syndrome is present even if evidence for one of its symptoms is absent, so any *and*-like function would not be quite appropriate and *soft-min* is the function of choice. With *soft-max*, only a fraction of the potential evidence is sufficient to lead to strong activity in the consequent, but unlike *or* no single piece is alone enough. The following rule provides a reasonable example usage of *soft-max*:

$\forall x$:*Person*[SOFTMAX*(tall(x) 500, athletic(x) 800, practiceDaily(x,Basketball))* $\Rightarrow$ *goodAt(x,Basketball)*]

Here each factor can contribute significantly to the result, but none is really sufficient to draw much of a conclusion absent some other support.

5 Conclusion

We have discussed how causal and referential coherence can arise within a neurally plausible system as a result of spontaneous activity in a network. The network's structure reflects the causal model of the environment and when the nodes in the network are activated to reflect a given state of affairs, the network spontaneously combines evidence, seeks coherent explanations, and makes likely predictions. The time taken to perform an inference is simply proportional to the depth of the causal derivation and is otherwise independent of the size of the causal model. The state of coherence is reflected as reverberation around *closed loops*. The reverberating pattern of rhythmic activity also codes dynamic bindings via synchronous activity. Coherence arises in SHRUTI as a result of (i) flexible evidence combination, (ii) inhibitory interactions among sibling entities, types and rules, (iii) short-term increase in link weights resulting from short-term potentiation, and (iv) the dynamic merging and instantiation of entities.

References

[1] Ajjanagadde, V. (1991) Abductive reasoning in connectionist networks. TR WSI 91-6, Wilhelm-Schickard Institute, University of Tubingen, Tubingen, Germany.

[2] Bliss, T.V.P. and Collingridge, G.L. (1993) A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* 361, 31–39.

[3] Driankov, D. Helendoorn, H. and Reinfrank, M. 1996. *An Introduction to Fuzzy Control.* Springer-Verlag: Berlin.

[4] Friedman, N. and Goldszmidt, M. (1996) Learning Bayesian Networks with Local Structure. In *Proceedings of the Twelfth Conference on Uncertainty in Artificial Intelligence.*

[5] Keenan, J. M., Baillet, S. D., & Brown, P. (1984) The Effects of Causal Cohesion on Comprehension and Memory. *Jrnl of Verbal Learning and Verbal Beh.*, 23, 115-126.

[6] Kintsch, W. (1988) The Role of Knowledge Discourse Comprehension: A Construction-Integration Model. *Psych. Rev.*, Vol. 95, 163-182.

[7] McKoon, G., & Ratcliff, R. (1980) The Comprehension Processes and Memory Structures Involved in Anaphoric Reference. *Jrnl. of Verbal Learning and Verbal Beh.*, 19, 668-682.

[8] J. Pearl. 1988. *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufmann.

[9] Potts, G. R., Keenan, J. M., & Golding, J. M. (1988) Assessing the Occurrence of Elaborative Inferences: Lexical Decision versus Naming. *Journal of Memory and Language*, 27, 399-415.

[10] Shastri, L. (1999) Advances in SHRUTI — A neurally motivated model of relational knowledge representation and rapid inference using temporal synchrony., *Applied Intelligence.* To appear.

[11] Shastri, L. & Ajjanagadde V. (1993) From simple associations to systematic reasoning. *Behavioral and Brain Sciences*, 16:3 p. 417-494.

[12] von der Malsburg, C. (1981) The correlation theory of brain function. Internal Report 81-2. Department of Neurobiology, Max-Planck Institute for Biophysical Chemistry, Gottingen, Germany.

[13] Zadeh, L. 1996. *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A Zadeh,* G. Klir & B. Yuan (eds), World Scientific: Singapore, River Edge, N.J.

Security and Performance of Storage Area Networks

Qiang Li
Department of Computer Engineering
Santa Clara University, Santa Clara, CA 94043, USA
Email: qli@sunrise.scu.edu

May Allam
Department of Computer Engineering
Santa Clara University, Santa Clara, CA 94043, USA
Email: mallam@scudc.scu.edu

Abstract – *Attaching storage devices to the network provides direct transfer between clients and storage. This new distributed storage architecture model has a potential to offer high bandwidth, low latency, scalability, and availability of storage to clients as well as multiple servers. Realizing this new technology require careful consideration in the design of the storage devices, the choice of a network technology, and the design of the high-level file system. Moreover, promoting the storage devices from I/O peripherals to network peripherals imposes new security concerns. This paper describes the performance gain of this new storage architecture as well as the design and security issues involved. A brief survey of the current research is also presented.*

Key Words: Storage Area Network, Network Attached Storage, Network Security, File system, Distributed Systems.

1. Introduction

Data sharing between a working group is a fundamental aspect of today's organizations. Typically, the working group has a number of workstations connected by a local area network such as Ethernet, FDDI, or ATM. One or more workstations have large-capacity storage devices attached to them and are dedicated for the storage and retrieval of data that need to be shared. These dedicated workstations are called servers; the other workstations, which are majority, are called clients. The storage devices are attached to the server by an I/O bus such as SCSI as illustrated in Figure 1.

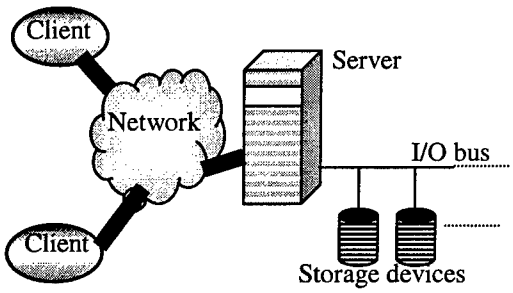


Figure 1: Server Attached Storage.

Typically, the data is represented as files, and a distributed file system, such as NFS (Network File System) [20] or AFS (Andrew File system) [21], abstracts the distributed files into a common file system for the applications running on the client machines. The distributed file system provides access to the remote files by sending requests to a file server. All data travelling between clients and storage devices must be stored and forwarded by the server which frequently gets overloaded and becomes a bottleneck limiting thus the data bandwidth offered to the clients. The problem aggravates as the size of the transferred data increases which is the case for many I/O bound applications such as multimedia or data mining applications [2].

Although, the server attached storage architecture, described above and shown in Figure 1, is the most common architecture in today's offices and campuses, it limits the performance and availability of the system to the server's I/O channel capacity and load.

The key to decreasing the server workload and thus increasing the effective bandwidth offered to the clients is to have direct transfer between clients and storage, at least for large data sets. This requires attaching the storage devices directly to the network as illustrated in Figure 2. In this architecture, a storage device is promoted from an I/O peripheral to a network peripheral. The architectural model is generally referred to as the Storage Area Network or SAN. We call the storage devices in question the SAN devices.

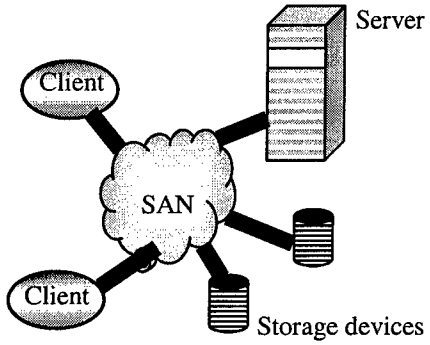


Figure 2: Network Attached Storage.

Direct transfer between clients and storage improves scalability, that is, the bandwidth linearly increases by

increasing the number of clients and storage devices, up to the limit of the bandwidth of the network, not the limit of server architecture. Scalability can be further improved through data striping which is now easier and cheaper to implement over storage instead of servers.

However, exposing the storage devices on the network introduces new security risks not encountered when storage was connected to a trusted file server through a private bus like SCSI.

To realize this storage architecture, the following issues must be considered:

1. The design of the network-attached storage device and the functionality that the device should provide.
2. The selection of a network capable of carrying the high bandwidth gained by this new storage architecture.
3. The distributed file system. The file system services need to be repartitioned between storage devices, clients, and servers. The file system must be efficient enough to pass the increased gain in bandwidth all the way to the clients' applications.

The design aspects of SAN devices are presented in section 2; the different networking options available for SAN are described in section 3. Section 4 and 5 discuss the file system and security issues for SAN respectively. Section 6 discusses how to extend the functionality of storage devices through programmability for supporting certain applications. Section 7 concludes this paper.

2. SAN Device

Externalizing a storage device from a server and directly exposing it on the network imposes new functionality on the storage device such as how to communicate with a requester, what operations to provide to the requester, etc. This new functionality can be categorized as follows:

- **Communication:** Being directly connected to the network, a SAN device must support a communication protocol to communicate with clients and servers. The NASD project at CMU used RPC over UDP/IP for their prototype. Their results showed that 90-97% of the time spent in reading/writing 64 KB of data was spent in communications [1]. A light-weight communication mechanism is definitely needed for SAN. This issue is still under research. The Netstation project at USC favors Internet protocols (TCP/UDP/IP) for Network-Attached peripherals because of their support for heterogeneous networks and high-connectivity [4].

- **Self-Management:** Achieving self-management on storage devices is important for scalability issues. Traditionally, storage devices are block-addressable. Their storage space is represented as an ordered set of fixed-length data blocks called sectors and it is the up to the file system to manage the fixed-length blocks. SAN devices take the responsibility of self-management by abstracting their internal organization into variable-length objects [1,3]. An object on such a storage device consists of an ordered set of sectors associated with a unique identifier. Data is referenced by the identifier to the object and an offset. An object also has a set of attributes associated with it. The Storage System Program at HPL works on achieving Quality-of-Service (QoS) by associating an object with QoS attributes such as capacity, performance capabilities, and a set of availability and reliability metrics to satisfy QoS requirements of the application workload [6].

- **Operational:** SAN Devices must provide an interface to process requests such as read data, write data, set an object attribute, etc. They can also provide richer retrieve/store semantics beyond simple read/write such as compressed/decompressed write/read operations [7], atomic transactions, and lock management (for concurrency) [6].

- **Access Control:** Since traditional storage devices are privately connected to a server, they execute every command they receive without worrying about any authorization. However, SAN devices must decide if a request for an operation should be granted or denied. It is the high-level file system that defines the access-right policy.
When a client request to perform an operation, two actions must be taken: access decision and enforcement of the decision. It is important to differentiate between the two actions because they can be done by two different entities as in the NASD project at CMU which introduced the concept of "Asynchronous Oversight" where access decisions are made once by a server and are asynchronously enforced by the drives [1].

- **Network Security:** Networks are inherently insecure. An adversary can eavesdrop, modify, or fake a request or a response. The concerns about security are well known to programmers of distributed systems, but not common issues for designers of I/O peripherals. SAN devices must provide cryptographic capabilities for privacy, authentication, and integrity. Secret keys used by the cryptographic algorithm must

be safely stored on the drives on a tamper-resistant hardware. Software implementations of cryptographic algorithms are prohibitive. The best solution is to add a secure coprocessor on the drive that can perform the cryptographic operations at an adequate speed, preserving thus the offered bandwidth, and can also encapsulate the cryptographic keys [9]. More details on access control and security can be found in section 4.

It is better to design the storage device such that it offers a fixed interface independent of a file system. This gives users the freedom to select a filesystem that best suit their applications and allows the storage device and filesystem to evolve independently [8].

3. Storage Area Network

The choice of a network technology is critical to the performance of the network attached storage architecture. The architectural model of SAN is somewhat independent to the physical layer and the link layer of the network. In general the network should satisfy SAN's requirements on bandwidth, latency and reliability [10,11,12,13].

SAN has the following components: SAN-interfaces, SAN-interconnects, and SAN-fabrics [11]. SAN-interfaces are generally ESCON, SCSI, HIPPI, or Fibre Channel. Like LANs and WANs, SAN-interconnects have routers, hubs, switches, and gateways. The most common SAN-fabrics are Switched SCSI, Fibre Channel Switched, and Switched SSA. SAN interconnects link SAN interfaces to SAN fabrics as shown in Figure 3.

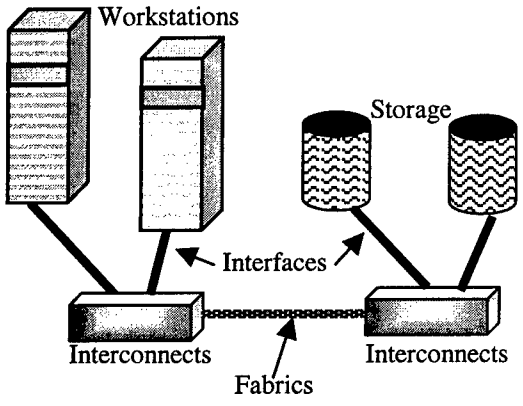


Figure 3: Components of a SAN Network.

Among the different SAN interfaces, the best fit for network attached storage is Fibre Channel because of its three desirable features: speed, distance, and connectivity. The current Fibre channel operates at a transfer rate of 1 Gigabits/s and can span up to 10km distance with single mode fibre. [12].

Fibre Channel offers a single standard interface capable of simultaneously, supporting both data channel and network connections. Fibre Channel can be used to carry a number of data channels and network protocols such as ATM (Asynchronous Transfer Mode), FDDI (Fibre Distributed Data Interface), Ethernet, HIPPI (High Performance Parallel Interface), SCSI and IPI over a single medium and with the same hardware connection [13].

4. Distributed File systems and Storage Area Network

Widely used distributed file systems such as AFS [21] and NFS [20] were designed mainly for the server-attached storage architecture described in section1. However, as the SAN technology gets increasingly omnipresent, it becomes reasonable to consider distributed file systems that exploit the network-attached storage architecture and address new issues such as synchronization and parallelism and provide better availability and scalability.

Distributed file systems that allow direct data transfer between clients and storage devices and allow the data on the storage device to be shared by more than one client are broadly called *Shared File Systems* [18].

Shared file systems that allow large files to be striped into subfiles on several storage devices and allow parallel scattering and gathering of files are called *Parallel File Systems* (PFS).

The design of shared file systems may differ in many aspects:

- **Inclusion of a sever**: A shared file system may be serverless, that is clients can perform all operations directly on the storage devices without the need of a file server. This category of shared file systems is called *symmetric*. If the shared file system requires a file server, often called file manager, to provide the clients with information or authorization for accessing the storage devices, the file system is called *asymmetric*. Although symmetric shared file systems allow the design of storage devices to be independent of the file system [8], they are vulnerable to file manager failure.

- **Lock management:** To allow more than one client to access the same data simultaneously, a locking mechanism must be provided to ensure mutual

exclusion. The locking can be performed either on clients or storage devices. Asymmetric shared file system give the option of performing the locking on the file server. Although, locking performed on clients has the advantage of balancing the lock management among clients, the mechanisms to recover from client failures are complex and may limit the overall scalability. Locking on storage devices is easier and faster than client-based locking.

The following two subsections present two shared file systems:

- The Global File System at UMN
- NASD Parallel File Sustem at CMU.

4.1 Global File System (GFS)

Global file system or GFS is a project carried at the University of Minnesota [19]. GFS is a symmetric shared file system (serverless), with device-based locks. It provides parallel access to storage devices through UNIX file system semantics: *open*, *close*, *read*, *write*, and *fcntl*.

In GFS, the storage devices form a global pool called Network Storage Pool (NSP) which is partitioned into a number of subpools. Each subpool may be configured with different characteristics. A system administrator performs the partitioning and configuration.

GFS employs a SAN to connect clients to the network storage pool and offers a thin protocol layer for communication between clients and storage. Their preliminary results show that GFS scaled well for 4 clients.

4.2 NASD Parallel File System

NASD is a project carried out by the Parallel Data Lab at Carnegie Mellon University [1]. NASD is an asymmetric storage architecture; it uses a file manager for managing global naming, access control decisions and consistency.

They first attempted to port NFS and AFS to the NASD environment but since NFS and AFS didn't exploit the potential high bandwidth offered by the NASD architecture, they implemented a parallel filesystem (NASD PFS) capable of making large parallel requests to files striped across several storage devices. However, they introduced an additional server called storage manager responsible of handling concurrency, mapping and redundancy. The extra server hides the file striping and allocations from the file manager by presenting itself as a virtual NASD drive storing the entire file. To access a striped file, clients first contact the file manager, which directs them to the virtual drive i.e. storage manager, which in turn directs them to the physical storage devices storing the actual subfiles.
Their results showed that they achieved an aggregate bandwidth that scales linearly. More on CMU's NASD can be found in the following section

5. Security and Network Attached Storage

Exposing storage devices to the insecure network imposes the security concerns involved in any distributed system, namely: privacy, authentication, and key management.

- **Privacy:** Privacy is achieved through encryption algorithms such as DES [16]. Data and/or control can be encrypted while being transmitted to protect against eavesdropping by adversaries. Furthermore, if the device is stored in a physically unsafe location, data can be stored in an encrypted form on the storage device.

- **Authentication:** Message authentication is achieved through cryptographic algorithm such as HMAC [17] to protect data and/or control parts of the exchanged messages against "man_in_the_middle" attacks.

- **Key management:** All cryptographic algorithms require the usage of keys and the security of the algorithms depends on keeping those keys private. The best way to store the keys is to keep them on a tamper-proof hardware.

Software implementations of cryptographic algorithms are time-consuming. It is best to use hardware support such as a secure co-processor for performing the cryptographic operations efficiently and for storing cryptographic keys safely [9].

5.1 Access Control

Another security concern of the network-attached storage architecture is adherence to the high-level file system's access right policy.

Access rights define the types of operations that a client can perform on a particular set of data. If the storage device has to make the access decisions by itself, the design of the storage device will become file system-dependent, which is undesirable.

A better approach would be to let the storage device enforce an access decision taken by another file system-dependent entity such as a file server. In other words, the access right functionality is distributed between the storage device and the file server. The server makes a decision; the device applies it. This leads to another security concern: how the server conveys to the storage the access decision through the unsecured network. Hostile users can eavesdrop and/or modify the network traffic and can fake requests and/or responses.

The next 2 subsections describes the security approach of two projects:

- The Netstation project at USC.
- The NASD project at CMU.

A comparison of the two approaches concludes this section.

5.2 Security of ISI's Netstation

Netstation is a project carried out by the Information Sciences Institute at the University of Southern California. The goal of this project is to demonstrate that gigabit LANs can effectively replace the system bus in conventional workstations [4,14].

The project proposes a mechanism called DVD (derived virtual devices) for secure shared access of client to network-attached peripherals. The system components, involved in security, are a Kerberos server, a ticket-granting server (TGS), disks, a file manager server called "storage manager", and clients. The system relies on Kerberos for authenticating disks, the storage manager, and clients.

When a client attempts to open a file for reading, the following procedure occurs (see Figure 4):

1-4) The client authenticates itself to Kerberos and acquires a ticket from TGS to access the storage manager.
5) The client sends a request to the storage manager.
6) The storage manager sends a request to the disk to create a DVD (derived virtual disk) that includes only the required file. The request includes the access rights for the new DVD that is who can access it and what operations are permitted.
7) The disk creates the requested DVD and sends an acknowledgment to the storage manager.
8) The storage manager returns to the client a handle for this newly created DVD.
9-12) If this is the first time the client accesses the disk, the client must request a ticket from TGS to access the disk. Otherwise, this step is omitted.
13) The client sends a request to the disk.
14) The disk validates the request then sends the requested data to the client.

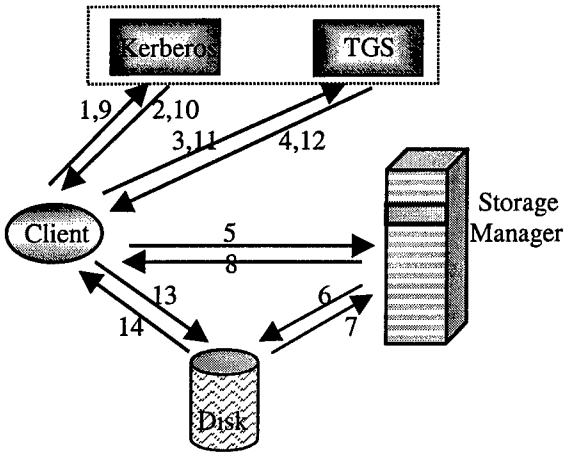


Figure 4: Netstation procedure for opening a file for reading.

5.3 Security of CMU's NASD

The goal of the NASD project at CMU is to design a high-bandwidth, scalable, and cost-effective storage architecture [1,2,6,8].

The project proposes an access control mechanism called "asynchronous file system oversight" where once an access right is granted for a client, this client can use this right over and over without further consultation.

The system components are clients, disks, and a file manager server. The file manager does the file system-dependent functionality such as name-space management and access-control decisions. Access rights for a client on an object are made once by the file manager and are enforced asynchronously by the drive.

The procedure for reading a file is illustrated in Figure 5 and consists of the following steps:

1-2) The first time a client accesses the file, the client requests a capability from the file manager. The file manager validates the request according to the file system's access policy and issues a capability to the client. Once the capability for accessing a file is acquired by the client, this step is omitted for future accesses to this file.
3) The client uses the capability in sending requests to the drive.
4) The drive validates the request then sends the requested data to the client.

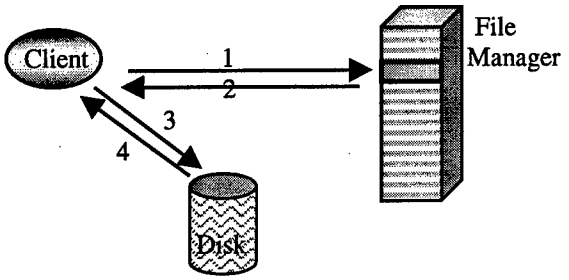


Figure 5: NASD procedure for reading a file.

The design of the capability uses cryptography to allow the drive to verify that the file manger generated the capability and that the capability was not modified on its way. This is achieved by having the drive share secret keys with the file manager. Each object on the disk has a key associated with it called working key. The file manager uses the working key of an object to sign the capability given to the user to access this particular object. The capability includes the type of operations that can be performed by the client on the drive.

5.4 A Comparison of ISI's Netstation and CMU's NASD Security

The following is a comparison between the two security approaches:

- **Number of messages:** The number of messages exchanged in the Netstation project is clearly larger than the number of messages exchanged by CMU's NASD. However, 8 of 14 messages shown on Figure 4 are for client authentication by Kerberos. CMU's NASD does not specify how the file manager authenticates the client. In ISI's Netstation, the file manager informs the drive directly about the access rights and gets an acknowledgement from the drive; in CMU's NASD, the file manager passes the access rights to the drive though the client eliminating thus 2 messages.

- **Persistence of access rights:** In ISI's Netstation, DVDs are dynamically created and destroyed when a client opens or closes a file respectively, that is, a client's access right to a file are lost once the file is closed. In CMU's NASD, the access right given to a client through a capability persists as long as the capability did not expire (capabilities are time-limited). Because of the long persistence of capabilities on CMU's NASD, a revocation mechanism is also provided to be able to cancel a capability issued to a client for an object when the object's access rights are changed.

- **Drive States:** ISI's Netstation requires the drive to have more states because of oustanding *open* requests. CMU's NASD does not keep track of any "per client" information.

Although NASD offers a better security mechanism in many aspects than Netstation, it also has its drawbacks. NASD security mechanism is suitable for the case where a request accesses only a single object. However, if the drive supports "batch-style" requests i.e. a single request can access a large number of objects, the security mechanism might become cumbersome.

It might be useful to consider a security mechanism that relies on public keys. Digital signature using public keys offers the advantage of non-repudiable communications. Certificate servers are becoming widely used in organizations and can be used to issue the keys instead of the file manager.

6. Mobile Agents on Network Attached Storage Devices

There is a potential computation power embedded in the storage devices that needs to be exploited. Applications that process a massive amount of data can take advantage of this computation power by sending code (mobile agents) to execute remotely on the storage devices near the data instead of downloading large data over the network. Storage disks that can accept code for execution are called *Active Disks*.

To realize the full benefits of active disks, the following issues need to be considered:

- **Parallelism:** One storage device by itself may not have a processor as powerful as those on workstations (where the application code runs), but the combined computation power of the processors on all the storage devices is definitely high. (The data is striped anyway over the storage devices).

- **Communications:** the storage device must provide an interface for receiving mobile agents and provide a communication mechanism so the downloaded code can communicate the results.

- **Safe execution:** the storage device must provide a safe environment for executing the mobile agents. The execution environment must ensure that the downloaded code does not violate the access right policy imposed by the file system, is not corrupting

memory or the execution environment itself, and is not consuming the drive's resources.

- **Application type:** Applications that can benefit from sending code to execute on the storage devices are those that include operations which scan a large amount of data to extract a relatively small amount e.g. exhaustive search, data mining, and contour tracing of images. The high-level file system can also benefit from sending code to execute on the drive. Moreover, code can be added to the drive to optimize and customize its standard interface to the high-level file system e.g. provide a response for hints on future read operations or provide a user authentication mechanism.

Executing functions on storage and near the data has been exploited for a long time in database systems. Recently, UCSB [23] and CMU [22] have exploited the idea for storage disks.

7. Conclusion

SAN provides direct transfer between clients and storage, which results in a high and scalable bandwidth to clients. Striping data over storage devices further enhances scalability.

Exposing the storage devices on the network requires them to provide sound security mechanism to protect clients' data. Hardware support for cryptographic operations is needed to preserve the bandwidth offered by the storage device.

Widely distributed file systems such as NFS and AFS fail to provide the high aggregated bandwidth provide by SAN technology. Applications that need to utilize the full performance of SAN should replace the old file systems with new ones that can support parallel I/O.

8. References

[1] Gibson, G. et al., "A Cost-Effective, High-Bandwidth Storage Architecture", Proc. of the 8th Conference on Architectural Support for Programming Languages and Operating Systems, 1998.

[2] Gibson, G. et al., "File Server Scaling with Network-Attached Secure Disks", Proceedings of the ACM International Conference on Measurement and Modeling of Computer Systems (Sigmetrics '97), 1997.

[3] Anderson, D., "Object-Oriented Devices: Description of Requirements", NASD Technology, www.nsic.org/nasd/technol.html, June 1998.

[4] Hotz, S. et al., "Internet Protocols for Network-Attached Peripherals", Proc. Of the 6th NASA Goddard Conference on Mass Storage Systems and Technologies, 1997.

[5] Borowsky, E. et al., "Using attribute-managed storage to achieve QoS", 5th Intl. Workshop on Quality of Service, Columbia Univ., New York, June 1997.

[6] Khalil, A. et al., "Scalable Concurrency Control and Recovery for Shared Storage Arrays", CMU-CS-99-111, February 1999.

[7] Burrows, M. et al, "On-line Data Compression in a Log-structured File System", Proc. of the 5th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS-V), October 1992.

[8] Gobioff, H. et al., "Security for Network Attached Storage Devices", CMU-CS-97-185, October 1997.

[9] Yee, B. and Tygar, D.,"Secure coprocessors in electronic commerce applications." Proc. Of 1st USENIX Workshop on Electronic Commerce, New York, July 1995.

[10] Horst, R. and Garcia, D., "ServerNet SAN I/O Architecture", Hot Interconnects V, 1997.

[11] Peterson M., "Storage Area Networking the next network", Strategic Research Corp., www.sresearc.com/105527.htm, January 1998.

[12] Wilson, S., "Managing a Fibre Channel Storage Area Network", Fibre Channel Community, www.fcloop.org/SAN/whitepapers/managesan/default.htm, November 1998.

[13] "Networking Fibre Channel - A Guide to Fibre Channel Storage Area Networking (SAN)", www.fcloop.org/SAN/whitepapers/networkfc/default.htm.

[14] Van Meter, R. et al., "Derived Virtual Devices: A secure Distributed File system", Proc. Of the 5th NASA Goddard conference on Mass Storage Systems nd Technologies, September 1996.

[15] Neuman B.C. and Ts'o, T., "Kerberos: An Authentication Server for Computer Networks", IEEE Communications, Volume 32, Number 9, pages 33-38, September 1994.

[16] Kaufman, C., Perlman, R., Speciner, M., "Network Security", Prentice Hall, 1995.

[17] Krawczyk, H., et al., "HMAC: Keyed-Hashing for Message Authentication", RFC 2104, Network Working Group, February 1997.

[18] O'Keefe, M.T., "Shared File Systems and Fibre Channel", Proc. Of the 6th NASA Goddard on Mass storage and Technologies, March 1998.

[19] Soltis, S.R., et al., "The Global File System", Proc. Of the 5th NASA Goddard on Mass Storage and Technologies, September 1996.
[20] Sun Microsystems, "NFS: Network File System Protocol Specification", RFC 1094, Network Working Group, March 1989.
[21] Campbell, R., "Managing AFS: The Andrew File System", Prentice Hall, February 1998.
[22] Riedel, E. et al., "Active Storage for Large-Scale Data Mining and Mulimedia", Proc. Of the 24th VLDB, 1998.
[23] Acharya, A., Uysal, M., and Saltz, "Active Disks", Technical Report TRC98-06, March 1998.

A Statistical Approach to Situation Assessment

Daniel McMichael*
The Centre for Sensor Signal and Information Processing
Signal Processing Research Institute Building
Mawson Lakes, South Australia, 5095, Australia
e-mail: dwm@cssip.edu.au
web: http://www.cssip.edu.au/~dwm

Abstract

Situation assessment is the discrimination of concise summary descriptions of the state of uncertain parametric models. This paper considers the problem of providing summary descriptions of dynamic state histories of multitarget models. It shows that under the product-of-marginals approximation of the posterior distribution of these models it is possible to evaluate the likelihood of asynchronous multistage situations involving multiple interacting elements. Such situations include individual and multiple target manoeuvres.

A dynamic programming algorithm is described that solves the problem of associating the stages of multiple situation elements with discrete time instants. The search space is much reduced by reparameterising the problem as an optimisation of the interaction times. The problem of associating tracks with situation elements is solved by selective enumeration. Methods are provided for eliminating a priori *the vast preponderance of uninteresting possibilities, so that they never need to be calculated.*

Keywords: situation assessment, tracking, association, dynamic programming, viterbi algorithm.

1 Introduction

A situation assessment is a summary of a complicated dynamic system appropriate for high level decision making. The methods used in situation assessment have tended to be informal, and are usually based on various techniques for automated reasoning. However, methods are beginning to emerge that enable such inferences to be computed by applying rigorous probabilistic approaches. Koller and Pfeffer have proposed a manual design scheme, the *object oriented Bayesian network* [1]. This idea provides a mechanism for packaging Bayesian networks into subnetworks so that much unnecessary design detail can be avoided. However, the key algorithmic question is how to construct such networks automatically in real-time [2].

The purpose of this paper is to propose statistical models suitable for inferring complex behaviours and situations from *tracks* of one or more targets generated from reports from one or more sensors. These models can be interpreted as variable structure Bayesian networks; specialised inference schemes are required.

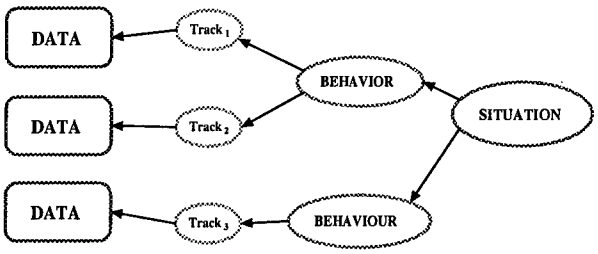


Figure 1: *A situation picture expressed in terms of tracks, behaviours and situations.*

Situations are assessed at two levels: target behaviours, and patterns of behaviours. Figure 1 shows the Bayesian network of a particular association of a situation to behaviours and of behaviours to tracks and (without showing any detail) of tracks

*The authors gratefully acknowledge the support of the Defence Science and Technology Organisation of Australia, and of the Sir Ross and Sir Keith Smith Fund.

to data. Algorithms for automatic track/data association have been in practical use for many years [3, 4, 5, 6], but higher-level associations have been assigned manually. To facilitate automatic methods, the problem can be parameterised and transformed into a problem of statistical inference over the space of association parameters. Such parameters are shown in figure 2. They switch the interaction between parents and their possible children. We refer to Bayesian networks that contain such association variables, which reside in the network, but also determine its effective structure, as *association networks* (section 3.2). Such structures are an example of those with state dependent structure investigated by Boutilier *et al.* [7]. The situation

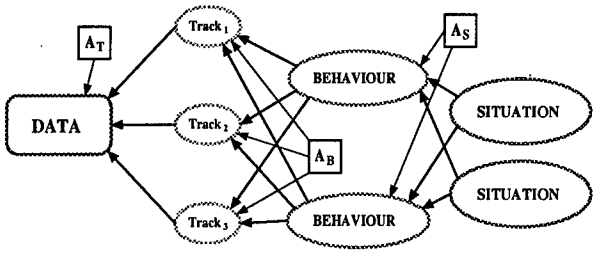


Figure 2: *Introducing explicit association variables to parameterise which data, tracks, behaviours and situations are related.*

assessment problem is therefore one of marginalising over association networks to find the marginal probabilities of the occurrence of a set of candidate situations. We address the problem by approximation. We evaluate the maximum likelihood behaviour/track associations, marginalise approximately over the track states, and then reapply the same approach at the situation/behaviour level. In the technical sections below, the exposition is simplified by omitting the behaviour level, and situations are inferred direct from the posterior distribution of the tracks.

2 A Process Model

Situations comprise *situation elements* which are entities (like targets) that may interact. Situations proceed by discrete stages $i = 1, ..., N_s$ and are modelled by tied state transition diagrams, such as that shown in figure 4. The state trajectory of element e during stage i is $Y_e(i)$. The ties indicate interaction between the states of the elements. Situations can also be represented by Bayesian networks. These networks can be associated with tracks or behaviours to form an overall model.

Figure 3 shows a situation/behaviour that takes place on a physical network of roads. Three elements are involved; a vehicle to supply fuel (dashed), which rendezvous with the other two attack vehicles at points 1 and 2 respectively; the latter two (black and grey) approach the target and escape afterwards along different routes.

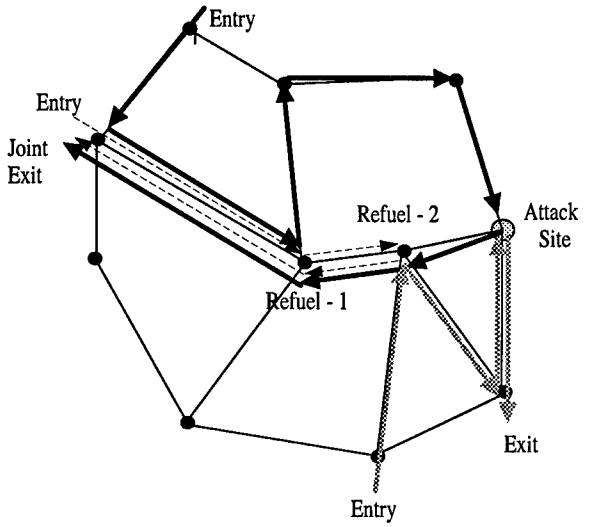


Figure 3: *A situation on a physical network: three elements are involved; one to supply fuel (dashed), which rendezvous with the other two elements at points 1 and 2 respectively; the latter two (black and grey) approach the target and escape afterwards along different routes.*

This physical event sequence is shown in abstraction in figure 4. The ties indicate element interactions for refuelling and attack.

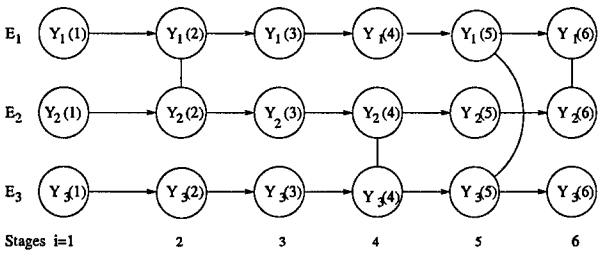


Figure 4: *An example of a process model for situa-*

tion elements corresponding to figure 3.

3 Extracting Situations from Posterior Distributions of Target Tracks

Let Z^k be the available batch data up to time k from all sensors. Let the state histories of all the targets up to time k be X^k, i.e.,

$$X^k = \{\{X_j(t)\}_{t=0}^{k}\}_{j=1}^{N_T}, \quad (1)$$

where j is the target index, N_T is the total number of targets. Although it is not a requirement of the algorithms described, it is notationally convenient that we assume that there is no target birth or death.

The posterior distribution $p(X^k|Z^k)$ is the distribution of the set of target histories given the data up to time k. The process of tracking approximates this distribution by its product-of-marginals:

$$p_m(X^k|Z^k) = \prod_{j=1}^{N_T}\prod_{t=0}^{k} p(X_j(t)|Z^t) \quad (2)$$

Usually, trackers further approximate $p(X_j(t)|Z^t)$.

Formally $p_m(X^k|Z^k)$ can be generated either by direct marginalisation or by iterative proportional scaling. In practice, $p(X^k|Z^k)$ is never calculated.

3.1 Notation for Situations

Let a situation S contain elements $E_1, ..., E_{N_E}$. Let $Y_e = \{Y_e(i)\}_{i=0}^{N_s}$ where i is the stage of the situation, e is the index of the stage in the situation, $Y_e(i)$ is the state of target e at stage i of the scenario. Let $Y = \{Y_e\}_{e=1}^{N_E}$. Note that many discrete time instants may be associated with the same stage.

3.2 Association Networks

The associations between situation states and track states can be represented within a single Bayesian network structure as an association network.

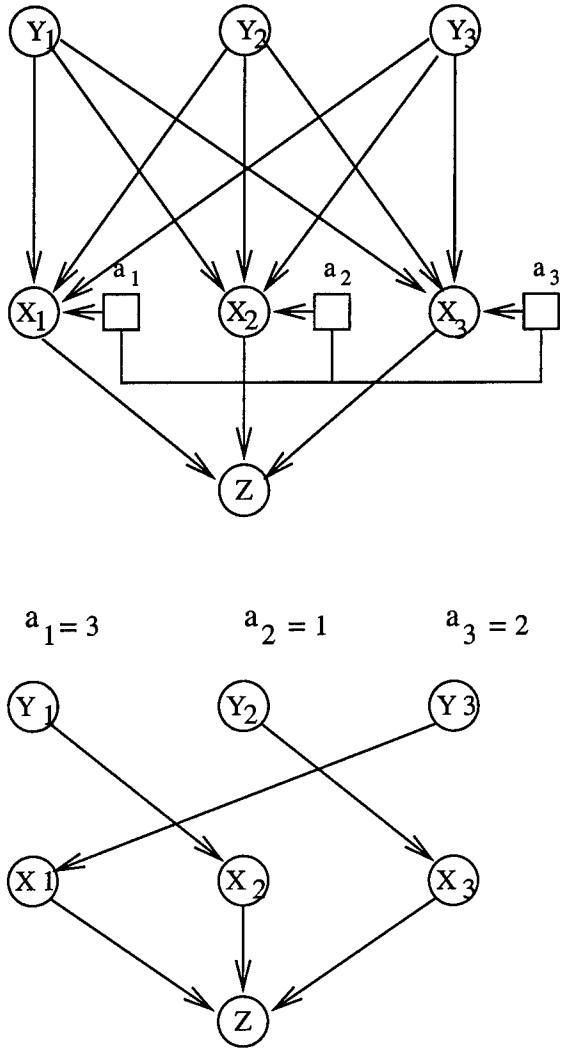


Figure 5: *A fragment of an association network for relating the states of three tracks to the states of three situation elements. The lower diagram shows the Bayesian network in the case when* $a_1 = 3$, $a_2 = 1$, *and* $a_3 = 2$. *The time variable is not introduced, for simplicity of presentation.*

Association networks contain nodes that determine the operative links of the Bayesian network. These links are associations, for example between tracks and situation elements. Figure (5) (top) shows a typical association network. If $a_1 = 3$, $a_2 = 1$, and $a_3 = 2$, the operative network reduces to that shown in figure (5) (bottom).

3.3 The likelihood of a Situation

To carry out inference we need the likelihood function of the situation S given the data Z^k. This re-

quires that we eliminate both the track states X^k and the situation states Y. Formally, we may write:

$$p(X^k|S) = \int p(X^k|Y)p(Y|S)dY, \qquad (3)$$

and thereby evaluate the

$$\begin{aligned} p(Z^k|S) &= \int p(Z^k|X^k)p(X^k|S)dX^k \\ &= \int \frac{p(X^k|Z^k)p(Z^k)}{p(X^k)}p(X^k|S)dX^k \\ &= p(Z^k)\int \frac{p(X^k|Z^k)}{p(X^k)}p(X^k|S)dX^k. \end{aligned} \qquad (4)$$

We parameterise the distribution $p(Y|S)$ by the set of variables Γ_S to obtain $p(Y|\Gamma_S)$. If the situation contains no cross chain links, and all the stages are independent

$$p(Y|\Gamma_S) = \prod_{e=1}^{N_E}\prod_{i=1}^{N_s} p(Y_e(i)|\gamma_{ei}), \qquad (5)$$

where γ_{ei} are situation parameters and $\Gamma_S = \{\{\gamma_{ei}\}_{e=1}^{N_E}\}_{i=1}^{N_s}$. We will examine situation inference in this case first, and then return to the problem of inferring situations with interacting elements.

We need to express e and i in terms of j and k. To do this we introduce the association variables a_j and b_k, where a_j is the association element associated with target j, and b_k is the situation stage associated with time k.

Where a track is not associated with a target, $a_j = 0$; $X_0(i)$ is a null state associated with nuisance tracks. Integrating out the situation states:

$$p(X^k|\Gamma_S) = \prod_{j=1}^{N_T}\prod_{t=0}^{k} p(X_{a_j}(b_t)|\gamma_{a_j b_t}). \qquad (6)$$

If the prior distribution of the track states is conditioned on parameters Θ then we may write $p(X^k)$ as $p(X^k|\Theta)$ and

$$p(X^k|\Theta) = \prod_{j=1}^{N_T}\prod_{t=0}^{k} p(X_j(t)|\Theta), \qquad (7)$$

and, applying the product-of-marginals approximation (2), equation (4) becomes

$$p(Z^k \mid S, a, b, r_0, \Gamma_S, \Theta) \approx p(Z^k)\prod_{j=1}^{N_k}\prod_{t=0}^{k} \int \frac{p(X_j(t)|Z^t)p(X_j(t)|\gamma_{a_j b_t})}{p(X_j(t)|\Theta)}dX_j(t). \qquad (8)$$

This integral is large if the track state distributions generated from the data and those derived from the situation template overlap. The "overlap" function is defined to be

$$O(S,j,t,e,i,\Gamma_S,\Theta) = \int \frac{p(X_j(t)|Z^t)p(X_j(t)|\gamma_{ei})}{p(X_j(t)|\Theta)}dX_j(t). \qquad (9)$$

Such functions are relatively easy to calculate, especially if the track states are Gaussian and/or discrete. The log-likelihood of the data is therefore

$$\log p(Z^k \mid S, a, b, \Gamma_S, \Theta) \approx \log p(Z^k) + \sum_{j=1}^{N_T}\sum_{t=0}^{k} \log O(S, j, t, a_j, b_k, \Gamma_S, \Theta). \qquad (10)$$

This likelihood is conditioned on the time–stage and track–element association variables a and b. To facilitate a fast algorithm we require that the situation parameters, Γ_S to be known *a priori*: soft situation templates are formulated in advance.

4 A Dynamic Programming Algorithm for Situation Assessment

The product-of-marginal approximation yields the log situation likelihood, $\log p(Z^k|S, a, b, \Gamma_S, \Theta)$, in a decomposable form to which dynamic programming can be applied. We consider a lattice in which the state at each time k is the stage of the situation. (In this treatment, we assume that all the situation elements transition between stages together) For each element-track configuration a, we maximise over b by dynamic programming.

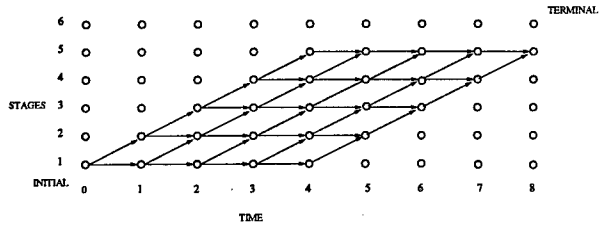


Figure 6: *Possible stage transitions subject to the order constraint.*

The *order constraint* imposes a massive reduction in the computational load: instead of requiring $N_s^2 k$ calculations of the overlap functions, only $(2N_s - 1)k$ such calculations are needed. If known, initial and terminal conditions can likewise save large amounts of computation.

The optimum path maximises $p(Z^k|S, a, b, \Gamma_S, \Theta)$ over b; it is found as follows:
Forward Phase: At each stage, find the maximum partial sum $\log p(Z^t|S, a, b, \Gamma_S, \Theta)$ at time t. Do this recursively by finding the maximum, for each value of i, of the log–likelihood by enumerating all possible transitions leading to state i. There are *at most two* (see figure 6).
Backward Phase: From the terminating condition backtrack down the trellis, recording the optimum path.

The target-element associations still have to be enumerated.

4.1 Prior Deletion

Large numbers of possible values of a can be excluded by doing a prior evaluation of each possible track–element association [it may be sufficient to use simple approximate methods for this operation]. All implausible pairwise associations can be deleted. The result of this thinning is a reduced set of possible targets that are candidates for association with each track element. If the set of possible tracks for element E_e is $\mathcal{T}_e$, the total number of target-element hypotheses is $\leq |\mathcal{T}_1|.|\mathcal{T}_2|.....|\mathcal{T}_{N_E}|$.

5 Interacting Situation Elements

In this case, the function $p(Y|\Gamma_S)$ does not factorise into separate terms for each element. Interacting groups of elements generate single factors.

At time t let the N_T elements have g_t interacting groups and let the m-th of these have state $X_m(t)$, which is the concatenation of all the states of the elements in the m-th interaction group. Hence

$$p(Z^k \mid S, a, b, \Gamma_S, \Theta) = p(Z^k) \prod_{t=0}^{k} \prod_{m=1}^{g_t} \int \frac{p(X_m(t)|Z^t) p(X_m(t)|\Gamma_S, a, b)}{p(X_m(t)|\Theta)} dX_m(t). \tag{11}$$

Aside from this small change, necessitating that the chains for all the elements be evaluated as one, the dynamic programming algorithm for b is unchanged.

5.1 Asynchronous Interacting Situation Elements

So far, we have only considered the case in which the stages of all the situation elements are synchronised. We now relax this constraint.

The associations between situation element stages and times may be different for each element. The stage-time association variable b_k is now also indexed by the situation element to which it refers, and becomes b_{ek}. Equation (10) becomes:

$$\log p(Z^k \mid S, a, b, \Gamma_S, \Theta) \approx \log p(Z^k) + \sum_{j=1}^{N_T} \sum_{t=0}^{k} \log O(S, j, t, a_j, b_{a_j k}, \Gamma_S, \Theta). \tag{12}$$

If elements do not interact, the chains for each of the elements can be solved separately. The target-element associations a are decided using ML or MAP criteria by enumeration after prior deletion.

Sparsely interacting elements pose the richest problem. Consider the situation with two elements shown in figure 7. To avoid needless complexity in presentation, we require that: *interacting stages begin and end simultaneously*. This constraint can be relaxed without difficulty. The trellis diagram becomes 2-dimensional, as the state space at each time is the pair of stages of the two elements. We identify the possible stage transitions on a rectangle with arrows pointing from states at time t to

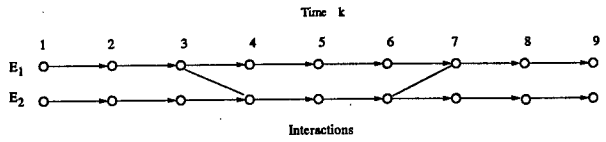


Figure 7: *Transitions and interactions for a two-element situation with two physical interactions.*

states at time $t + 1$. Figure 8 shows the permitted transitions. Self transitions are also allowed.

A solution could be obtained by extending the state space of the process accordingly (to the Cartesian product of the stage spaces of all the elements). This approach leads to representation of vast numbers of unreachable states. Alternatively, the problem could be resolved onto a more efficient representation.

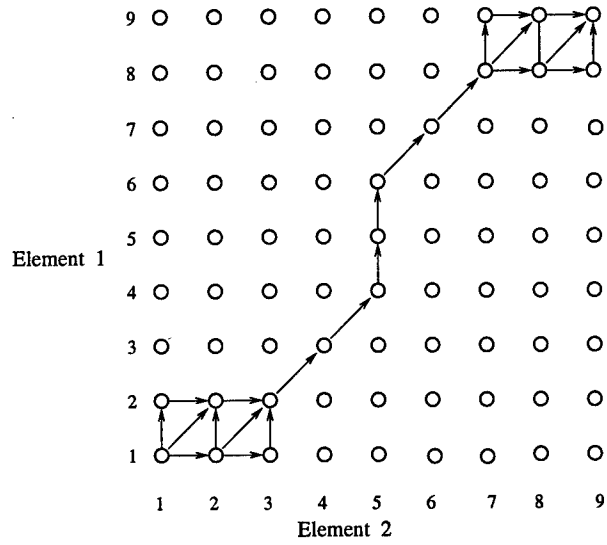


Figure 8: *Permissible transitions for the transition diagram of figure 7.*

We can easily recognise that the majority of the transition possibilities are generated by the sections where the elements are independent. However these sections can be handled efficiently by a single element dynamic programming algorithm. Unfortunately, the beginnings and endings of these sections are unknown in time, but they can easily be found.

5.2 An Algorithm

1. Evaluate optimum trellis paths and corresponding contributions to the overall criterion for all possible start and end times of each independent section.

2. By enumeration or otherwise, find the optimum set of start times for each interacting stage of the situation e.g., $[T_1\ T_2\ \ldots\ T_n]$.

3. Back track and apply dynamic programming for the intervening independent sections to complete the entire optimum allocation of stages of all elements to times.

5.3 Prior Deletions

1. Remove impossible pairings of targets to elements by reference to the observed track interactions. Delete target-element associations where any such pairings occur.

2. Remove impossible element-target associations by first marginalising the element models and testing as in section 4.

6 Evaluating the Overall Situation

The situations we have discussed are summary descriptions of subsections of the observed data. We have seen how to evaluate their likelihood approximately. Discriminating between multiple situations is achieved by simple Bayesian inference using these approximate likelihoods:

$$p(S|Z^k,\Gamma,\Theta) = \frac{P(Z^k|S,\Gamma_S,\Theta)P(S)}{\sum_{\{\tilde{S}\}} P(Z^k|\tilde{S},\Gamma_{\tilde{S}},\Theta)P(\tilde{S})}, \quad (13)$$

where $\Gamma = \{\Gamma_S\}_{\{S\}}$. This Bayesian solution is approximated by substituting $P(Z^k|S,\Gamma,\Theta)$ by $P(Z^k|S,\hat{a}_S,\hat{b}_S,\Gamma_S,\Theta)$, where the $\hat{a}_S$ and $\hat{b}_S$ are estimates. Appropriate distributions $p(X_j(k)|S)$ are required for the tracks that are not associated with a situation element. Without these $P(Z^k|S,\Gamma,\Theta)$ cannot be calculated.

7 Discussion

The product-of-marginals approximation facilitates the use of dynamic programming to evaluate large numbers of time-stage associations efficiently. However, consideration of even simple scenarios, indicates that computational speeds may still be too slow. It appears that the key to making the procedures described here run fast is the elimination of the vast majority of possible association hypotheses from consideration. Various rules for prior deletion have been proposed.

The product-of-marginals approximation is likely to be significantly wrong, but although the situation probabilities will be over–estimates, it also appears likely that discrimination of the most likely situations will be possible.

Acknowledgment

The author is most grateful to Dr. Nickens Okello for his helpful comments and, not least, for typesetting much of the text and drafting many of the diagrams.

References

[1] D. Koller and A. Pfeffer. Object-oriented bayesian networks. In *Proceedings of the 13th Annual Conference on Uncertainty in AI*, pages 302–313, Providence, RI, 1997.

[2] D.W. McMichael. Statistical models for situation awareness. invited talk, Information Decision and Control 99, Adelaide, Australia, February 1999.

[3] S. B. Blackman. *Multiple-target tracking with radar application.* Artech House, Norwood, MA, 1986.

[4] Y. Bar-Shalom and X.-R. Li. *Estimation and tracking: principles, techniques and software.* Artech House, 1993.

[5] Y. Bar-Shalom and X.-R. Li. *Multitarget--multisensor tracking: principles and techniques.* YBS, 1995.

[6] R.L. Streit, editor. *Studies in probabilistic multi-hypothesis tracking and related topics*, volume SES-98-01. US Naval Undersea Warfare Centre Division, Newport, Rhode Island, 1998.

[7] C. Boutilier, N. Friedman, M. Goldszmidt, and D. Koller. Context-specific independence in Bayesian networks. In E. Horvitz and F. Jensen, editors, *Proc. Twelfth Annual Conference on Uncertainty in Artificial Intelligence*, pages 115–123, Portland, Oregon, August 1996. Morgan Kaufmann.

Dempster-Shafer Belief Propagation in Attribute Fusion

P. Korpisaari and J. Saarinen
Signal Processing Laboratory
Tampere University of Technology
P.O.Box 553
33101 Tampere
Finland
Email: pkorpi@cs.tut.fi
Tel: +358 3 3653392
Fax: +358 3 3653095

Abstract ***A complete evidential mapping approach in the contect of Dempster-Shafer reasoning is applied to target identification task in the case of hierarchically dependent attributes. The mapping defines causal links for all possible evidence sets. Such causal links are then used to propagate evidencies through variable chains. Finally, the independent evidencies propagated to aircraft type node are combined with Dempster's rule of combination. The complete evidential mapping is compared to more common many-valued mapping by analyzing the identification probabilities after detection serieses.***

Keywords: Dempster-Shafer reasoning, target identification

1 Introduction

We address a target identification problem in a case of hierarchically formed attributes. A root variable of the hierarchical attribute tree is aircraft type. This attribute is of particular interest in the case of target identification. Other nodes are attributes which explain the root node. These other nodes may have indirect impact to the type of the target. The problem is to infer the type of the target (root node) given attributes (descendant nodes).

A common approach is to use the Dempster-Shafer theory for reasoning the target type. Dempster introduced concepts of lower and upper probabilities in his pioneering work [1] which was refined by Shafer [2]. These works generated an inferencing method known Dempster-Shafer reasoning which has got great interest in area of approximate reasoning and its applications.

A frame of discernment contains all given attributes. The elements of the frame of discernment are the leaf attributes of the tree. All ancestor attributes can be described as unions of these attributes. This approach assumes multivalued mappings in the attribute tree i.e. causal matrices that link the descendant attributes to their ancestors contain only ones and zeros. By using this approach to describe to attributes' hierarchical structure the conventional Dempster's rule of combination can be used for mass function evaluations. Such an approach is described by Bogler [4] who applies Dempster-Shafer reasoning for target identification. This approach is commented by Buede [5] who finally gives a comprehensive comparision of the Dempster-Shafer reasoning and Bayesian reasoning approaches in the context of target identification [6]. A Bayesian approach is represented also in [3].

We apply the complete evidential mapping approach described by Liu et.al. [17]. This approach enables soft causal links between the sets. It also defines additional causal

links to complete the causal links to all possible evidence sets.

2 Attribute Fusion

A motivation of attribute fusion is twofold:

- identification of target's aircraft type
- resolving ambiguities in observation-to-target associations

In this paper we concentrate on the target identification process. This process can be understood as a part of the whole data-association scheme.

2.1 Target Identification

Target identification concerns determination of the target's aircraft type. The identification can be done by direct aircraft type detections or with supporting attribute detections. Attribute detections do not identify directly the aircraft type. They induce a set of possible aircraft types that could have the detected property. The direct aircraft type detections and the attribute detections include uncertainties and inconsistencies which makes the identification problem untrivial.

2.2 Attributes' Hierarchy

Attributes' internal relationships define a hierarchical attribute structrure. The leaf attributes can be used to explain their ancestor attributes' values. Furthermore, these ancestors may explain some other attributes. Finally, this attribute structure converges to aircraft type which is a root node of the hierarchical attribute tree.

We investigate two different kind of hierarchical structures:

- naive trees
- simple hierarchical trees

In naive trees the root node is aircraft type and all other attributes are direct descendants of the root node. Simple hierarchical trees may contain several layers i.e. the direct descendants of the aircraft type node (root node) may have own descendants.

3 Dempster-Shafer Reasoning

Dempster-Shafer method combines two different mass functions m_1 and m_2 into one mass function m^*. The combination is carried out based on seeking consensus between the mass functions. Given two sets D and E and their mass function values $m_1(D)$ and $m_2(E)$ a joint mass function value is $m_1(D)m_2(E)$ which is assigned to a set C that contains all elements common to D and E, i.e. $C = D \cap E$. Finally, masses of empty intersections are assigned to zero and the remaining non-zero mass function values are normalized to one. The following formula presents the idea described above and it is known as a Dempster's rule of combination:

$$m(C) = \frac{\sum_{D_i \cap E_j = C} m_1(D_i)m_2(E_j)}{1 - \sum_{D_i \cap E_j = \emptyset} m_1(D_i)m_2(E_j)} \quad , \quad (1)$$

where summations over joint mass functions indicates that the final mass function value $m(C)$ is a sum of all joint masses of intersections equal to C.

3.1 Multivalued Mappings

Both sets D and E have to belong to a same power set 2^{θ_X}. Thus, they describe the same discrete variable X. Dempster's rule of combination does not directly apply to reasoning from one variable to another. One way to overcome this is to map a set in 2^{θ_X} for all possible values y_i of Y:

$$\mathbf{f} : y_i \rightarrow D_j \;\; ; \; y_i \in \theta_Y \, , \, D_j \in 2^{\theta_X} \quad . \quad (2)$$

Such a mapping is called as a multivalued mapping. It induces a set D_j for y_i which may present a detection. The mapping assigns a mass equal to y_i:s mass to D_j; $m(D_j) = m(y_i)$. Since D_j is defined in 2^{θ_X} its mass $m(D_j)$ can be combined with other mass functions in 2^{θ_X}. Hence, this approach enables reasoning for X given a value of another variable Y.

Multivalued mapping links together two frames of discernments T and Θ. Here T is a set of attribute's possible values and Θ is

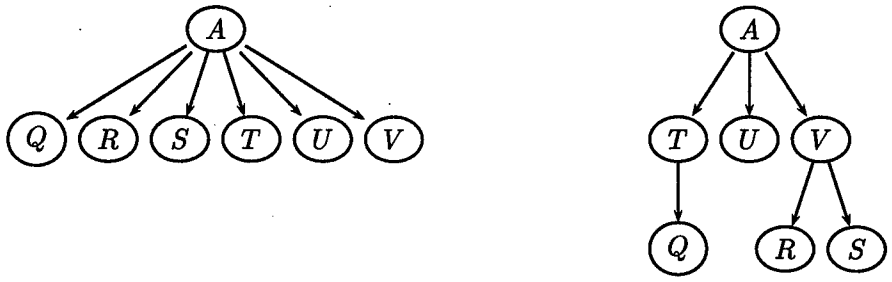


Figure 1: A naive tree and a simple hierarchical tree.

Table 2: Mapping from attributes to sets of aircraft types.

		Sets of aircraft types $A \in 2^{\Theta}$			
Attributes		$\{a_1, a_2\}$	$\{a_1, a_3\}$	a_2	$\{a_2, a_3\}$
T	t_1	1	0	0	0
	t_2	0	0	1	0
	t_3	0	0	0	1
U	u_1	0	1	0	0
	u_2	0	0	0	1

a set of all aircraft types. Multivalued mapping defines for each element $t \in T$ corresponding subsets in Θ. It is assumed that these subsets are not empty. There is no such value $t \in T$ which has not even one corresponding element in Θ. In other words the set of possible attribute values T does not contain such a value that is impossible for the given set of aircraft types Θ. This kind of mapping represents the information that defines all possible attribute values $t \in T$ for each element (aircraft type) $a \in \Theta$. One example of aircraft types' possible properties is given in Table 1. '1' indicates that the attribute on the row is possible for the aircraft type on the column.

An attribute detection yields a probability distribution $P(T)$ for the corresponding attribute T. Additionally a confidence value of the detection itself will be taken into account. This is done by addressing a certain amount of mass to the whole frame of discernment of the corresponding attribute. The mass $m(T)$ equals to $1 - c$ where c is a given confidence value of the detection. Thus the detection induces two focal elements; one for the subset A corresponding to the detection and one for the whole frame of discernment Θ.

3.2 Evidential Mapping

Attribute T has N possible values:

$$T = \{t_1, t_2, \ldots, t_N\} \tag{3}$$

A subset $T_i \in 2^{\Theta_T}$ contains n values from T:

$$T_i = \{t_{i_1}, t_{i_2}, \ldots, t_{i_n}\} \tag{4}$$

A basic matrix (BM) is a matrix that links each single attribute value $t_i \in T$ to sets of aircraft types. An element of BM is denoted as $b(t_i, A_j)$ and it describes a strength of the causal link between attribute value t_i and the set $A_j \in 2^{\Theta_A}$.

$$\begin{array}{c} \\ t_1 \\ t_2 \\ t_3 \end{array} \begin{array}{c} \begin{array}{cccc} A_1= & A_2= & A_3= & A_4= \\ \{a_1,a_2\} & \{a_3\} & \{a_1,a_3\} & \{a_1,a_4\} \end{array} \\ \begin{bmatrix} 0.1 & 0.9 & 0 & 0 \\ 0.4 & 0 & 0.6 & 0 \\ 0.1 & 0 & 0.8 & 0.1 \end{bmatrix} \end{array} \tag{5}$$

An image of an attribute value t_i in Θ_A is a set of aircraft types that may possibly have

Table 1: Aircraft types' properties for attributes Q, R, S, T, U and$V$

		Aircraft types θ_A			
Attributes		a_1	a_2	a_3	a_4
T	t_1	1	0	1	0
	t_2	0	1	1	1
U	u_1	1	1	1	1
	u_2	1	1	0	1
	u_3	1	1	0	0
	u_4	1	1	0	1
V	v_1	1	0	0	1
	v_2	0	1	0	1
	v_3	0	0	1	1
Q	q_1	1	0	1	0
	q_2	1	0	0	0
	q_3	0	0	0	0
	q_4	0	1	1	1
	q_5	1	1	1	1
R	r_1	1	1	1	1
	r_2	0	0	1	1
	r_3	0	0	1	0
S	s_1	1	0	0	1
	s_2	0	1	0	1
	s_3	0	1	1	1
	s_4	0	0	1	0

attribute value t_i. Image of t_i is denoted as A^i:

$$A^i = \bigcup_{j=1}^{r} \{A_j \ ; \ b(t_i, A_j) > 0\} \qquad (6)$$

Construction of complete evidential mapping matrix (CEM)has following steps:

1. Expand BM so that its rows contain all subsets $T_i \in 2^{\Theta_T}$ except the empty set $\emptyset$.

2. Copy causal links in BM directly to CEM.

A set of causal links from single attribute values $t_{i_k} \in T_i$ to one set of aircraft types A_j, $j = 1 \ldots r$ is denoted as M_{ij}:

$$M_{ij} = \{b(t_{i_k}, A_j)\}_{k=1}^{n} \qquad (7)$$

For all pairs of newly created sets T_i and sets A_j, $j = 1 \ldots r$ that are defined in the basic matrix perform the following steps.

1. Calculate average of the causal links M_{ij}. Denote this average with b^*_{ij}:

$$b^*_{ij} = \frac{1}{n} \sum_{k=1}^{n} b(t_{i_k}, A_j) \qquad (8)$$

2. • If all causal links in M_{ij} are nonzero establish a causal link from T_i to A_j. A strength of this link $c(T_i, A_r)$ is b^*_{ij}.
• Otherwise, if at least one of the causal links in M_{ij} is zero define a set of aircraft types A_r which is a union of the image sets of all single attribute values in T_i:

$$A_s = \bigcup_{t_{i_k} \in T_i} A^{i_k} \qquad (9)$$

Add a column into CEM for A_s if such does not exist yet. Set $c(T_i, A_r) = b^*_{ij}$. If a column for A_s exists already set $c(T_i, A_r) \leftarrow c(T_i, A_r) + b^*_{ij}$.

A complete evidential mapping matrix (CEM) that is produced from the basic matrix given on the previous page is the following:

	A_1	A_2	A_3	A_4	A_5	A_6
t_1	0.1	0.9	0	0	0	0
t_2	0.4	0	0.6	0	0	0
t_3	0.1	0	0.8	0.1	0	0
$\{t_1,t_2\}$	0.25	0	0	0	0	0.75
$\{t_1,t_3\}$	0.1	0	0	0	0.9	0
$\{t_2,t_3\}$	0.25	0	0.7	0	0.05	0
$\{t_1,t_2,t_3\}$	0.2	0	0	0	0.8	0

A_1=$\{a_1,a_2\}$
A_2=$\{a_3\}$
A_3=$\{a_1,a_3\}$
A_4=$\{a_1,a_4\}$
A_5=$\{a_1,a_2,a_3,a_4\}$
A_6=$\{a_1,a_2,a_3\}$

3.3 Belief Propagation

Complete evidential mapping matrices are used to propagate beliefs from the evidence node to all its ancestors [17]. The belief is propagated on a basis of sets rather than single elements. The complete evidential mapping matrix defines a belief mapping to hypothesis space for all possible evidence sets.

Thus, CEMs can be used to propagate beliefs through a chain of variables. The leaf variable induces sets to its parents and these sets are mapped to third variable using the corresponding CEM matrix.

Belief propagation procedure contains the following steps:

1 Create basic matrix (BM) from the set of known rules.

2 Construct complete evidential mapping matrix (CEM) from BM.

3 CEM is used for belief propagation based on the given propagation rules.

A propagation through two CEM matrices can be simplified to one CEM matrix mapping. Let C_1 be a CEM matrix from X to Y and C_2 be a CEM matrix from Y to Z. One set of Y-values is assigned to each column of C_1 and hence C_1 introduces a set of sets in Y. The rows of C_2 are defined for all possible sets in Y. A new matrix C_2' is formed based on C_2 by picking the rows corresponding to sets induced by C_1. Now there exists a row in C_2' for each column of C_1. Further it is assumed that the rows of C_2' are ordered similarly to the column order of C_1. Now the belief propagation from X to Z through Y can be expressed by one single CEM matrix instead of two cascaded CEM matrices. This CEM matrix performing the mapping is a matrix product of the two separate matrices:

$$CEM_{X \to Z} = C_1 \times C_2' \qquad (10)$$

This principle extends easily to a chain of variables where the resulting CEM matrix is a product of several CEM matrices.

4 Simulations

In the simulations the aim was to compare the target identification performance of the two represented methods. Simulations included 50 different detection series which contained observations from seven different attributes. The causal relationships obey the hierarchical tree illustrated in Fig. 1. It was assumed that number of different target types is 4. Each target type had own parameters in the attribute tree. All attributes are discrete and their number of possible values is between 2-5. The detection distributions for hierarchical tree are shown in Fig. 3. Each column corresponds to aircraft type and each row is one of the seven attributes. The top row describes the aircraft type itself and thus the distributon on the top are the direct aircraft type distributions.

Each simulation run produced a belief distribution to aircraft type node. These distributions describe the identification probabilities. Two distinct analysis were carried out to describe the identification performancies. First, one aircraft type with largest belief value was selected as an identified type at end of the detection period. The identified type distributions were compared by picking the detection serieses made from the same aircraft type. These serieses should produce a distribution with very high peak at the correct aircraft type value. The distributions are illustrated in Fig. 4. The top row is a distribution for multivalued mapping and the bottom row is for complete evidential mapping.

Another analysis was made by collecting distribution of ith target type's probabilities at the end detection period against the correct target type. Such distributions are given in Figs. 5. and 6. The (i,j)th distribution in these tables describe how the probability of ith target is distributed when the detections are made from the jth target. Thus the diagonal elements of the table of distributions correspond to correct identifications.

References

[1] A. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *Annals of Mathematical Statistics*, vol. 38, pp. 325–339, 1967.

[2] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton University Press, 1976.

[3] P. J. Nahin and J. L. Pokoski, "NCTR plus sensor fusion equals IFFN or can

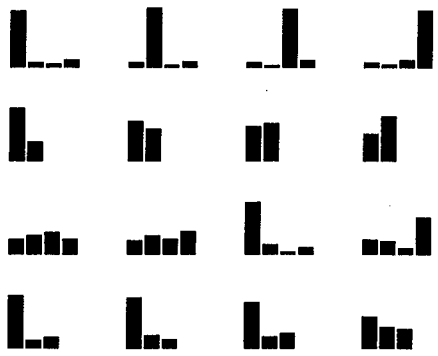


Figure 2: Distributions of attribute detections in a case of naive tree.

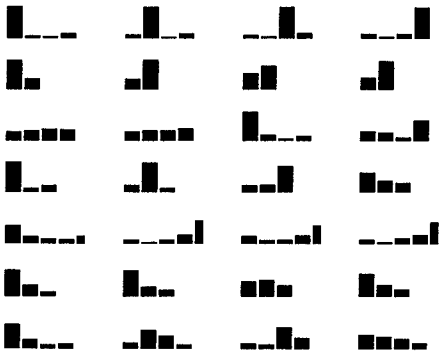


Figure 3: Distributions of attribute detections in a case of hierarchical tree.

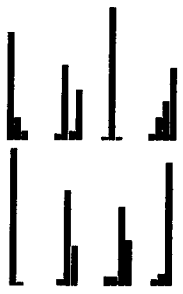


Figure 4: Distributions of identifications.

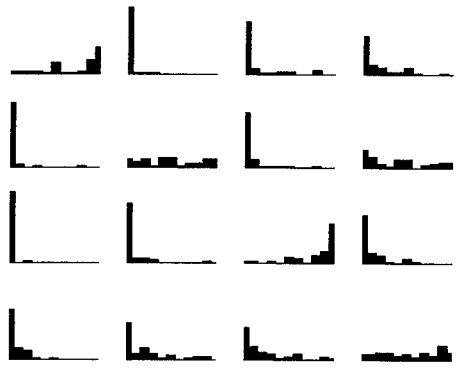


Figure 5: Distributions of identifications. Multivalued mapping.

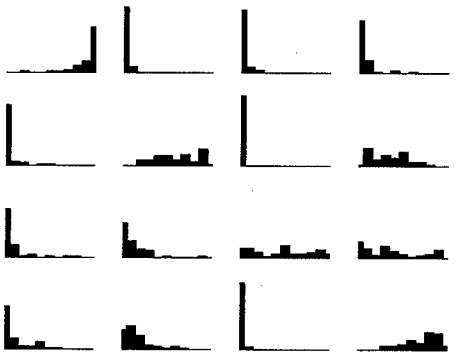


Figure 6: Distributions of identifications. Complete evidential mapping.

two plus two equal five?," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 16, pp. 320–337, May 1980.

[4] P. Bogler, "Shafer-Dempster reasoning with applications to multisensor target identification systems," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 17, pp. 968–1011, November 1987.

[5] D. Buede, "Shafer-Dempster and Bayesian reasoning: A response to "Shafer-Dempster reasoning with applications to multisensor target identification systems"," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 18, pp. 1009–1011, November 1989.

[6] D. Buede and P. Girardi, "A target identification comparision of Bayesian and Dempster-Shafer multisensor fusion," *IEEE Transactions on System, Man, and Cybernetics - Part A: Systems and Humans*, vol. 27, pp. 569–577, September 1997.

[7] H. Hau and R. Kashyap, "Belief combination and propagation in a lattice-structured inference network," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, pp. 45–57, January 1990.

[8] J. Chao, C. Cheng, and C. Su, "A moving target detector based on information fusion," in *Record of the IEEE 1990 International Radar Conference*, pp. 341–344, 1990.

[9] G. Shafer, "Perspectives on the theory and practice of belief functions," *International Journal of Approximate Reasoning*, vol. 4, pp. 323–362, 1990.

[10] G. Shafer, "Rejoinders to comments on "perspectives on the theory and practice of belief functions"," *International Journal of Approximate Reasoning*, vol. 6, pp. 445–480, 1992.

[11] L. Hong and A. Lynch, "Recursive temporal-spatial information fusion with applications to target identification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, pp. 435–445, April 1993.

[12] Z. Luo and D. Li, "Multi-source information integration in intelligent systems using plausibility measures," in *Proceedings of the 1994 IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems*, pp. 403–409, 1994.

[13] J.-B. Yang and M. Singh, "An evidential reasoning approach for multiple-attribute decision making with uncertainty," *IEEE Transactions on systems, Man, and Cybernetics*, vol. 24, pp. 1–17, January 1994.

[14] P. Valin, J. Couture, and M.-A. Simard, "Position and attribute fusion of radar, esm, iff and datalink for aaw missions of the canadian patrol frigate," in *Proceedings of the 1996 IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pp. 63–71, 1996.

[15] L.-W. Jang and J. Chao, "An information fusion algorithm for data association in multitarget tracking," in *Proceedings of the First Australian Data Fusion Symposium ADFS'96*, pp. 119–124, 1996.

[16] R. Murphy, "Dempster-Shafer theory for sensor fusion in autonomous mobile robots," *IEEE Transactions on Robotics and Automation*, vol. 14, pp. 197–206, April 1998.

[17] W. Liu, J. Hughes, and M. McTear, *Advances in the Dempster-Shafer Theory of Evidence*, ch. Representing heuristic knowledge and propagating beliefs in the Dempster-Shafer theory of evidence, pp. 441–471. John Wiley & Sons, Inc., 1994.

A Novel Minimal Norm Based Learning Subspace Method①

De-Shuang Huang

Beijing Institute of System Engineering, P. O. Box 9702-19, Beijing 100101, China
Tel: 86-010-66356616, Fax: 86-010-64879213, email: bisenet@bisel.bise.ac.cn

Abstract: *This paper proposes a novel* minimal norm based learning subspace method (*MNLSM*), *which can satisfy the requirements of being insensitive to the order of presentation of the training samples. This MNLSM is applied to recognition of simulating high resolution radar (HRR) targets (two for ships, one for chaff). Experimental results show that the performance of proposed MNLSM such as rate of correct recognition and convergence speed is satisfactory.*

Keywords: Self-Supervised Learning, Subspace, Pattern Recognition, Minimal Norm, Disordered Learning, Radar Targets.

1. Introduction

The *learning subspace method* (LSM) proposed by Kohonen in 1978[1], in essence, is an adaptive method of extracting principal components of pattern vectors from each class. This approach assumes the class labels for all input samples to be known, and uses Hebbian rule to update the basis vectors corresponding to each subspace. So it is also called the self-supervised neural network approach [2], which designs each subspace in terms of the label for each sample. However, the essential drawbacks to the LSMs are sensitive to the order of presentation of the input samples, in other words, the prior learned samples which might be recorded in the basis vectors of the corresponding subspace may be offset or forgotten by the learning of the late-coming samples, which leads to total performance decreasing [2, 3, 4, 6]. In 1982, E. Oja et. al proposed the *averaging learning subspace method* (ALSM)[3,4]which can avoid the sensitiveness to the order of presentation of the input samples. But it needs to compute three conditioned correlation matrices and their eigenvalue decompositions which leads to the convergence speed much decreasing[5]. To avoid or reduce the defects for those existing methods, this paper proposes a novel self-supervised learning subspace methods, called *minimal norm based learning subspace method* (MNLSM), which are not sensitive to the order of presentation of the input samples, and much improve the convergence speed.

This new LSM, to verify its validity, is applied to the recognition of high resolution radar (HRR) targets (two for simulating ships and one for simulating chaff). The experimental results support our claims.

2. A Novel Minimal Norm Based Learning Subspace Method

2.1 General Presentation

① This work was supported by Grants 69705001 from NSF and DSR of China

suboptimal) solutions. The training sample for each iteration needs to be selected not only "randonly" but also some specific criterion based from the training sample set. In fact, it is easily thought that for some class, the training sample with the minimal orthogonal projection length on its own subspace can be selected to design corresponding subspaces. Thus this method is called as *minimal norm based learning subspace method* (MNLSM). Fig. 1 depicts the scheme of the learning process for MNLSM.

2.3 Minimal Norm Based Learning Subspace Iterating Algorithm

Firstly, assume that, at the kth iteration, the sample with minimal norm from the$i$th subspace is selected as

$$\mathbf{x}_m^{(i)} = \text{argmin}\{\delta(\mathbf{x}_j^{(i)}) = (\mathbf{x}_j^{(i)T} P^{(i)} \mathbf{x}_j^{(i)})^{\frac{1}{2}} , j = 1,2,\cdots,N_i\} \quad (4)$$

where argmin$\{\cdot\}$ denotes the operator of selecting the training sample with minimal orthogonal projection length on its own subspace, $\delta(\mathbf{x}_j^{(i)})$ the orthogonal projection length of $\mathbf{x}_j^{(i)}$ on the ith subspace.

So, the sample with minimal norm $\mathbf{x}_m^{(i)}$ is used to learn its own subspace with the positive manner and learn other subspaces with the negative manner. According to Eq. (1), the iterating formulae for the MNLSM can be stated as follows

$$L_k^{(i)} = (I + \mu_m^{(i,i)} \mathbf{x}_m^{(i)} \mathbf{x}_m^{(i)T}) L_{k-1}^{(i)}, i = 1,2,\cdots,K \quad (5)$$

$$L_k^{(j)} = (I - \mu_m^{(j,i)} \mathbf{x}_m^{(i)} \mathbf{x}_m^{(i)T}) L_{k-1}^{(j)}, \quad j \neq i \quad (6)$$

Generally, the above learning process could be unlimitedly gone on, but after several iterations, the formed subspace might become stable. The iterating algorithm for the MNLSM is summarized as follows:

Algorithm **Minimal Norm Based Learning Subspace Iterating Algorithm**

Step 1 $k = 1$, select the dimensionality $p^{(i)}$, the termination accuracy η and learning coefficient $\mu^{(i)}(i = 1,2,\cdots,c)$, set the initial basis vectors of the c subspaces, and compute the orthogonal projection matrixes $P_k^{(i)}(i = 1,2,\cdots,c)$.

Step 2 for each pattern vector $\mathbf{x}_j^{(i)}$ of the ith class, compute its orthogonal projection length (norm) on its own subspace

$$\delta(\mathbf{x}_j^{(i)}) = (\mathbf{x}_j^{(i)T} P^{(i)} \mathbf{x}_j^{(i)})^{\frac{1}{2}} , i = 1,2,\cdots,c; j = 1,2,\cdots,N_i\}$$

Step 3 select the training sample with minimal norm from the training sample set of each class

$$\mathbf{x}_m^{(i)} = \text{argmin}\{\delta(\mathbf{x}_j^{(i)}) , i = 1,2,\cdots,c; j = 1,2,\cdots,N_i\}$$

Step 4 rotate its own subspace with the positive manner in terms of Eq. (5) and rotate other subspaces with the negative manner in terms of Eq. (6) using the training sample with minimal norm.

Step 5 compute the averaging orthogonal projection length, T_d, of all training samples with minimal norms $\mathbf{x}_m^{(i)}$ from the ith class on their own subspaces according to Eq. (3)

Step 6 if $T_d \geqslant \eta$, skip to **Step 8**; else, continue to **Step 7**.

Step 7 k=k+1, return to **Step 2**.

Step 8 stop.

Note that whole iteratinng process will be terminated until the termination accuracy η is satisfied.

3. Experimental Results

The simulating range profiles of radar tar-

Assume that the ith class ω_i has N_i pattern vectors $\{\mathbf{x}_k^{(i)} \in R^d,\ k = 1, 2, \cdots, N_i;\ i = 1, 2, \cdots, c\}$, respectively. The self-supervised learning subspace method proposed by Kohonen is stated as follows[1]

$$\begin{cases} L_k^{(i)} = (I + \mu_k^{(i)} \mathbf{x}_k^{(i)} \mathbf{x}_k^{(i)T}) L_{k-1}^{(i)} \\ L_k^{(j)} = (I - \mu_k^{(j)} \mathbf{x}_k^{(i)} \mathbf{x}_k^{(i)T}) L_{k-1}^{(j)} \\ (j \neq i = 1, 2, \cdots, c) \\ L_k^{(i)} = L[\mathbf{u}_1^{(i)}(k), \mathbf{u}_2^{(i)}(k), \cdots, \mathbf{u}_{p^{(i)}}^{(i)}(k)] \end{cases} \tag{1}$$

where $\mu_k^{(i)}$, $\mu_k^{(j)}$ are the positive learning coefficients. Generally, $|\mu_k^{(i)}| < 1/\|\mathbf{x}\|_k^2$ and $|\mu_k^{(j)}| < 1/\|\mathbf{x}\|_k^2$); T denotes matrix or vector transpose; $\mathbf{x}_k^{(i)} = [x_1^{(i)}(k), x_2^{(i)}(k), \cdots, x_d^{(i)}(k)]^T$; I is an unit matrix; $L_k^{(i)} = L[\cdot]$ indicates the ith subspace composed of $p^{(i)}$ basis vectors $\mathbf{u}_n^{(i)}(k)(n = 1, 2, \cdots, p^{(i)})$ at instant k.

It should be, however, noted that the basis vectors $\mathbf{u}_n^{(i)}(k)$ $(n = 1, 2, \cdots, p^{(i)})$ should be kept orthonormal in the learning process of the LSM. Usually, the orthonormal approach available is the Gram-Schmidt one. Supposed that the converged orthogonal projection matrices corresponding to c subspaces are $P^{(i)}(i = 1, 2, \cdots, c)$, respectively. For an arbitrary input vector $\mathbf{x}$, the classification rule of the self-supervised LSM for pattern recognition is that if[1]

$$\|P^{(i)}\mathbf{x}\|^2 = \mathbf{x}^T P^{(i)} \mathbf{x} = \sum_{j=1}^{p^{(i)}} (\mathbf{x}^T \mathbf{u}_j^{(i)})^2 = \max_j \|P^{(i)}\mathbf{x}\|^2 \tag{2}$$

classify $\mathbf{x}$ in class i, i.e. $\mathbf{x} \in \omega_i$.

Assume that the confidence coefficient for stopped iterating of subspaces is η $(0.5 \leqslant \eta \leqslant 1)$, if the averaging orthogonal projection of an arbitrary pattern vector $\mathbf{x}$ on the c subspaces, at the kth iterating

$$T_k = \frac{1}{c} \sum_{i=1}^{c} \mathbf{x}^{(i)T} P_k^{(i)} \mathbf{x}^{(i)} \tag{3}$$

satisfies $T_k \geqslant \eta$, c subspaces are thought to have converged to the given accuracy, the iterating will be stopped[2].

2.2 Basic Idea of Minimal Norm Based Learning Subspace Iterating Algorithm

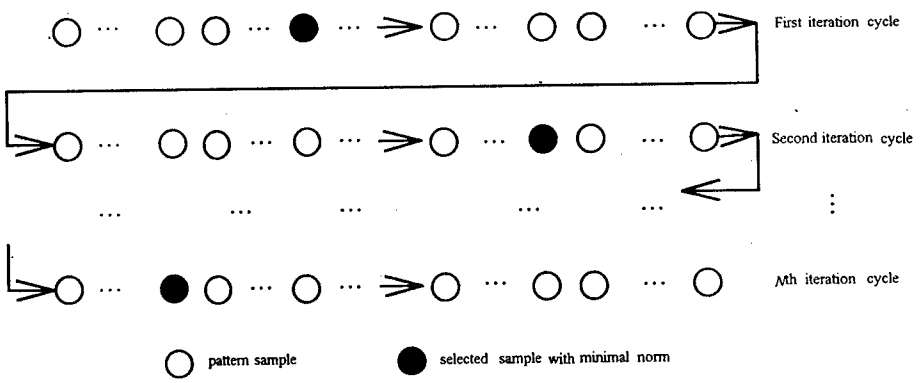


Fig 1. Scheme of the learning process for MNLSM

To avoid the effect of the order of presentation of the input samples on the classification performance, the training sample for each iteration can be selected "randonly" from training sample set. However, in doing so, it is very difficult for the learning subspace to converge to the optimal (or

gets (ship 1, ship 2 and chaff) are used as the experimental data that will be classified by the self-supervised LSM. Assume that the resolution of radar is $\triangle R = 7.5m$, the radar echoes of multiple resolution cells of targets relative to radar in the range of 3000 m to 3480 m are measured, where the azimuths of targets are changed at intervals of $\triangle\theta = 0.5°$. As a result, for each class 50 range profiles whose dimensionality is 64 are obtained. In addition, in experiment the 50 range profiles in the time-domain are transformed to those in the frequency-domain by Fourier transform. These transformed experimental data are also used to train corresponding subspaces.

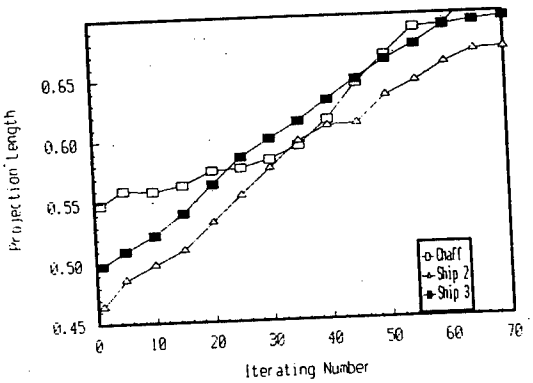


Fig 2. Dynamic learning processes of three training samples with minimal norms respectively from three classes for MNLSM classifier

Assume that the numbers of strong scatter centres of ship 1 and ship 2 are 9 and 7, respectively. In the light of the selection method of dimensionality of a subspace dicussed in [2], the dimensionalities of the subspaces corresponding to ship 1, ship 2 and chaff are selected as 10, 8 and 2, respectively. Before experiments, let all training sample vectors be normalized to unit vectors. Regardless of the properties of individual training sample vector, assume that the learning coefficient μ is selected as follows[2]

$$\mu(\mathbf{x}) = \alpha\sqrt{1 - \|\hat{\mathbf{x}}\|^2} \qquad (7)$$

where $\hat{\mathbf{x}}$ is the orthogonal projection vector of the training sample vector $\mathbf{x}$ on its own subspace; α is an adjusting coefficient. Generally, $0 < \alpha < 1$. In the practical training process $\mu_m^{(j,i)}$ is assumed to be $\mu(\mathbf{x}_m^{(i)})$, i. e., $\mu_m^{(i,i)}$.

In simulation, the obtained 50 sample vectors in the two domains for each class are divided into training sets which consist of 25 odd numbered samples and testing sets which consist of 25 even numbered samples. Assume that for each class one sample randomly selected from the 25 training samples is used to design the initial subspace. Fig. 2 shows the dynamic learning processes of three training samples with minimal norms from three classes, respectively. Table 1 gives the testing recognition results of testing samples from three classes in the two domains with the iterative number.

Table 1 The testing recognition results of testing samples from three classes in the two domains with the iterative number for MNLSM classifier

Iterative Number		10	20	30	40	50	60	70	80	90	100
Time domain	chaff	23.4%	34.2%	45.5%	56.7%	61.2%	68.7%	75.1%	81.2%	81.6%	82.1%
	ship 1	19.5%	31.5%	37.3%	45.6%	59.3%	66.4%	73.2%	76.6%	78.5%	79.2%
	ship 2	23.3%	33.7%	43.4%	49.7%	61.8%	65.2%	70.8%	73.4%	76.9%	78.4%
Frequency domain	chaff	12.5%	20.4%	26.7%	34.6%	37.5%	41.3%	48.9%	54.2%	58.5%	61.5%
	ship 1	8.6%	15.6%	21.4%	30.4%	36.6%	42.3%	51.5%	58.6%	61.3%	63.4%
	ship 2	6.7%	13.2%	21.2%	26.2%	31.5%	35.5%	40.2%	48.1%	52.6%	57.2%

In addition, assume that three kinds of self-supervised LSMs, i. e., LSM, ALSM and

MNLSM, are used to classify the simulating range profiles (time and frequency domains) of three classes. After Eq. (3) is satisfied ($\eta = 0.8$), the 25 testing samples in the two domains are used to test the rate of correct recongnition, the classification results are shown in Table 2.

Table 2 Comparisons of correct recognition rates of testing samples from three classes in the two domains using the three classification methods of LSM, ALSM and MNLSM

Classification Method		LSM	ALSM	MNLSM
Rate of Recognition (Time-Domain)	chaff	79. 8%	83. 4%	82. 8%
	ship 1	78. 9%	84. 3%	83. 7%
	ship 2	78. 3%	83. 7%	82. 4%
Rate of Recognition (Frequency-Domain)	chaff	60. 4%	64. 2%	64. 3%
	ship 1	61. 2%	63. 9%	64. 1%
	ship 2	60. 6%	64. 7%	63. 9%

From the above experiments, it can be found that the MNLSM possesses the better performance whether the convergence speed or the correct recognition rate.

4. Conclusion

Learning subspace method (LSM) for pattern recognition is one of efficient self-supervised learning neural network classifiers. This paper, based on the LSMs proposed by Kohonen, proposed a novel self-supervised LSM with higher correct classification rate and less computation time, i. e., *minimal norm based learning subspace method* (MNLSM), which is not sensitive to the order of presentation of the input samples. To verify the validities for this method, this paper discussed applying this method to recognition of simulating high resolution radar (HRR) targets. Experimental results show that the performance of proposed MNLSM such as rate of correct recognition and convergence speed is satisfactory.

Reference

[1] T. Kohonen., Self-Organization and Associative Memory. Springer-Verlag Berlin Heidelberg New York, 1989.

[2] D. S. Huang, The Study of Recognition Technique Based on One-Dimensional Image of Radar Targets. Ph. D dissertation, Xidian University, 1992.

[3] M. Kuusela, et al., "The averaging learning subspace method for spectral pattern recognition," Proc. of 6th IC on PR, Munich, pp. 134-137, 1982.

[4] E. Oja, et al., "The ALSM algorithm -an improved subspace method of classification," Pattern Recognition, PR-16, no. 4, pp. 421-427, 1983.

[5] E. Oja, Subspace Methods Of Pattern Recognition. Research Studies Press Ltd. 1983.

[6] D. L. Mensa, High Resolution Radar Imaging. Artech House, Inc., 1981.

Author Index