MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

MERL Annual Report 2003

Waters, R.C.

TR2003-00 June 30, 2003

Abstract

MERL Annual Report

July 2002 through June 2003

Welcome to Mitsubishi Electric Research Laboratories (MERL), the North American corporate R&D arm of Mitsubishi Electric Corporation (MELCO). In this report, you will find descriptions of MERL as a whole, our two laboratories (MERL Research and MERL Technology), and most importantly our current projects.

Mitsubishi Electric Research Laboratories

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Mitsubishi Electric Research Laboratories (MERL)

Annual Report

July 2002 through June 2003

TR2003-00

(Published August 2003)

Welcome to Mitsubishi Electric Research Laboratories (MERL), the North American corporate R&D arm of Mitsubishi Electric Corporation (MELCO). In this report, you will find descriptions of MERL as a whole, our two laboratories (MERL Research and MERL Technology), and most importantly our current projects.

201 Broadway, Cambridge, Massachusetts 02139
617.621.7500

Table of Contents

Mitsubishi Electric Research Laboratories

Mitsubishi Electric Research Laboratories (MERL) is the North American arm of the central research and development organization of Mitsubishi Electric Company (MELCO). MERL conducts application-motivated basic research and advanced development in computer and communications technology.

MERL's mission—our assignment from MELCO—is twofold.

- To generate highly significant intellectual property (papers, patents, and prototypes) in areas of importance to MELCO.
- To locate organizations within MELCO that can benefit from this technology and through close partnership with them, significantly impact MELCO's business.

MERL's vision—our goal for ourselves—is also twofold.

- To be one of the world's premiere research laboratories, significantly advancing the frontiers of technology and making lasting impacts on the world.
- Within our areas of expertise, to be the prime source of new technology for MELCO.

MERL focuses on two key technology sectors:

Human/computer interaction (HCI) – broadly construed to include computer vision, computer graphics, speech interfaces, data analysis and novel interaction devices.
Digital communications - featuring video processing and wireless communication.

MERL is small enough to be agile and flexible in the dynamic marketplace of ideas. However, we gain leverage from the size, recognition and diversity of our strong global parent. We turn our technical achievements into business successes by partnering with MELCO's business units and with other labs in MELCO's global R&D network.

We are strongly involved in the R&D community and standards activities, maintaining long-standing cooperative relationships with a number of research universities including MIT, CMU, Georgia Tech, Princeton, Columbia, the University of Paris, Dublin City University, ETH Zurich and the City University of London. We encourage our staff to be involved in their professional communities via conferences, papers, and continuing professional development.

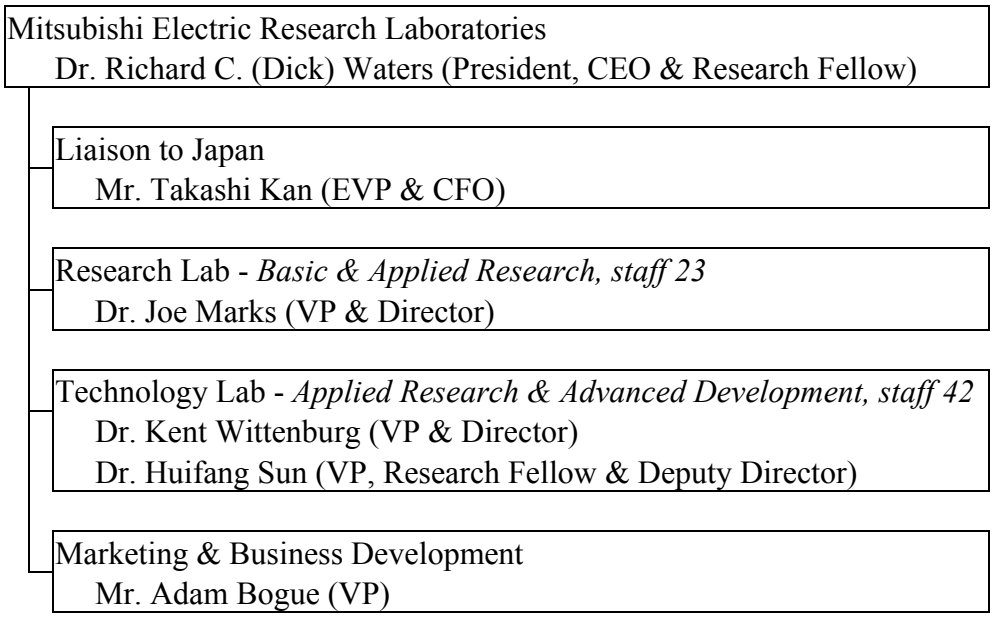
MERL consists of two laboratories, which share the same space in Cambridge Massachusetts and collaborate closely to achieve groundbreaking results. Our output ranges from papers and patents, through proof-of-concept hardware and software prototypes, to industry-first products. The Headquarters operation includes a small marketing and business development department to help assess the potential market impact of our work and an in-house patent department to speed the filing of patents.

This Annual report is a snapshot of MERL's web site. For additional and updated information please visit "http://www.merl.com".

Dick Waters

President, MERL

The following shows the basic organization of MERL. The six members of the top management team work closely together, guiding all aspects of MERL's operation.



Richard C. (Dick) Waters
Ph.D., Massachusetts Institute of Technology, 1978
President, Chief Executive Officer & Research Fellow

Dick Waters received his Ph.D. in artificial intelligence (AI). For the next 13 years he worked at the MIT AI Lab as a Research Scientist and co-principal investigator of the Programmer's Apprentice project. Dick was a founding member of MERL's Research Lab in 1991. As a MERL researcher, his work centered on multi-user interactive environments for work, learning and play. For this work, he was made a MERL Research Fellow in 1996. In January 1998, Dick became Director of MERL's Research Lab. In December 1999, he became CEO of MERL as a whole. In addition to his duties at MERL, Dick is currently a member of the board of directors of the Computer Research Association.

Takashi Kan
M.S., Tohoku University, 1978
Executive Vice President Chief Financial Officer & Chief Liaison Officer

Takashi Kan joined MELCO in 1978. In the 80s and 90s, Takashi worked on a variety of computer-related projects involving computer architecture, parallel computing, digital broadcasting and image processing. From 1987 to 1988, he was a visiting associate professor at the University of Illinois at Urbana-Champaign. Before coming to MERL in 2002, Takashi rose to the position of general manager of the Multimedia Laboratory within MELCO's Information Sciences R&D Center (Johosoken).

Joe Marks
Ph.D., Harvard University, 1991
Vice President; Director Research Lab

Prior to Joining MERL in 1994, Joe Marks worked at two other research labs: Bolt Beranek & Newman and Digital Equipment Corporation's Cambridge Research Laboratory. In addition, Joe was an adjunct lecturer at Harvard University. As a researcher at MERL, Joe's primary focus was on computer graphics, user interfaces, and heuristic optimization. In these areas, he plays a strong role in the scientific community including being: Chair of ACM SIGART, Associate Editor for ACM Transactions on Graphics, and Papers Chair for SIGGRAPH 2004. Joe became Associate Director of the MERL's Research Lab in 1999 and Director in 2000.

Kent Wittenburg
Ph.D., University of Texas at Austin, 1986
Vice President; Director Technology Lab

Before Joining MERL in 2001, Kent Wittenburg worked at the Microelectronics and Computer Technology Corporation (MCC), Bellcore, and Verizon/GTE laboratories. His research encompassed a variety of Human-Computer Interaction (HCI) technologies including rapid serial visual presentation, multidimensional information visualization, and natural language. He managed groups in natural language interfaces and Internet technologies prior to joining MERL as group manager of speech and HCI. Kent was promoted to Laboratory Director in 2002.

Huifang Sun
Ph.D., University of Ottawa, 1986
Vice President & Research Fellow; Deputy Director Technology Lab

After receiving his Ph.D., Huifang Sun became an Associate Professor at Fairleigh Dickinson University. In 1990 Huifang moved to the Sarnoff Research Laboratory where he became Technology Leader of Digital Video Communication and did extensive work on MPEG standards. In 1995, Huifang joined MERL as the leader of our video efforts, becoming a Deputy Lab Director in 1997. In recognition of his long and productive career in video processing Huifang was made an IEEE Fellow in 2001. He was made a MERL Research Fellow in 2003.

Adam Bogue
B.S., MIT, 1986; MBA, MIT Sloan School, 1990
Vice President

Adam Bogue had 15 years of industry experience before joining MERL. This included 3 years at GenRad Inc, where Adam was responsible for managing a new line of automatic test equipment. Subsequently, Adam spent 7 years at Active Control eXperts Inc. beginning as Director of Sales and Marketing and ending as Vice President, Core and New Business Unit helping to grow ACX into a successful Inc. 500 company. Adam came to MERL in June of 2000 to lead our Marketing and Business Development effort.

Mitsubishi Electric

Number 126 on Fortune magazine's 2003 list of the world's 500 largest corporations, Mitsubishi Electric Corporation (MELCO) has approximately $30 billion in annual sales and more than 100,000 employees in 35 countries. Like most Japanese companies, the lingering malaise of the Japanese economy coupled with the worldwide slump in semiconductors and cell phones has lead to losses. However, these losses have been relatively small and energetic restructuring (including the merging of most of MELCO's semiconductor operations with Hitachi) has given MELCO a firm foundation for growth once the world economy recovers.
MELCO is composed of a wide range of operations. The business units with sales of $1 billion or more are listed below in order of estimated 2003 revenue. (The rightmost column shows the abbreviated Japanese business unit nicknames commonly used by MELCO insiders.)

Mitsubishi Electric Diversified Electrical and Electronics Manufacturer		MELCO
Living Environment & Digital Media Equipment Air Conditioners, Refrigerators, TVs, DVDs, LCD Projectors	(Shizuoka, Kyoto)	Lihon
Social Infrastructure Systems Power Equipment, Plant Control, Transportation	(Kobe, Itami)	Shakaihon
Communication Systems Wired Communications, Broadcast Communications, Cell Phones	(Kamakura, Itami)	Tsuhon
Automotive Equipment Alternators, Engine Controllers, Car Stereos, Car Navigation	(Himeji, Sanda)	Shahon
Building Systems Elevators, Escalators, Building Monitoring	(Inazawa)	Biruhon
Factory Automation Programmable Logic Controllers, Industrial Machine Tools	(Nagoya)	FAhon
Electronic Systems Satellites, Radar, Military Systems	(Kamakura, Itami)	Denshihon
Information Systems and Services Turnkey Information Systems, Computer Hardware	(Tokyo, Kamakura)	IShon
Semiconductors Optical and Radio Frequency Semiconductors	(Kita Itami)	Hanpon

Together, these nine business units produce approximately three quarters of MELCO's revenue. Because information technology is important to each of the business units, MERL works with them all.
Nearly 10% of MELCO's sales are in North America and many of MELCO's business units have North American subsidiaries. MERL seeks to work directly with these subsidiaries, particularly when they have substantial local design and manufacturing as well as sales.
The US operations with sales of $100 million or more are listed below in order of estimated 2003 revenue. The largest of these (MDEA) is part of Lihon and has sales of approximately $1B.

Mitsubishi Digital Electronics America, Inc.	(MDEA)
Design, Manufacturing & Sales: Lihon (Los Angeles, Mexicali MX)	
High Definition Projection Televisions, DVDs, VCRs	

Mitsubishi Electric Automotive America, Inc.	(MEAA)
Manufacturing & Sales: Shahon (Detroit, Mason OH)	
Auto Parts	

Mitsubishi Electric United States, Inc.	(MEUS)
Sales: Several BUs (Los Angeles, Sunnyvale & other cities)	
Semiconductors, Air Conditioning, Elevators	

Mitsubishi Electric Power Products, Inc.	(MEPPI)
Design, Manufacturing & Sales: Shakaihon (Pittsburgh)	
Power Transmission Products	

Mitsubishi Electric Automation, Inc.	(MEAU)
Sales & Installation: FAhon (Chicago)	
Factory Automation Equipment	

It is worthy of note that there are over 30 major independent companies in the world that use the word "Mitsubishi" in their names. These companies include the Mitsubishi Trading Company, Mitsubishi Motors, Mitsubishi-Tokyo Financial Group, Mitsubishi Heavy Industries and Mitsubishi Chemical (all five of which are also on the Fortune Global 500 list—Numbers 10, 118, 187, 206 & 317 respectively). They have shared roots in 19th century Japan; however, these companies have been separate for many years and MELCO has been separate from all of them since MELCO's founding in 1921.

Mitsubishi Electric Corporate R&D

MERL is one of five laboratories in MELCO's global corporate R&D network. The chart below summarizes the primary activities of these labs. MERL pursues collaborations with all these labs. (The rightmost column shows the Japanese nicknames commonly used by insiders.)

Corporate R&D Headquarters: Dr. H.Ogata (Director), Mr. K.Kuroda (GM), 20 people (Tokyo) Managing MELCO's R&D	Hatsuhon
Advanced Technology R&D Center (ATC) Research & Advanced Development: Dr. K.Kyuma (GM), 900 people (Itami) Materials, Semiconductor Devices, Electrical & Mechanical Engineering	Sentansoken
Information Technology R&D Center (ITC) Advanced Development: Dr. H.Koezuka (GM), 900 People (Ofuna) Information Systems, Communications, Opto-Electronics	Johosoken
Industrial Design Center (IDC) Advanced Development: Mr. K.Chiba (GM), 100 people (Ofuna) Industrial Design, Usability Studies	IDken
Mitsubishi Electric Research Laboratories (MERL) Research & Advanced Development: Dr. R.Waters (CEO), 80 people (MA) Computer Vision, Speech Interfaces, HCI, Digital Audio & Video Communications	MERL
Mitsubishi Electric Information Technology Centre Europe (ITE) Advanced Development: Mr. R.Nishii (CEO), 50 people (France & England) Wireless Communications, Digital Audio & Video	ITE

Historical Note

The history of MERL is a story of separate labs with separate histories coming together into an organization that has become much more than the sum of its parts.

In 1984, MELCO's computer business unit founded a laboratory in Waltham MA called Horizon Research inc. (HRI). HRI was created to design IBM compatible computers, which MELCO then produced and sold. However, in the late 80's MELCO decided to withdraw from this business. In the ensuing years, Horizon transformed itself into a software-oriented laboratory working with multiple parts of MELCO. In consonance with this, HRI left the computer business unit and became part of MELCO's central R&D unit (CR&D) in the early 90s.

In 1991, MELCO's central R&D organization founded a research laboratory in Cambridge MA called Mitsubishi Electric Research Laboratories (MERL). The charter of this lab was to do fundamental research in computer science, with the goal of leading MELCO into the future with innovative new technologies.

In 1993, MELCO's audio/visual business unit founded a laboratory in Princeton NJ called the Advanced Television Laboratory (ATL). ATL was created to develop a chip set capable of decoding US HDTV signals. ATL worked on this in close collaboration with Bell Labs and MELCO's semiconductor business unit. After a couple of years, ATL moved to Murray Hill NJ to be closer to Bell Labs. In the mid-90s, as the chip set was being completed, ATL began to branch out into other digital video technologies and digital communications in general. In addition, ATL created a satellite operation in Sunnyvale CA focusing on digital broadcasting and home networking. As these changes were going on, ATL was transferred from its original parent business unit to CR&D.

In 1995, CR&D created a new US entity called Mitsubishi Electric Information Technology Center America (MEITCA or more simply ITA) in order to gather CR&D's various labs in the US together into a single entity. HRI was incorporated into ITA becoming the Horizon Systems Lab (ITA/HSL). ATL was incorporated into ITA becoming ITA/ATL. The following year, MERL was incorporated into ITA as well, becoming ITA/MERL.

The ITA that resulted in 1996 had a great deal of promise. By combining advanced development and basic research into a single organization, ITA made it easier for technological advances to be perfected and transferred into MELCO business. Further, by combining expertise from a wide range of areas and experience with a wide range of MELCO business units into a single organization, ITA was in a good position to understand a wide range of technical trends, market needs, and MELCO business opportunities.

However, like any organization created by the combination of pre-existing parts, ITA faced many obstacles blocking the realization of its full potential. In particular, the natural momentum was for ITA/HSL, ITA/ATL, and ITA/MERL to continue operating separately, following their own goals and agendas, with collaboration being the exception rather than the rule. This fundamental problem was significantly aggravated by the geographic separation of the labs.

The first inter-lab collaborations within ITA were between HSL and MERL. These led to the creation of a new engineering group in ITA, the Volume Graphics Organization (VGO), with the charter of designing a real time volume-graphics rendering chip that could be hosted on a PC plug in board. VGO was based on research carried out at ITA/MERL and began as a project within ITA/HSL before growing so large that it was made a separate laboratory within ITA in 1998. From its inception through 1998, VGO shared space with ITA/MERL and interacted

closely with ITA/MERL researchers. However, in 1999 VGO moved to separate quarters in Concord MA as sales of its first product were beginning.

The technical success of VGO notwithstanding, ITA in 1999 was an extremely fragmented organization, with some 115 people divided between 5 locations: VGO in Concord MA, HSL in Waltham MA, MERL in Cambridge MA, ATL in Murray Hill NJ and an ATL branch in Sunnyvale CA—making intra-ITA collaboration difficult. The overwhelming impression was of an organization with many small parts that was little (if any) more effective than the sum of those parts. However, in late 1999, the tide began to turn toward the unified and invigorated US research operation that exists today.

The first wave of change involved three steps. (1) The ATL branch facility in Sunnyvale was closed and the people moved to NJ. (2) The HSL building in Waltham was closed and the people moved to the Cambridge building. (3) Finally, to provide a symbol of greater unification, ITA was renamed Mitsubishi Electric Research Labs (MERL)—Mitsubishi Electric Information Technology Center America never having been a very satisfactory name—and the individual labs were given geographic names replacing the names they had independently operated under: ATL became the MERL Murray Hill Lab (MERL/MHL), HSL became the MERL Cambridge Systems Lab (MERL/CSL), the research lab became the MERL Cambridge Research Lab (MERL/CRL) and VGO became MERL Concord. In 2001, the volume graphics group left MERL becoming part of the business unit selling its product. This left a more streamlined MERL, with people in just two locations: Cambridge and Murray Hill.

Having CSL and CRL in the same building led to greatly increased collaboration between the two labs and was a great success. In 2003, it was decided to take the final logical step, closing the building in NJ and moving these people to the Cambridge building as well. Further, in the interest of greater collaboration and greater management efficiency, the two advanced development labs, MHL and CSL, were merged into a single new lab called the MERL Technology lab (MTL). In analogy with this name, CRL was renamed the MERL Research Lab (MRL).

Like all real histories, the story above is convoluted with many twists and turns. However, through it runs a steady thread of MELCO's strong commitment to research and advanced development in the US. Now that everybody is in one place, where they can work together to the greatest advantage, the stage is set for MELCO to reap maximum advantage from this commitment.

MERL Research Lab

The MERL Research Lab (MRL) pursues basic research in applied computing. Our efforts are directed towards applications of practical significance, but our time horizon is long (five or more years) and our appetite for technical challenge and risk is high. Located in Cambridge, Massachusetts, hometown of Harvard University and the Massachusetts Institute of Technology, the laboratory currently has a technical staff of 23. The permanent staff is enriched by visiting scientists, consultants, academic collaborators, and student interns. Our colleagues within MELCO provide other opportunities for collaboration and enrichment. In particular, our close cooperation with the MERL Technology Lab (MTL) has recently produced several successful, synergistic projects. We also participate actively in external research communities, exposing our work to critical peer review, and publishing our results as quickly as possible.

Our primary areas of research are computer vision, audio and image processing, computer graphics, natural language processing, speech processing, data analysis, decision and control, networked and wireless communication, and human-computer interaction. These areas were chosen both because of the opportunities that they currently offer for technological innovation, but also because of their high relevance to many MELCO businesses. Within these areas, we strive to advance the state of the art and to achieve business impact through the invention of new products and services. On both counts, we have been very successful of late. A selection of project highlights from this past year include:

- *Adaptive-distance fields for digital typography*: Every type of electronic display is a potential business opportunity for this new method of representing and displaying scalable text. Although this project emerged only recently, it is built. on fundamental research going back many years.
- *Data analysis and data mining*: New fundamental data-analysis and data-exploration algorithms generated several major publications and helped launch a new business for MELCO. This work was recognized with a MELCO CR&D Award in FY2002.
- *Projected displays*: The potential synergy of projectors and cameras was explored in several projects at MRL. DiamondTouch, a multi-user touch surface for projected displays, was the focal point for a related body of work conducted in collaboration with MTL. Collectively, these projects produced some notable publications (including a paper at SIGGRAPH 2003, the premier computer-graphics conference), sparked multiple dialogs with interested parties in MELCO business units, and was recognized with a MELCO Johosoken Award in FY2002, a MELCO IP Award in FY2002, and selection as a finalist in the IT Hardware section of the 2003 World Technology Network Awards.
- *Security and surveillance technologies*: In collaboration with colleagues at MTL, researchers at MRL worked on fingerprint recognition, face detection, face recognition, 3D face recognition, far-field human detection, people counting, human tracking, and non-photorealistic rendering of surveillance video. These projects again produced several significant publications (including an orally presented paper at ICCV 2003, the premier computer-vision conference) and generated interest at several MELCO business units.
- *Group elevator control*: The first major new research direction in group-elevator control in more than 20 years resulted in a Best Paper award at the 2003 International Conference on Automated Planning and Scheduling (ICAPS) conference.
- *LED-based communication and chemical sensing*: Joint research with MTL resulted in new uses for the humble LED. Although at an early stage, this work has generated

several promising leads at MELCO and resulted in a paper at Ubicomp 2003, the premier conference on ubiquitous computing.

Successful basic research in an industrial setting requires many things: a first-rate team of researchers recruited from a global pool; a long-term commitment to open-ended, risky exploration; research themes that resonate well with the parent company and that are ripe for commercial exploitation; and effective tech-transfer mechanisms whereby research ideas can lead to business impact. These elements are all in place at MERL Research and we look to the future with much optimism.

Joe Marks
Director, MERL Research Lab

MERL Technology Lab

The programs at MERL's Technology Laboratory (MTL) are deliberately intertwined with those at MRL. However, where MRL is composed largely of individual researchers with a time horizon of five or more years, MTL is devoted to the practical realization of technology innovations in a one-to-four-year time frame. This goal requires significant structure and coordination. To facilitate this, MTL is organized into four groups, each with about 10 technical staff. These teams produce early concept prototypes and patents along with MRL researchers. However, MTL goes beyond this to advanced development, resolving key technical issues that stand in the way of transforming a gleam in a researcher's eye into a full-bodied form with business impact.

The work at MTL is grounded in two key ways. First, we focus on real-world business problems posed by MELCO business units to drive technological research. This is done in close collaboration with MELCO business units in the U.S. and Japan, and with the other CR&D labs in Japan and Europe. We strive to build particularly close relationships with our Japanese colleagues since they are often the key to achieving impact for our work on the business bottom line. Second, we participate strongly in international standards bodies, making key contributions in areas of importance to MELCO.

A particular strength of MTL is its emphasis on collaboration across conventional academic and engineering disciplines. Expertise in the lab ranges all the way from chip-level and board-level hardware through low-level communication protocol layers to all levels of software and applications. Our disciplines range from communications and video processing to software systems featuring data analysis, computer vision, speech recognition, and human interfaces. The lab's strength is most evident in projects that require contributions from multiple disciplines such as projects in automotive multimedia, video surveillance & analysis, sensor networks, and interactive retail display environments.

The following paragraphs summarize the activities in the four MTL groups: Computer Vision Applications & Devices, Digital Communications & Networking, Digital Video Technology & Systems and Human-centric & Analytic Systems.

The Computer Vision Applications & Devices group is devoted to building software libraries and systems for a variety of applications that make use of the world-class computer vision research emerging from MERL as a whole. These applications include security, safety of elevator operations, analysis and enhancement of imagery, calibration of multiple projector sources, and demographic analysis from video data. This group leads a multi-year project funded from MELCO Corporate R&D that is focused on building software libraries that are aligned with a variety of business units in the area of computer-human observation. The group also focuses on innovative devices and their interaction off the desktop. One focus for this year was development of new forms of interaction in retail environments that incorporate multiple projectors and sensors.

The Digital Communications & Networking group focuses on high-speed digital communications systems and advanced networking. The work spans next generation mobile wireless communications systems, wireless local area and home networks, and high speed Ethernet. A corporate project on ultra wideband (UWB) technology includes participation by all of MELCO's corporate laboratories and is being led by this group. The team creates new concepts and technologies, new algorithms, and new designs that enhance MELCO's IP portfolio. The group is particularly active in industry standards such as 3GPP, UWB, ZigBee

and the IEEE standards 802.11, 802.15 & 802.16. The team collaborates with MELCO's business units to develop competitive communications and networking products such as ones relating to the Simple Control Protocol (SCP) for Power Line Communication (PLC).

The Digital Video Technology & Systems group develops technologies for multimedia coding and content analysis. It develops video compression and transcoding algorithms and software as well as methods for audio/video content analysis and indexing. Key contributions include MPEG 2 to 4 transcoding and compression for high-definition TV as well as video summarization, event detection, and content mining. The team has been active in the development of MPEG and other standards, including MPEG-2, MPEG-4, MPEG-7, MPEG-21, JVT, and IETF. The group works closely with Japan to bring these technologies to the marketplace for video products and services in the consumer, business, and government arenas.

The Human-Centric & Analytic Systems group is focused on developing advanced software systems to enable new user interfaces to digital information. This area encompasses speech interfaces multimodal interfaces, collaboration systems, data-mining, and sound classification. In the speech arena, the group supports the business units by testing and evaluating speech engines as well as developing novel paradigms for speech and multimodal user interfaces. In the area of collaboration systems it focuses on software for new off-the-desktop devices such as MERL's DiamondTouch multi-user touch technology. Other efforts include data-mining in support of recommendation engines and sound classification for industrial and consumer applications. This group works closely with MELCO's industrial design center and also with the automotive, information systems, electronics, and public utility systems business units of MELCO.

Kent Wittenburg
Director, MERL Technology Lab

Recent Awards and Commendations

The high caliber of MERL's staff and research is evident in a variety of ways. Four are shown below. The first is the members of our staff that are Fellows of technical societies. The second and third are the best paper awards and technology awards received from outside organizations in the past two years. The fourth is awards received from MELCO in the past two years for MERL's contribution to MELCO products.

Current Technical Society Fellows

Dr. Charles Rich, Fellow American Association for Artificial Intelligence.

Dr. Candace L. Sidner, Fellow American Association for Artificial Intelligence.

Dr. Huifang Sun, Fellow Institute of Electrical and Electronic Engineers.

Best Paper Awards

Nikovski, D.; Brand M.E., "Decision-Theoretic Group Elevator Scheduling", *International Conference on Automated Planning and Scheduling (ICAPS)*, June 2003. [Best applied research paper at the conference.]

Yin, P.; Vetro, A.; Liu, B.; Sun, H., "Drift Compensation for Reduced Spatial Resolution Transcoding", *IEEE Transactions on Circuits and Systems for Video Technology*, ISSN: 1051-8215, Vol. 12, Issue 11, pp. 1009-1020, November 2002. [Best full-length paper of the year in the journal.]

Almers, P.; Tufvesson, F.; Edfors, O.; Molisch, A.F., "Measured Capacity Gain Using Water-filling in Frequency Selective MIMO Channels", IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), September 2002, Vol. 3, pp. 1347-1351. [Best paper at the conference.]

Vetro, A.; Hata, T.; Kuwahara, N.; Kalva, H.; Sekiguchi, S.;, "Complexity-Quality Analysis of Transcoding Architectures for Reduced Spatial Resolution", *IEEE Transactions on Consumer Electronics*, ISSN: 0098-3063, Vol. 48, Issue 3, pp. 515-521, August 2002. [Second best full-length paper of the year in the journal.]

Vetro, A.; Hata, T.; Kuwahara, N.; Kalva, H., "Complexity-Quality Analysis of MPEG-2 to MPEG-4 Transcoding Architectures", *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 130-131, June 2002. [Second best paper at the conference.]

Brand, M.E., "Morphable 3D Models from Video", *IEEE Computer Vision and Pattern Recognition (CVPR)*, Vol. 2, pp. 456, December 2001. [Best paper at the conference.]

Moghaddam, B.; Nastar, C.; Pentland, A., "Bayesian Face Recognition with Deformable Image Models", *IEEE International Conference on Image Analysis and Processing (ICIAP)*, pp. 26-35, September 2001. [Winner of the 2001 *Pierre Devijver* Prize from the International Association of Pattern Recognition (IAPR).]

Lesh, N.B.; Rich, C.; Sidner, C.L., "Collaborating with Focused and Unfocused Users Under Imperfect Communication," *International Conference on User Modeling (UM),* July 2001, pp. 64-73. [Outstanding Paper Award.]

Technology Awards From Outside Organizations

In 2003, DiamondTouch was selected by the World Technology Network as one of 5 finalists for selection as the new technology with the greatest likely long-term significance in the category "Information Technology – Hardware".

Awards From MELCO

In 2003, MERL staff received a "CR&D" award for the addition of data mining and collaborative filtering capabilities to MELCO's Diaprism database product through the invention of Imputatitve Incremental Singular Value Decomposition (IISVD).
In 2003, MERL staff received a "Johosoken" award for their contribution to the technology underlying multi-projector display walls through the development of algorithms supporting seamless projection.
In 2003, MERL staff received an "Intellectual Property" award for the patent underlying MERL's DiamondTouch multi-user touch surface, computer interface device.
In 2001, MERL staff received a "CR&D" award for multiple contributions to the MPEG-7 standard in the areas of audio and video.

Technical Staff

Ali Azarbayejani
Ph.D., Massachusetts Institute of Technology, 1997
Principal Technical Staff
MERL Technology Lab

Ali Azarbayejani's doctoral research was on vision-based computational 3D geometry and underlying nonlinear probabilistic methods. In 1997, he founded Alchemy 3D Technology to develop technology and software based on his computer vision research and led the development of new markets in the film and video post-production industry for vision-based software. In 2002, he began consulting in software design and development and in 2003, he joined MERL with interests in technology, software, and business development.

Paul Beardsley
Ph.D., Oxford University, 1992
Research Scientist
MERL Research Lab

Paul Beardsley is a computer vision researcher. He obtained a DPhil from Oxford for work in applications of projective geometry to computer vision. He is currently working on a geometry-based vision library in C++. The animation shown above left consists of synthetically-generated images which were created in a former project on tracking head and eye motions. Current projects are on stereo-vision for surveillance, and on a hand-held 3D scanner.

Ghulam Bhatti
Ph.D., Boston University, 1998
Member Technical Staff
MERL Technology Lab

Ghulam Bhatti received his Ph.D. in Computer Science, specializing in Distributed and Parallel Discrete Event Simulation. He joined MERL in November 2000. Previously, Ghulam has worked as a Sr. Software Engineer at Evare LLC, Inc, developing software for a network switch, named as Winternet, and implementing an RSA cryptographic scheme. He also worked at Excel Tech. Ltd. (XLTEK) as an Embedded Software Engineer, developing embedded software for a portable EEG device. Currently, Ghulam is working on Home Networking and Digital TV. His interests include algorithms, embedded software development, and networking.

Emmanuelle Bourrat
M.S., Université de Rennes1, 1999
Member Technical Staff
MERL Technology Lab

Emmanuelle Bourrat graduated in 1999 from IRISA (Institut de Recherche en Informatique et Systèmes Aléatoires), Université de Rennes1, France. She earned a Master's degree in computer science and digital images. She was an intern at MERL from February 2000 to August 2000, where she worked on a tool to visualize and segment scans of the head. She joined MERL as a

full-time employee in August 2000. Her main interest is in computer vision and medical imaging. She is currently working on the CHO (Computer Human Observation) project.

Matthew Brand
Ph.D., Northwestern University, 1994
Research Scientist
MERL Research Lab

Matthew Brand studies unsupervised learning from sensory data. One goal is to make machines that learn to realistically mimic and augment human performances; another is to optimize systems that serve the conflicting interests of large numbers of people. Recent results include spectral solutions for reconstructing manifolds from samples, decision-theoretic elevator group control, a linear-time online SVD, video-realistic synthesis of humans and nature scenes, recovery of nonrigid 3D shape from ordinary video, and an entropy optimization framework for learning. Brand has been named one of the top innovators of his generation (Technology Review 1999) and one of industry's top "R&D stars" (Industry Week 2000). Recent academic honors include best paper awards in computer vision (CVPR2001) and scheduling (ICAPS2003).

Dirk Brinkman
J.D., Suffolk University Law School
Patent Counsel
MERL Headquarters

Dirk Brinkman's undergraduate and Masters work was in Medical Physics. Prior to joining MERL in 1998, he spent most of his career at Digital Equipment Corporation, first as an engineer and product manager in the Medical Systems Group and then as a Patent Attorney for Digital's Research Laboratories in Cambridge MA and Palo Alto CA.

Johnas Cukier
M.Sc., from Polytechnic Institute of New York, 1985
Principal Member Technical Staff
MERL Technology Lab

Johnas Cukier received his B.Sc. degree in Physics and Computer Science in 1983 from New York University and his M.Sc. degree in Electrical Engineering in 1985 from Polytechnic Institute of New York. He joined MERL in 1996, working on digital systems for CATV, RF microwave transmitters/receivers, and front-end advanced TV receivers. His current interests are in advanced Digital Networking and Digital Signal Processing.

Andrew J. Curtin
J.D., Suffolk University Law School, 1997
Associate Patent Counsel
MERL Headquarters

Andrew Curtin received his B.S. in Marine Engineering from the Massachusetts Maritime Academy and his J.D. from Suffolk University Law School. Prior to joining MERL in 2001, he spent six years as an engineering officer aboard U.S. flag merchant ships engaged in world-wide trade, followed by two years as an attorney in private practice.

Paul Dietz
Ph.D., Carnegie Mellon University, 1995
Principal Technical Staff
MERL Technology Lab

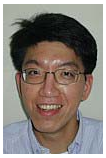
Paul Dietz is an over-educated, academic refugee who seeks happiness through creating clever devices and systems. Most recently, Paul headed up the electrical engineering efforts at Walt Disney Imagineering's Cambridge R&D lab where he worked on a wide variety of projects including theme park attractions, systems for the ABC television network and consumer products. Since joining MERL in 2000, Paul has been leading efforts developing new user interface technologies.

Ajay Divakaran
Ph.D., Rensselaer Polytechnic Institute, 1993
Senior Principal Member Technical Staff
MERL Technology Lab

Ajay Divakaran was an Assistant Professor with the Department of Electronics and Communications Engineering, University of Jodhpur, 1985-86 and a Research Associate at the Department of Electrical Communication Engineering, Indian Institute of Science, in Bangalore, India in 1994-95. He was a Scientist with Iterated Systems Inc., Atlanta, GA from 1995 to 1998 before joining MERL in 1998. He has been an active contributor to the MPEG-7 video standard. His current research interests include video analysis, summarization, indexing and compression, and related applications. He currently serves on program committees of key conferences in the area of multimedia content analysis.

Alan Esenther
M.Sc., Boston University, 1993
Principal Technical Staff
MERL Technology Lab

Alan Esenther enjoys HCI (human-computer interaction), distributed software development, graphical user interfaces, Internet technologies, and creating new applications before dawn. Recent work includes touch applications that support multiple concurrent users (think multiple mice), rapid image presentation for video browsing, and instant co-browsing (lightweight real-time distributed collaboration using unmodified web browsers). Previous work includes interactive web page generation for mobile agents, an email gateway for offline surfing, transaction monitors, kernel-level volume management, and microprocessor development. Interests include rock climbing, adventure travel, and learning new technologies.



James Fang
B.Sc., Columbia University, 1992
Member Technical Staff
MERL Technology Lab

James Fang received his B.Sc. from Columbia University in 1992 and did some graduate work there before joining Mitsubishi Electric in 1995. He worked on consumer televisions for three years before transferring to MERL in 1998. He is currently working on digital wireless communications.

George Fang
B.Sc., California Institute of Technology, 1990
Member Technical Staff
MERL Technology Lab

George Fang received his B.Sc. degree from California Institute of Technology and became a member of MELCO's Kyoto Works in 1990. During the ten years working in Japan, he was a hardware engineer designing analog and digital consumer televisions for the American market and coordinated joint design efforts between Japan and the United States. He joined MERL in February of 2001 with research objectives in wireless and network technologies.

Clifton Forlines
Master of HCI, Carnegie Mellon University, 2001
Research Associate
MERL Research Lab

Clifton Forlines is Research Associate at MERL Research Lab. His research interests include the design and evaluation of novel user interfaces. Current research projects span from three-dimensional presentation of and navigation through recorded digital video, to collaborative tabletop user interfaces, to using hand-held projectors for augmented reality. He is currently leading the evaluation of three projects, MediaFinder, TimeTunnel, and DiamondSpin. Before coming to MERL, Clifton worked on Carnegie Mellon's Alice project, which aimed at teaching programming to children through building interactive 3D worlds.

Sarah Frisken
Ph.D., Carnegie Mellon University, 1991
Senior Research Scientist
MERL Research Lab

Sarah Frisken (formerly Gibson) has research interests in computer graphics, volume visualization and physically based modeling. She has led a team of researchers and students to build a knee arthroscopy simulator that incorporates high-quality rendering, haptic feedback and physical modeling to simulate interactions between surgical tools and a computer model derived from 3D MRI data. Her current work is with Adaptively Sampled Distance Fields, a general representation of shape for computer graphics, which provides intuitive manipulation, and editing of smooth surfaces with fine detail. Applications include digital sculpting, volumetric effects for rendering, color gamut representation, CNC milling, and rapid prototyping.

Daqing Gu
Ph.D., SUNY Stony Brook, 1999
Principal Technical Staff
MERL Technology Lab

Daqing Gu received the BE degree from Tsinghua University, Beijing, China in 1987; the MS and Ph.D. degrees in electrical engineering from the State University of New York at Stony Brook, Stony Brook, NY in 1996 and 1999, respectively. He joined MERL in 1999, and is currently a principal member of technical staff. He has been involved in many wireless communications and networking projects, and has many publications in the field. His current research interests include IEEE802.11 standardizations, QoS in wireless communications and networks, multimedia home networking and MIMO-OFDM technologies.

Jianlin Guo
Ph.D., Windsor University, 1995
Principal Technical Staff
MERL Technology Lab

Jainlin Guo received his Ph.D. from Windsor University in 1995. He worked at Waterloo Maple for a year and a half as a software developer and joined MERL in 1998. He has published seven research papers and his primary research interests include home networks, digital broadcasting, and wireless computing.

Bret Harsham
Massachusetts Institute of Technology
Principal Technical Staff
MERL Technology Lab

Bret Harsham joined MERL in January 2001 to pursue interests in speech user interfaces and speech-centric devices. Prior to joining MERL, Bret spent 3 1/2 years at Dragon Systems designing and implementing handheld and automotive speech products. Earlier, he was a principal architect of a Firewall and Virtual Private Network product. Bret's other technical interests include distributed architectures, knowledge representation and language theory.

Jyhchau (Henry) Horng
Ph.D., Polytechnic University, 1998
Principal Technical Staff
MERL Technology Lab

Henry Horng received the Ph.D. from Polytechnic University in 1998. He has worked as a research assistant at Polytechnic and as software developer and lecturer for CCIT, Taiwan. Henry joined MERL in 1999. His primary research interests include digital communication and signal processing.

Frederick J. Igo, Jr.
B.A., LeMoyne College
Senior Principal Technical Staff
MERL Technology Lab

Fred Igo's professional interests are in software development and its process. Fred joined MERL in 1985 and has worked on various software technology, including Distributed Computing, Distributed OLTP, Message Queuing, Mobile Agents, OLAP/MDDB and Data Mining. Prior to joining MERL Fred worked at IPL systems.

Michael Jones
Ph.D., Massachusetts Institute of Technology, 1997
Principal Technical Staff
MERL Technology Lab

Mike Jones joined MERL in the fall of 2001 after 4 years at the Digital/Compaq Cambridge Research Laboratory. Mike's main area of interest is computer vision, and he is particularly interested in using machine learning approaches for solving computer vision problems. He has focused on algorithms for detecting and analyzing people in images and video such as face detection, skin detection and facial analysis using morphable models. Recent Projects include Fast Face Detection using a Cascade of Detectors.

Ignacio Kadel-Garcia
Massachusetts Institute of Technology, 1981-1988
Systems and Network Engineer
MERL Headquarters

Nico (short for Ignacio) is a member of the Computer Network Services Group. He supports UNIX and Linux software and hardware, and various networking services for MERL. His previous experience includes Systems Engineering for Akamai Technologies. Nico also provided computer support and designed analog instruments for cochlear implant research for a dozen years at the Cochlear Implant Research Lab of the Massachusetts Eye and Ear Infirmary.

Mamoru Kato
B.S., Tokyo University, 1991
Principal Technical Staff
MERL Technology Lab

Mamoru Kato received his B.S. degree in Electrical Engineering from Tokyo University and became a member of Mitsubishi Electric Corporation in 1991. He worked on IA servers and network security products as a hardware engineer before he joined MERL in December of 2002. His current interest is in sensor network and data mining.

Hao-Song Kong
Ph.D., Sydney University, 1998
Member Technical Staff
MERL Technology Lab

Hao-Song Kong has research interests in neural networks, digital signal processing, image processing, video transcoding, video transmission and networking. He worked at Motorola Australian Research Center over 4 years before joining MERL in 2001. His current projects are DVD recording, real-time video streaming platform development and QoS provisioning research for the next generation wired and wireless IP networks.

Christopher Lee
Ph.D., Carnegie Mellon University, 2000
Visiting Scientist
MERL Technology Lab

Christopher Lee is a graduate of the Robotics Ph.D. program at Carnegie Mellon University. His research is motivated by the potential of robots to work with and to learn from people. His work utilizes technology from robotics, artificial intelligence, machine learning, and related fields. His previous research includes the derivation of new mathematical models for representing human motion, and space robotics.

Darren Leigh
Ph.D., Harvard University, 1998
Research Scientist
MERL Research Lab

Darren Leigh's research interests range from electronic hardware and embedded systems to signal processing, RF and communications. Before coming to MERL, he worked on the Harvard University/Planetary Society Billion-channel ExtraTerrestrial Assay (Project BETA), a search for microwave signals from extraterrestrial civilizations (SETI). Other previous research included 3D microscopic scanning, desktop manufacturing and network architectures for multimedia. His current research includes the DiamondTouch multi-user touch technology, sensor networks and a plethora of gadgets.

Neal Lesh
Ph.D., University of Washington, 1998
Research Scientist
MERL Research Lab

Neal Lesh's research efforts aim to improve (or at least ease) the interaction between people and computers. His research projects include interactive optimization (the HuGS project), collaborative interface agents (the COLLAGEN project), and collaborative navigation of digital data (the Personal Digital Historian project). Before coming to MERL, he was a graduate student at the University of Washington with Oren Etzioni, and and a postdoc with James Allen at the University of Rochester.

Sergei Makar-Limanov
Ph.D., Stanford University, 1994
Principal Technical Staff
MERL Technology Lab

Sergei Makar-Limanov received his Bachelors degree from University of Chicago, and his PhD in Mathematics from Stanford University. Before joining MERL, Sergei worked on computer aided manufacturing for PTC Corporation; and supply management for Kewill PLC; and most recently on designing automated software scalability testing tools at Empirix Corporation. Sergei has joined MERL in May 2001 to work on the Concordia project. He is now working with the vision group on creating a database framework for video surveillance applications.

Koji Miyahara
M.Sc., Kyushu University, 1988
Senior Principal Technical Staff
MERL Technology Lab

Koji Miyahara received his B.Sc. and M.Sc. degrees in Information Science from Kyushu University in 1986 and 1988, respectively. He joined Mitsubishi Electric Corp. in 1988. He worked at the Information Technology R&D Center of Mitsubishi Electric Corp before he moved to MERL in 2002. He was a visiting researcher at the University of California, Irvine, from 1999 to 2000. His research interests include user interface, intelligent agents and information filtering.

Baback Moghaddam
Ph.D., Massachusetts Institute of Technology, 1997
Research Scientist
MERL Research Lab

Baback Moghaddam's research is in computational vision with focus on probabilistic visual learning, statistical modeling and pattern recognition with applications in biometrics and computer-human interfaces. Prior to MERL, Dr. Moghaddam was at the Vision and Modeling Group at the MIT Media Laboratory where he developed an automatic vision system that won DARPA's 1996 "FERET" face recognition competition. His previous research included fractal image compression, segmentation of synthetic aperture radar (SAR) imagery as well as designing a zero-gravity experiment for laser annealing of amorphous silicon which was flown aboard the US space shuttle in 1990.

Andreas F. Molisch
Ph.D., Technical University Vienna, 1994
Principal Member Technical Staff
MERL Technology Lab

Andy Molisch received his M. Sc., Ph.D., and habilitation degrees from the TU Vienna, Austria, in 1990, 1994, and 1999, respectively. His current research interests are multiple-antenna systems, wireless channel measurement and modeling, ultrawideband systems, and OFDM. He is active in standardization (IEEE 802.15, 3GPP, COST273), and has authored or co-authored two books, five book chapters, some 50 journal papers, and numerous conference papers. He is a senior member of IEEE.

Yves-Paul Nakache
M.Sc., E.S.I.E.E., 2000
Member Technical Staff
MERL Technology Lab

Yves-Paul Nakache received a French Engineering diploma equivalent to M.Sc. degree in Electrical Engineering in 2000 from E.S.I.E.E. (Ecole Supérieure d'Ingénieurs en Electrotechnique et Electronique), Paris, France. He joined MERL in 2000, where he is currently a Member of the Technical Staff. He works on Interference Cancellation and 3G CDMA systems. His current interests are in speech processing and wireless communications.

Daniel Nikovski
Ph.D., Carnegie Mellon University, 2002
Research Scientist
MERL Research Lab

Daniel Nikovski's research is focused on algorithms for reasoning, planning, and learning with probabilistic models. His current work is on the application of such algorithms to hard transportation problems such as group elevator control. Recent results include an efficient algorithm for exact estimation of expected waiting times of elevator passengers, a look-ahead algorithm for balancing the waiting times of existing and future passengers, a risk-averse scheduler for reducing the fraction of passengers waiting excessively, and a maximum-entropy algorithm for estimating complete origin-destination traffic matrices.

Philip Orlik
Ph.D., SUNY Stony Brook, 1999
Member Technical Staff
MERL Technology Lab

Philip Orlik received the B.E. degree in 1994 and the M.S. degree in 1997 both from the State University of New York at Stony Brook. In 1999 he earned his Ph. D. in electrical engineering also from SUNY Stony Brook. He joined MERL in August 2000, and is currently a member of technical staff. His research interests include wireless and optical communications, networking, queuing theory, and analytical modeling.

Kadir Peker
Ph.D., New Jersey Institute of Technology, 2001
Member Technical Staff
MERL Technology Lab

Kadir A. Peker received the B.S. degree from Bilkent University, Turkey in 1993, the M.S. degree from Rutgers University in 1996, and the Ph.D. degree from New Jersey Institute of Technology in 2001, all in Electrical Engineering. He worked at MERL as an intern for more than a year, and joined as a member of technical staff in 2000. His Ph.D. dissertation was on content based video indexing and summarization using motion activity. He has contributed to MPEG-7, published conference and journal papers, and filed patents based on his doctoral work. He also worked on home networks and multimedia networking for a year. He attended UPnP working groups during this period. His current research interests include video indexing, browsing and summarization, video presentation techniques, and video mining.

Georgiy Pekhteryev
M.Sc., from Kharkiv Aviation Institute, Ukraine, 1994
Member Technical Staff
MERL Technology Lab

Georgiy Pekkhteryev received M.Sc. degree in Aviation Engineering in 1994 from Kharkiv Aviation Institute, Kharkiv, Ukraine. He joined MERL in 2002, where he is currently a Member of the Technical Staff. His current interests are in wireless and wired home networking, network technologies.

Ron Perry
B.Sc., Bucknell University, 1981
Research Scientist
MERL Research Lab

Ron Perry joined MERL as a Research Scientist in 1998. Prior to that, he was a consulting engineer at DEC developing a three-dimensional rendering ASIC called Neon. Ron has consulted for many companies including Kodak, Atex, Adobe, Quark, and Apple over the last 20 years, developing software and hardware products in the areas of computer graphics, imaging, color, and desktop publishing. Ron's research interests include fundamental algorithms in computer graphics with occasional excursions in numerical analysis and protein folding.

Hanspeter Pfister
Ph.D., State University of New York at Stony Brook, 1996
Associate Director and Senior Research Scientist
MERL Research Lab

Hanspeter Pfister is Associate Director and Senior Research Scientist at MERL in Cambridge, MA. He is the chief architect of VolumePro, Mitsubishi Electric's real-time volume rendering hardware for PCs. His research interests include computer graphics, scientific visualization, and computer architecture. His work spans a range of topics, including point-based graphics, 3D photography, volume graphics, and computer graphics hardware. Hanspeter Pfister received his Ph.D. in Computer Science in 1996 from the State University of New York at Stony Brook. He received his M.S. in Electrical Engineering from the Swiss Federal Institute of Technology (ETH) Zurich, Switzerland, in 1991.

Erik Piip
Manager, Computer Network Services
MERL Headquarters

Erik is the manager of the Computer Network Services Group. The group supports MERL's computing and network infrastructure and end-users. Erik is responsible for identifying MERL wide strategic and tactical enhancements. In past lives, Erik worked for Digital Equipment in multiple roles; service delivery, support, fault management and analysis, and server management design. Other interests include using and building large telescopes and instrumentation and Amateur Radio.

Fatih Porikli
Ph.D., Polytechnic University, 2002
Member Technical Staff
MERL Technology Lab

My research interests are in the areas of video processing, computer vision, aerial image processing, 3-D depth estimation, texture segmentation, robust optimization, network traffic management, multi-camera systems, data mining, and digital signal filtering. I have received my Ph.D. degree from Polytechnic University, Brooklyn, NY in 2002. Before I joined to MERL in 2000, I worked for Hughes Research Labs, Malibu, CA (1999) and AT&T Research Labs, Holmdel, NJ (1997).

Stanley Pozerski
BA Computer Systems, Daniel Webster College 1987
Systems Network Administrator
MERL Headquarters

Stan's interests have followed the application of computers to a variety of manufacturing tasks including using PDP-11's to demonstrate control of multiple reactor chemical processes, using PCs for production testing, manufacturing chemicals and controlling multi-axis rotary assembly machines. More recently, Stan has been Systems Administrator of a CIM system for a semiconductor facility performing shop floor scheduling, data collection, and process monitoring. Currently, Stan supports Windows and Linux clients and servers, networking, and the wide variety of PC applications used at MERL.

Bhiksha Raj
Ph.D., Carnegie Mellon University, 2000
Research Scientist
MERL Research Lab

Dr. Bhiksha Raj joined MERL as a Staff Scientist. He completed his Ph.D. from Carnegie Mellon University (CMU) in May 2000. Dr. Raj works mainly on algorithmic aspects of speech recognition, with special emphasis on improving the robustness of speech recognition systems to environmental noise. His latest work is on the use of statistical information about speech for the automatic design of filter-and-sum microphone arrays. Dr. Raj has over thirty conference and journal publications and is currently in the process of publishing a book on missing-feature methods for noise-robust speech recognition.

Ramesh Raskar
Ph.D., University of North Carolina at Chapel Hill
Research Scientist
MERL Research Lab

Ramesh Raskar joined MERL as a Research Scientist in 2000. Prior to that, he was at the Office of the Future group at UNC's Computer Graphics lab. As part of his dissertation, he developed a framework for projector based 3D graphics by treating a projector as the dual of a camera. Current work includes topics from non-photorealistic rendering, computer vision and intelligent user interfaces. He is a member of the ACM and IEEE.

Charles Rich
Ph.D., Massachusetts Institute of Technology, 1980
Distinguished Research Scientist
MERL Research Lab

The thread connecting all of Dr. Rich's research has been to make interacting with a computer more like interacting with a person. As a founder and director of the Programmer's project at the MIT Artificial Intelligence Lab. in the 1980s, he pioneered research on intelligent assistants for software engineering. Dr. Rich joined MERL in 1991 as a founding member of the Research Lab. For the past several years, he has been working on a technology, called Collagen, for building collaborative interface agents based on human discourse theory. Dr. Rich is a Fellow and past Councilor of the American Assoc. for Artificial Intelligence. He was Program Co-Chair the 2004 Int. Conf. on Intelligent User Interfaces.

David Rudolph
M.S., University of Illinois, 1989
Principal Technical Staff
MERL Technology Lab

David Rudolph joined MERL in 1990. During this time he has contributed to several systems software projects, including the gm80 simulator. His last 6 years have been spent developing the "Network Replication" project, which is now being successfully marketed by Veritas Software Corporation as VVR (Veritas Volume Manager). This project was the 1998 winner of the Corporate R&D GM's Award for Excellence. Before joining MERL, David spent three years at Data General, interrupted by a two-year stint at the University of Illinois, where his research interests were system software and performance analysis for massively parallel architectures.

Kathy Ryall
Ph.D., Harvard University, 1997
Principal Technical Staff
MERL Technology Lab

Kathy Ryall's research interests focus on human-computer interaction, user interfaces and improving human-computer collaboration. Her current research is on the design of interfaces and interaction techniques to support multi-user collaboration on shared surfaces. She currently leads the DiamondTouch project, developing the infrastructure for MERL's multi-user, multi-touch technology, and coordinating collaborations with external groups. Prior to joining MERL, Kathy was an Assistant Professor of Computer Science at University of Virginia for three years.

Zafer Sahinoglu
PhD, New Jersey Institute of Technology, 2001
Technical Staff
MERL Technology Lab

Zafer received his B.S. in E.E. from Gazi Uni., Ankara, M.S. in BME and Ph.D. in EE from NJIT. He was awarded the Hashimoto Prize in 2002. He worked at AT&T Shannon Labs in 1999, and joined MERL in March 2001 where he is currently with the Video Technology. His research areas includes home networking, QoS in video streaming and multicasting, wireless image sensor networks, traffic self-similarity and biomedical signal processing. He has significant contributions to emerging MPEG-21 and ZigBee standards.

Bent Schmidt-Nielsen
B.S. University of California at San Diego, 1971
Senior Principal Technical Staff
MERL Technology Lab

Bent Schmidt-Nielsen has seven years of experience at Dragon Systems in applying speech recognition to useful products. Here at MERL he is paying a lot of attention to making speech interfaces robust and usable. Bent has very broad interests in science and technology. Among many other activities he has taught genetics at the University of Massachusetts at Boston and he has been a leader in the development of an easy to use mass market database.

Derek Schwenke
M.S., Worcester Polytechnic Institute, 1988
Principal Technical Staff
MERL Technology Lab

Derek Schwenke received his B.S.E.E. from Tulane and M.S.C.S. from Worcester Polytechnic Institute. At Raytheon (1984-1988) in Marlboro, MA. Derek worked on image processing and satellite communications systems. At MERL (1988) Derek worked on the design and simulation of the M80 and PXB1 CPUs, and software development using the OSF-DCE/Encina platform. He co-developed the OpenMQ™ message queuing system for MELCO. Derek's worked on the Concordia™ mobile agent project's Java security architecture and the Location Aware Systems project. Currently he is working on FormsTalk™ and Collagen™ multimodal interfaces and is an active member of the W3C VoiceXML and Multimodal working groups.

Huairong Shao
Ph.D., Tsinghua University, 1999
Research Scientist
MERL Research Lab

Huairong Shao has research interests in adaptive and reliable multimedia communications, QoS provision for the next generation wired and wireless Internet, pervasive computing and collaborative systems. Before joining MERL Huairong Shao worked with Microsoft Research Beijing and Redmond. He received his Ph.D. in Computer Science from Tsinghua University.

Chia Shen
Ph.D., University of Massachusetts, 1992
Associate Director & Senior Research Scientist
MERL Research Lab

Chia Shen is Associate Director and Senior Research Scientist at MERL Research Lab. Her research interests span from shared user interfaces and HCI, to distributed real-time and multimedia systems. Her current research is on user interface design and interaction techniques for multi-user collaboration on shared surfaces. She is leading two projects, DiamondSpin and Personal Digital Historian, in this area. Previously, Chia Shen has lead the MidART research project which has been successfully incorporated into several large distributed industrial plant control systems. MidART is a real-time middleware for applications where humans need to interact, control and monitor instruments and devices in a network environment through computer interfaces.

Samuel Shipman
M.Sc., Carnegie Mellon University, 1985
Principal Technical Staff
MERL Technology Lab

Sam Shipman received the M.S. degree in Computer Science from Carnegie Mellon University and the B.S. from UNC-Wilmington. His technical interests and background are in real-time and distributed operating systems research and development. At MERL, he has worked on the Network Replication and Open Community projects, and on smaller efforts related to MPEG-7, interactive surroundings, and fingerprint recognition.

Candace Sidner
Ph.D., Massachusetts Institute of Technology, 1979
Senior Research Scientist
MERL Research Lab

Candy Sidner is an expert in user interfaces, especially those involving speech and natural language understanding, and human and machine collaboration. Before coming to MERL, she had been a research staff member at Bolt Beranek Newman, Digital Equipment Corp., and Lotus Development Corp., and a visiting scientist at Harvard University. She is currently working on applying speech understanding technology to collaborative interface agents in the COLLAGEN project. Dr. Sidner was Chair of the 2001 International Conference on Intelligent User Interfaces and is a past President of the Association for Computational Linguistics. She is also a Fellow and past Councilor of the American Association for Artificial Intelligence.

Paris Smaragdis
Ph.D., Massachusetts Institute of Technology, 2001
Research Scientist
MERL Research Lab

Paris Smaragdis joined MERL late 2002 as a research scientist. His main interests are auditory scene analysis and self-organizing computational perception. Before coming to MERL he was a postdoctoral associate at MIT, where he also obtained his PhD degree in perceptual computing. His most recent work has been on sound source separation, multimodal statistics and audio classification.

Jay Thornton
Ph.D., University of Michigan, 1982
Group Manager, Computer Vision Applications
MERL Technology Lab

Jay Thornton's degree program was Mathematical Psychology. His doctoral work focused on perception and vision, and his thesis concerned channels mediating color vision. After a post doc at the University of Pennsylvania, he worked for Polaroid Corporation, first in the Vision Research Laboratory and then as manager of the Image Science Laboratory. At Polaroid he worked on problems in color reproduction, image quality, image processing, and half toning. At MERL since January 2002, he manages the Computer Human Observation project, and is excited about the computer vision problems that arise when computers analyze, measure, count, detect, and recognize people.

Jeroen van Baar
M.Sc., Delft University of Technology, 1998
Research Associate
MERL Research Lab

Jeroen van Baar has a MSc in Computer Science from Delft University of Technology, the Netherlands. His interests are in the fields of Computer Graphics, Scientific Visualization, Computer Vision and HCI. He first came to MERL as an intern in 1997, and after finishing his Masters, he joined MERL full-time in 1999 as a Research Associate. The projects he has been working on include points as rendering primitives, automatic keystone correction for projectors, and multi-projector displays on both planar and curved surfaces.

Anthony Vetro
Ph.D., Polytechnic University, 2001
Senior Principal Technical Staff
MERL Technology Lab

Anthony Vetro joined MERL in 1996, where he is currently a Senior Principal Member of the Technical Staff. He received the Ph.D. degree in Electrical Engineering from Polytechnic University in Brooklyn, NY, and his current research interests are related to the encoding and transport of multimedia content, with emphasis on video transcoding, rate-distortion modeling and optimal bit allocation. He has published a number of papers in these areas and has been an active participant in MPEG standards for several years, where he is now serving as Editor for MPEG-21 Part 7, Digital Item Adaptation.

Joseph Woelfel
M.S., Rutgers University, 1992
Principal Technical Staff
MERL Technology Lab

Before joining MERL in February 2001, Joe worked at Dragon Systems, where he led small teams developing an extensible voice architecture. In the years before that, Joe worked on the development of a statistical process control software package at GE-Fanuc. Joe earned a B.S. in Physics from SUNY Albany, and an M.S. in Communication and Information Science from Rutgers University. As Project Engineer for the Surveillance product, he will work on MERL's Interactive Surroundings Initiative.

Peter Wolf
B.S., Yale University, 1983
Senior Principal Technical Staff
MERL Technology Lab

Peter is an expert in Speech Technologies and a broad range of Software Engineering tools and practices. While Peter's role is often that of a technical expert and principal engineer, his main interest is the definition and creation of new products and services, made possible by new technologies. Peter is currently exploring the use of speech recognition to retrieve information with applications for cellphones, PDAs, automobiles and home entertainment.

David Wong
Ph.D., University of Connecticut, 1991
Group Manager, E-Services
MERL Technology Lab

David Wong is group manager of Human-Centric and Analytic Systems, a group concerned with developing next generation UI and data analysis technologies. His technical interests include various flavors of agent technology, data mining and visualization techniques, and distributed computing infrastructures. Prior to joining MERL in 1994, he worked for 4 years on developing advanced distributed transaction processing systems at Digital Equipment Corporation. In a prior life, he also taught and conducted research in chemical engineering for 4 years. David received his B.Sc. degree from Brown University in 1980 and his Ph.D. in computer science from the University of Connecticut in 1991.

Christopher R. Wren
Ph.D., Massachusetts Institute of Technology, 2000
Research Scientist
MERL Research Lab

Christopher Wren's research area is Perception for Human-Computer Interaction. While Chris' recent work has focused on using computer vision techniques to create systems that are visually aware of the user, his current interests also extend to include audio processing and other sensing modalities. Prior to coming to MERL, he was at the Vision & Modeling Group at the MIT Media Laboratory. As part of his dissertation work, he developed a system for combining physical models with visual evidence in real time to recover subtle models of human motion.

Kazuhiro Yamamoto
M.S., Shibaura Institute of Technology, 1991
Japanese Liaison
MERL Headquarters

Kazuhiro Yamamoto joined MELCO in 1991 and had worked on Intellectual Property administration jobs for nine years, for example, software license administration, technical contract control, reviewing of research agreement, IP training to freshmen, at Information Technology R&D Center. He was legal staff of IP-Center, before he move to MERL in 2000. He received a M.S. from Shibaura Institute of Technology in 1991

Jonathan Yedidia
Ph.D., Princeton University, 1990
Research Scientist
MERL Research Lab

Jonathan Yedidia's graduate work at Princeton (1985-1990) and post-doctoral work at Harvard's Society of Fellows (1990-1993) focused on theoretical condensed-matter physics, particularly the statistical mechanics of systems with quenched disorder. From 1993 to 1997, he was a professional chess player and teacher. He then worked at the internet startup company Viaweb, where he helped develop the shopping search engine that has since become Yahoo's shopping service. In 1998, Dr. Yedidia joined MERL. He is particularly interested in the development of new methods to analyze graphical models. His work has applications in the fields of artificial intelligence, digital communications, and statistical physics.

William Yerazunis
Ph.D., Rensselaer Polytechnic Institute, 1987
Research Scientist
MERL Research Lab

William Yerazunis has worked in a number of fields including: optics, vision processing, and signal processing (for General Electric's jet engine manufacturing); computer graphics (at Rensselaer's Center for Interactive Computer Graphics); artificial intelligence and parallel symbolic computation (for DEC's OPS5, XCON, and the successor products such as RuleWorks); radioastronomy and SETI (at Harvard University), transplant immunology (for the American Red Cross), virtual and augmented reality, realtime sensing and ubiquitous computing, and realtime statistical categorization of texts (the CRM114 Discriminator anti-spam system). He has appeared on numerous educational television shows, and holds 19 U.S. patents.

Fangfang Zhang
M.S., Brandeis University, 2000
Member Technical Staff
MERL Technology Lab

Fangfang Zhang received her Master in Software Engineering from Brandies University. She joined MERL in Jan 2001. She is currently working on Concordia project and speech project. Prior to joining MERL, Fangfang worked at CMGI, as a member of web-dialup service development team.

Jinyun Zhang
Ph.D., University of Ottawa, 1991
Senior Principal Member Technical Staff
MERL Technology Lab

Jinyun received her Ph.D. in Electrical Engineering from the University of Ottawa in the area of digital signal processing, where she was also a Visiting Scholar and worked on digital image processing. Prior to this she was a teacher/lecturer at Tsinghua University, Beijing, China. Jinyun worked for Nortel Networks for the past ten years where she held engineering and most recently management positions in the areas of VLSI design, Advanced Wireless Technology Development, Wireless Networks and Optical Networks. She has a broad technical background, specializing in system design, DSP algorithms, and real-time embedded S/W for wireless communications and DWDM optical networks. Jinyun joined the MERL in 2001.

Recent Major Publications

The following lists the 222 major publications by members of the MERL staff over the past two years. (This is an average of more than 1.7 papers per technical staff member per year.) A publication is considered major if it appeared in a refereed journal, a referred conference proceedings, or some other significant publication such as a book. For completeness, the list includes 63 papers that have been accepted for publication in the near future.

An asterisk (*) appears before the 74 (26%) publications that were subject to highly stringent selection criteria where they were published. Some venues (such as major journals and certain key conferences) are very selective in what they publish and some (such as workshops and many conferences) are not. There are good reasons to publish something in a non-selective venue, the most important of which being that a given workshop or conference may be the best place at which to expose a particular piece of work to the scientific community. However, the mere appearance of a piece of work in a non-selective venue says little if anything about the quality of the work. In contrast, getting a piece of work into a highly selective venue is a mark of distinction that says a lot about the quality of the work in the eyes of the scientific community.

As a basis for assessing the selectivity of various venues, the list below uses acceptance rates. For instance, certain key conferences such as CHI, ICCV, and SIGGRAPH accept only 20% or less of the papers submitted to them, rejecting many papers that in fact describe fine work. In contrast, many workshops and regional conferences accept 80% or more of the papers submitted, taking everything but the truly awful. The list below puts an asterisk before a conference or workshop paper only if the acceptance rate was less than 25%, or the paper received a best paper award. In addition, asterisks appear before papers in major archival Journals.

2003

* Molisch, A.F.; Petrovic, R., "Reduction of the Error Floor of Binary FSK by Nonlinear Frequency Discriminators", *IEEE Transactions on Wireless Communications*, To Appear December 2003, (TR2003-41).

* Xie, L.; Xu, P.; Chang, S-F.; Divakaran, A.; Sun, H., "Structure Analysis of Soccer Video with Domain Knowledge and Hidden Markov Models", *Pattern Recognition Letters*, Vol. 24, Issue 15, To Appear December 2003.

* Peker, K.A.; Divakaran, A., "Framework for Measurement of the Intensity of Motion Activity of Video Segments", *Journal of Visual Communications and Image Representation*, Vol. 14, Issue 4, To Appear December 2003.

* Win, M.Z.; Chriskos, G.; Molisch, A.F., "Selective Rake Diversity in Multipath Fading with Arbitrary Power Delay Profile", *IEEE Transactions on Wireless Communications*, To Appear Winter 2003.

* Molisch, A.F.; Zhang, X., "FFT-Based Hybrid Antenna Selection Schemes for Spatially Correlated MIMO Channels", *IEEE Communication Letters*, To Appear Winter 2003.

* Sahinoglu, Z.; Orlik, P., "Optimum Transmit Power Control and Power Compensation for Error Propagation in Relay Assisted Wireless Networks", *IEEE Global Communications Conference (Globecom)*, To Appear December 2003.

* Almers, P.; Tufvesson, F.; Molisch, A.F., "Measurement of Keyholes and Capacities in Multiple-Input Multiple-Output (MIMO) Channels", *IEEE Global Communications Conference (Globecom)*, To Appear December 2003.

* Zhang, X.; Molisch, A.F.; Kung, S.Y., "Antenna Selection Under Spatially Correlated MIMO Channels", *IEEE Global Communications Conference (Globecom)*, To Appear December 2003.

* Yu, J.; Yao, Y.D.; Zhang, J.; Molisch, A.F., "Reverse Link Capacity of Power-Controlled CDMA Systems with Antenna Arrays in a Multipath Fading Environment", *IEEE Global Communications Conference (Globecom)*, To Appear December 2003.

* Pang, B.; Shao, H-R.; Gao, W., "An Admission Control Scheme for End-to-End Statistical QoS Provision in IP Networks", *Journal of Computer Science and Technology*, Vol. 18, No. 3, To Appear November 2003.

* Wittenburg, K.; Forlines, C.; Lanning, T.; Esenther, A.; Harada, S.; Miyachi, T., "Rapid Serial Visual Presentation Techniques for Consumer Digital Video Devices", *ACM Symposium on User Interface Software and Technology (UIST),* To Appear November 2003.

* Sahinoglu, Z.; Vetro, A., "Mobility Characteristics for Multimedia Service Adaptation", *EURASIP Journal: Image Communications*, To Appear Fall 2003, (TR2003-53).

* Sahinoglu, Z.; Orlik, P., "Regenerator vs. Simple Relay with Power Compensation for Error Propagation", *IEEE Communication Letters*, To Appear Fall 2003, (TR2003-54).

* Almers, P.; Tufvesson, F.; Molisch, A.F., "Measurement of Keyhole Effect in a Wireless Multiple-Input Multiple-Output (MIMO) Channel", *IEEE Communication Letters*, To Appear Fall 2003, (TR2003-43).

* Molisch, A.F., "A Generic Model for MIMO Wireless Propagation Channels in Macro- and Microcells", *IEEE Transactions on Signal Processing*, To Appear Fall 2003, (TR2003-42).

Xie, L.; Chang, S-F.; Divakaran, A.; Sun, H., "Unsupervised Mining of Statistical Temporal Structures in Video in Video Mining", *Video Mining*, Edited by Rosenfeld, A.; Doermann, D.; Dementhon, D., ISBN: 1402 075 499, To Appear October 2003.

Divakaran, A.; Peker, K.A.; Radharkishnan, R.; Xiong, Z.; Cabasson, R., "Video Summarization Using MPEG-7 Motion Activity and Audio Descriptors", *Video Mining*, Rosenfeld, A.; Doermann, D.; DeMenthon, D., To Appear October 2003.

* Dai, H.; Molisch, A.F.; Poor, H.V., "Downlink Capacity of Interference-Limited MIMO Systems with Joint Detection", *IEEE Transactions on Wireless Communications*, Vol. 2, To Appear Fall 2003.

* Molisch, A.F.; Win, M.Z.; Winters, J.H., "Reduced Complexity Transmit/Receive Diversity Systems", *IEEE Transactions on Signal Processing*, To Appear Fall 2003, (TR2003-44).

Smaragdis, P.; Brown, J.C., "Non-negative Matrix Factorization for Polyphonic Music Transcription", *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, To Appear October 2003.

Molisch, A.F.; Win, M.Z.; Winters, J.H., "Performance of Reduced-Complexity Transmit/Receive-Diversity Systems", *International Symposium on Wireless Personal Multimedia Communications (WPMC)*, ISSN: 1347-6890, Vol. 2, pp. 738-742, October 2003.

Forester, J.R.; Pendergrass, M.; Molisch, A.F., "A UWB Channel Model for Ultrawideband Indoor Communications", *International Symposium on Wireless Personal Multimedia Communications (WPMC)*, To Appear October 2003.

Molisch, A.F.; Nakache, Y.P.; Orlik, P.; Zhang, J.; Wu, Y.; Gezici, S.; Kung, S.Y.; Poor, H.V.; Li, Y.G.; Sheng, H.; Haimovich, A., "An Efficient Low-Cost Time-Hopping Impulse Radio for High Data Rate Transmission", *International Symposium on Wireless Personal Multimedia Communications (WPMC)*, To Appear October 2003.

Li, Y.G.; Molisch, A.F.; Zhang, J., "Channel Estimation and Signal Detection for UWB", *International Symposium on Wireless Personal Multimedia Communications (WPMC)*, To Appear October 2003.

Moghaddam, B.; Lee, J.; Pfister, H. Machiraju, R., "Model-Based 3D Face Capture with Shape-from-Silhouettes", *IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG)*, To Appear October 2003, (TR2003-84).

* Shakhnarovich, P.; Viola, P.; Darrell, T., "Fast Pose Estimation With Parameter Sensitive Hashing", *IEEE International Conference on Computer Vision (ICCV)*, To Appear October 2003.

* Levin, A.; Viola, P.; Freund, Y., "Unsupervised Improvement of Visual Detectors Using Co-Training", *IEEE International Conference on Computer Vision (ICCV)*, To Appear October 2003.

* Viola, P.; Jones, M.J.; Snow, D., "Detecting Pedestrians Using Patterns of Motion and Appearance", *IEEE International Conference on Computer Vision (ICCV),* To Appear October 2003, (TR2003-90).

* Shen, C.; Everitt, K.M.; Ryall, K., "UbiTable: Impromptu Face-to-Face Collaboration on Horizontal Interactive Surfaces", *International Conference on Ubiquitous Computing (UbiComp)*, To Appear October 2003.

* Dietz, P.H.; Yerazunis, W.; Leigh, D., "Very Low Cost Sensing and Communication Using Bidirectional LEDS", *International Conference on Ubiquitous Computing (UbiComp)*, To Appear October 2003, (TR2003-35).

* Choi, Y.S.; Molisch, A.F.; Win, M.Z.; Winters, J.H., "Fast Antenna Selection Algorithms for MIMO Systems", *IEEE Vehicular Technology Conference (VTC)*, To Appear October 2003.

Sidner, C.L.; Lee, C.; Lesh, N., "Engagement Rules for Human-Robot Collaborative Interactions", *IEEE International Conference on Systems, Man & Cybernetics (CSMC)*, To Appear October 2003, TR2003-50).

* Zhou, S.; Chellappa, R.; Moghaddam, B., "Visual Tracking and Recognition Using Appearance-Based Modeling in Particle Filters", *IEEE Transactions on Image Processing (TIP)*, To Appear Fall 2003.

* Liu, J.; Li, B.; Shao, H.S.; Zhu, W.; Zhang, Y-Q., "A Proxy-Assisted Adaptation Framework for Object Video Multicasting", *IEEE Transactions on Circuits and Systems for Video Technology*, To Appear October 2003.

Pasadakis, N.; Gaganis, V.; Smaragdis, P., "Independent Component Analysis for Deconvolution of Overlapping HPLC Aromatic Peaks of Oil", *International Conference of Instrumental Methods of Analysis (Modern Trends and Applications) (IMA)*, To Appear September 2003.

Xiong, Z.; Radhakrishnan, R.; Divakaran, A., "Generation of Sports Highlights Using Motion Activity in Combination with a Common Audio Feature Extraction Framework", *IEEE International Conference on Image Processing (ICIP)*, To Appear September 2003.

Xie, L.; Chang, S-F.; Divakaran, A.; Sun, H., "Feature Selection for Unsupervised Discovery of Statistical Temporal Structures in Video", *IEEE International Conference on Image Processing (ICIP)*, To Appear September 2003.

Guillamet, D.; Moghaddam, B.; Vitria, J., "Higher-Order Dependencies in Local Appearance Models", *IEEE International Conference on Image Processing (ICIP)*, To Appear September 2003, (TR2003-93).

Kong, H.S.; Vetro, A.; Sun, H., "Combined Rate Control and Mode Decision Optimization for MPEG-2 Transcoding with Spatial Resolution Reduction", *IEEE International Conference on Image Processing (ICIP)*, To Appear September 2003.

Porikli, F., "Inter-Camera Color Calibration by Cross-Correlation Model Function", *IEEE International Conference on Image Processing (ICIP)*, To Appear September 2003, (TR2003-103).

Peker, K.A.; Divakaran, A., "Extended Framework for Adaptive Playback-Based Video Summarization", *SPIE Internet Multimedia Management Systems IV*, To Appear September 2003.

Radhakrishnan, R.; Xiong, Z.; Raj, B.; Divakaran, A., "Audio-Assisted News Video Browsing Using a GMM Based Generalized Sound Recognition Framework", *SPIE Internet Multimedia Management Systems IV*, To Appear September 2003.

Molisch, A.F.; Zhang, X.; Kung, S.Y.; Zhang, J., "FFT-Based Hybrid Antenna Selection Schemes for Spatially Correlated MIMO Channels", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, To Appear September 2003.

Wu, Y.; Molisch, A.F.; Kung, S.Y.; Zhang, J., "Impulse Radio Pulse Shaping for Ultra-Wide Bandwidth UWB Systems", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, To Appear September 2003.

Gu, D.; Zhang, J., "A New Measurement-Based Admission Control Method for IEEE802.11 Wireless Local Area Networks", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, To Appear September 2003.

Huang, Q.; Gu, D.; Zhang, J., "Energy/Security Scalable Public-Key Mobile Cryptosystem", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, To Appear September 2003.

Lao, D.; Horng, J.H.; Zhang, J., "Throughput Analysis for W-CDMA Systems with MIMO and AMC", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC),* To Appear September 2003, (TR2003-48).

Horng, J.H.; Li, L.; Zhang, J.; "Adaptive Transmit Weights for Performance Enhancement in Space-Time Transmit Diversity Systems", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC),* To Appear September 2003.

Fossorier, M.; Palanki, R.; Yedidia, J., "Multi-Step Interactive Decoding of Majority Logic Decodable Codes", *International Symposium on Turbo Codes and Related Topics*, To Appear September 2003, (TR2003-107).

Gezici, S.; Fishler, E.; Kobayashi, H.; Poor, V., "A Rapid Acquisition Technique for Impulse Radio", *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, To Appear August 2003, (TR2003-46).

Suzuki, N.; Iwata, M.; Sasakawa, K.; Komaya, K.; Nikovski, D.; Brand, M., "A Calculation Method of Car Arrival Time With Dynamic Programming in Elevator Group Control", *Conference of the Institute of Electrical Engineers of Japan,* To Appear August 2003.

Molisch, A.F.; Zhang, J.; Miyake, M., "Time Hopping Versus Frequency Hopping in Ultra Wideband Systems", *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, To Appear August 2003, (TR2003-92).

Molisch, A.F., "MIMO Systems with Antenna Selection: An Overview", *IEEE Radio & Wireless Conference (RAWCON)*, To Appear August 2003.

* Moghaddam, B.; Tian, Q.; Lesh, N.; Shen, C.; Huang, T. S., "Visualization and User-Modeling for Browsing Personal Photo Libraries", *International Journal of Computer Vision*, To Appear Fall 2003.

* Raskar, R.; van Baar, J.; Beardsley, P.; Willwacher, T.; Rao, S.; Forlines, C., "iLamps: Geometrically Aware and Self-Configuring Projectors", *ACM Transactions on Graphics (SIGGRAPH)*, To Appear July 2003, (TR2003-23).

Sidner, C.L.; Lee, C.; Lesh, N.B., "The Role of Dialog in Human Robot Interaction", *International Workshop on Language Understanding and Agents for Real World Interaction, Hokkaido University, Japan*, To Appear July 2003, (TR2003-63).

Zhou, S.; Chellappa, R.; Moghaddam, B., "Adaptive Visual Tracking and Recognition using Particle Filters", *IEEE International Conference on Multimedia and Expo (ICME)*, To Appear July 2003, (TR2003-95).

Kong, H.S.; Vetro, A.; Sun, H., "Coding Mode Optimization for MPEG-2 Transcoding with Spatial Resolution Reduction", *SPIE Conference on Visual Communications and Image Processing (VCIP)*, Vol. 5150, pp. 501-511, To Appear July 2003, (TR2003-99).

Zhou, J.; Shao, H-R.; Shen, C.; Sun, M-T., "FGS Enhancement Layer Truncation with Minimized Intra-Frame Quality Variation ", *IEEE International Conference on Multimedia and Expo (ICME)*, To Appear July 2003, (TR2003-16).

Porikli, F.M.; Divakaran, A., "Multi-Camera Calibration, Object Tracking and Query Generation", *IEEE International Conference on Multimedia and Expo (ICME)*, To Appear July 2003, (TR2003-100).

Xie, L.; Chang, S-F.; Divakaran, A.; Sun, H., "Unsupervised Discovery of Multilevel Statistical Video Structures Using Hierarchical Hidden Markov Models", *IEEE International Conference on Multimedia and Expo (ICME)*, To Appear July 2003, (TR2003-101).

Vetro, A.; Haga, T.; Sumi, K.; Sun, H., "Object-Based Coding for Long-Term Archive of Surveillance Video", *IEEE International Conference on Multimedia and Expo (ICME)*, To Appear July 2003.

Cheng, H.; Zhang, X.M, Shi, Y.Q.; Vetro, A.; Sun, H., "Rate Allocation for FGS-Coded Video Using Composite Rate-Distortion Analysis", *IEEE International Conference on Multimedia and Expo (ICME)*, To Appear July 2003.

* Lee, J.; Moghaddam, B.; Pfister, H.; Machiraju, R., "Silhouette-Based 3D Face Shape Recovery", *Graphics Interface*, June 2003, (TR2003-81).

Gu, D; Zhang, J., "QoS Enhancement in IEEE802.11 Wireless Local Area Networks", *IEEE Communications Magazine*, ISSN: 0163-6804, Vol. 41, Issue 6, pp. 120-124, June 2003, (TR-2003-67).

* Lee, J-W.; Vetro, A.; Wang, Y.; Ho, Y-S., "Bit Allocation for MPEG-4 Video Coding with Spatio-Temporal Tradeoffs", *IEEE Transactions on Circuits and Systems for Video Technology*, ISSN: 1051-8215, Vol. 13, Issue 6, pp. 488-502, June 2003, (TR2003-75).

Yedidia, J.; Chen, J.; Fossorier, M., "Representing Codes for Belief Propagation Decoding", *IEEE International Symposium on Information Theory*, pp.176, June 2003, (TR2003-106).

Wittenburg, K., Lanning, T., Forlines, C.; Esenther, A., "Rapid Serial Visual Presentation Techniques for Visualizing a 3rd Data Dimension", *International Conference on Human-Computer Interaction (HCI)*, June 2003, TR2003-20).

* Moghaddam, B.; Guillamet, D.; Vitria, J., "Local Appearance-Based Models using High-Order Statistics of Image Features", *IEEE Computer Society Conference on Computer Vision & Pattern Recognition (CVPR)*, ISSN: 1063-6919, Vol. 1, pp. 729-735, June 2003, (TR2003-85).

Nakane, K.; Otsuka, I.; Esumi, K.; Murakami, T.; Divakaran, A., "A Content-Based Browsing System for HDD and/or recordable DVD Personal Video Recorder", *IEEE International Conference on Consumer Electronics (ICCE)*, June 2003, (TR2003-89).

Kong, H.S.; Vetro, A.; Sun, H., "Optimal MPEG-2 Transcoding for DVD Recording System", *IEEE International Conference on Consumer Electronics (ICCE)*, June 2003, (TR2003-76).

Guillamet, D.; Moghaddam, B.; Vitria, J., "Modeling High-Order Dependencies in Local Appearance Models", *Iberian Conference on Pattern Recognition & Image Analysis*, June 2003, (TR2003-94).

* Nikovski, D.; Brand, M., "Decision-Theoretic Group Elevator Scheduling", *International Conference on Automated Planning and Scheduling (ICAPS)*, June 2003, (TR2003-61). [Best applied research paper at the conference.]

Nikovski, D.; Brand, M., "Non-Linear Stochastic Control in Continuous State Spaces by Exact Integration in Bellman's Equations", *International Conference on Automated Planning and Scheduling (ICAPS)*, June 2003, (TR2003-91).

Frisken, S.; Perry, R., "Simple and Efficient Traversal Methods for Quadtrees and Octrees", *Journal of Graphics Tools*, Vol. 7, Issue 3, May 2003.

Sun, H.; Vetro, A.; Asai, K., "Resource Adaptation Based on MPEG-21 Usage Environment Descriptions", *IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2003, (TR2003-58).

Gu, D.; Zhang, J., "Evaluation of EDCF Mechanism for QoS in IEEE 802.11 Wireless Networks", *World Wireless Congress*, May 2003, (TR2003-51).

van Baar, J.; Willwacher, T.; Rao, S.; Raskar, R., "Seamless Multi-Projector Display on Curved Screens", *Eurographics Workshop on Virtual Environments (EGVE)*, ISBN: 3-905673-00-2, pp. 281-286, May 2003, (TR2003-26).

Shao, H-R; Shen, C.; Gu, D.; Zhang, J.; Orlik, P., "Dynamic Resource Control for High-Speed Downlink Packet Access Wireless Channel", *International Conference on Distributed Computing Systems Workshops*, pp. 838-843, May 2003, (TR2003-60).

Kalva, H.; Vetro, A.; Sun, H., "Performance Optimization of the MPEG-2 to MPEG-4 Video Transcoder", *SPIE Conference on VLSI Circuits and Systems*, Vol. 5117, pp. 341-350, May 2003, (TR2003-57).

Cassioli, D.; Win, M.Z.; Vatalaro, F.; Molish, A.F., "Effects of Spreading Bandwidth on the Performance of UWB Rake Receivers", *IEEE International Conference on Communications (ICC)*, Vol. 5, pp. 3445-3549, May 2003, (TR2003-65).

Sheng, H.; Orlik, P.V.; Hiamovich, A.M.; Cimini, L.J.; Zhang, J., "On the Spectral and Power Requirements for Ultra-Wideband Transmission", *IEEE International Conference on Communications (ICC)*, Vol. 1, pp. 738-742, May 2003.

Zhou, J.; Shao, H-R; Shen, C.; Sun, M-T, "Multi-Path Transport of FGS Video", *IEEE Packet Video*, April 2003, (TR2003-10).

Zhu, L.; Ansari, N.; Sahinoglu, Z.; Vetro, A.; Sun, H., "Scalable Layered Multicast with Explicit Congestion Notification", *IEEE International Conference on Information Technology Coding and Computing*, pp, 331-335, April 2003, (TR2003-32).

Vlasic, D.; Pfister, H.; Molinov, S.; Grzeszczuk, R.; Matusik, W., "Opacity Light Fields: Interactive Rendering of Surface Light Fields with View-Dependent Opacity", *ACM Symposium on Interactive 3D Graphics (I3D)*, ISBN:1-58113-645-5, pp. 65-74, April 2003.

Molisch, A.F., "Effect of Far Scatterer Clusters in MIMO Outdoor Channel Models", *IEEE Vehicular Technology Conference (VTC)*, ISSN: 1090-3038, Vol. 1, pp. 534-538, April 2003, (TR2003-38).

Nakache, Y-P; Molisch, A.F., "Spectral Shape of UWB Signals - Influence of Modulation Format, Multiple Access Scheme and Pulse Shape", *IEEE Vehicular Technology Conference (VTC)*, ISSN: 1090-3038, Vol. 4, pp. 2510-2514, April 2003, (TR2003-40).

Horng, J.H., Li, L., Zhang, J., "Adaptive Space-Time Transmit Diversity for MIMO Systems", *IEEE Vehicular Technology Conference (VTC)*, ISSN: 1090-3038, Vol. 2, pp. 1070-1073, April 2003, (TR2003-37).

Almers, P.; Tufvesson, F.; Karlsson, P.; Molisch, A.F., "The Effect of Horizontal Array Orientation on MIMO Channel Capacity", *IEEE Vehicular Technology Conference (VTC)*, ISSN: 1090-3038, Vol. 1, pp. 34-38, April 2003, (TR2003-39).

Porikli, F.M., "Road Extraction by Point-wise Gaussian Models", *SPIE AeroSense Technologies and Systems for Defense and Security*, April 2003, (TR2003-59).

* Sahinoglu, Z.; Orlik, P., "Power Efficient Transmission of Layered Video Through Wireless Proxy Servers", *The IEE - Electronics Letters*, ISSN: 0013-5194, Vol. 39, Issue 8, pp. 698-699, April 2003, (TR2003-52).

* Lesh, N.; Mitzenmacher, M.; Whitesides, S., "A Complete and Effective Move Set for Simplied Protein Folding", *International Conference on Research in Computational Molecular Biology (RECOMB)*, ISBN: 1-58113-635-8, pp. 188-195, April 2003, (TR2003-03).

Lin, S.; Vetro, A.; Wang, Y., "Rate-Distortion Analysis of the Multiple Description Motion Compensation Video Coding Scheme", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-6149, Vol. 3, pp. 401-404, April 2003, (TR2003-27).

Raj, B.; Whittaker, E.W.D;, "Lossless Compression of Language Model Structure and Word Identifiers", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-6149, Vol. 1, pp. 388-391, April 2003.

Xiong, Z.; Radhakrishnan, R.; Divakaran, A.; Huang, T.;, "Comparing MFCC and MPEG-7 Audio Features for Feature Extraction, Maximum Likelihood HMM and Entropic Prior HMM for Sports Audio Classification", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-6149, Vol. 5, pp. 628-631, April 2003.

Xiong, Z.; Radhakrishnan, R.; Divakaran, A.; Huang, T.S., "Audio Events Detection Based Highlights Extraction from Baseball, Golf and Soccer Games in a Unified Framework", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-6149, Vol. 5, pp. 632-635, April 2003.

Smaragdis, P.; Casey, M., "Audio/Visual Independent Components", *International Symposium on Independent Component Analysis and Blind Source Separation*, pp. 709-714, April 2003.

Porikli, F.M.; Tuzel, O., "Fast Object Tracking by Adaptive Background Models and Mean-Shift Analysis", *International Conference On Computer Vision Systems (ICVS)*, April 2003, TR2003-36).

* Seltzer, M.L.; Raj, B., "Speech Recognizer Based Filter Optimization for Microphone Array Processing", *IEEE Transaction on Signal Processing*, ISSN: 1070-9908, Vol. 10, Issue 3, pp. 69-71, March 2003.

* Shen, C.; Lesh, N.; Vernier, F., "Personal Digital Historian: Story Sharing Around the Table", *ACM Interactions*, ISSN: 1072-5520, Vol. 10, Issue 2, pp. 15-22, March/April 2003, (TR2003-04).

* Vetro, A.; Wang, Y.; Sun, H., "Rate-Distortion Modeling for Multiscale Binary Shape Coding Based on Markov Random Fields", *IEEE Transactions on Image Processing*, ISSN: 1057-7149, Vol. 12, Issue 3, pp. 356-364, March 2003, (TR2003-31).

Porikli, F.M., "Human Body Tracking by Adaptive Background Models and Mean-Shift Analysis", *IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, March 2003, (TR2003-36).

* Vetro, A.; Christopoulos, C.; Sun, H., "Video Transcoding Architectures and Techniques: An Overview", *IEEE Signal Processing Magazine*, ISSN: 1053-5888, Vol. 20, Issue 2, pp. 18-29, March 2003.

Porikli, F.M., "Sensitivity Characteristics of Cross-Correlation Distance Metric and Model Function", *The John Hopkins University Conference on Information Sciences and Systems (CISS)*, March 2003.

Zhu, L.; Sahinoglu, Z.; Cheng, G.; Vetro, A.; Ansari, N.; Sun, H.;, "Proxy Caching for Video on Demand Systems in Multicast Networks", *The John Hopkins University Conference on Information Sciences and Systems (CISS)*, March 2003, (TR2003-54).

* Zhang, X.M.; Vetro, A.; Shi, Y.Q.; Sun, H., "Constant Quality Constrained Rate Allocation for FGS Coded Video", *IEEE Transactions on Circuits and Systems for Video Technology*, ISSN: 1051-8215, Vol. 13, Issue 2, pp. 121-130, February 2003, (TR2003-30).

Yedidia, J.S.; Freeman, W.T.; Weiss, Y., "Understanding Belief Propagation and Its Generalizations", *Exploring Artificial Intelligence in the New Millennium*, ISBN 1558608117, Chapter 8, pp. 239-236, January 2003, (TR2001-22).

* Freeman, W.T.; Tenenbaum, J.B.; Pasztor, E.C., "Learning Style Translation for the Lines of a Drawing", *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 22, Issue 1, pp. 33-46, January 2003.

Cabasson, R.; Divakaran, A., "Automatic Extraction of Soccer Video Highlights Using a Combination of Motion and Audio Features", *SPIE Conference on Storage and Retrieval for Multimedia Databases*, Vol. 5021, pp. 272-276, January 2003.

Divakaran, A.; Radhakrishnan, R.; Xiong, Z.; Casey, M., "Procedure for Audio-Assisted Browsing of News Video Using Generalized Sound Recognition", *SPIE Conference on Storage and Retrieval for Multimedia Databases*, Vol. 5021, pp. 160-166, January 2003.

Yin, P.; Vetro, A.; Xia, M.; Liu, B., "Rate-Distortion Models for Video Transcoding", *SPIE Conference on Image and Video Communications and Processing*, Vol. 5022, pp. 467-488, January 2003, (TR2003-28).

Zhang, X.M.; Vetro, A.; Sun, H.; Shi, Y-Q., "Adaptive Field/Frame Selection for High-Compression Coding", *SPIE Conference on Image and Video Communications and Processing*, Vol. 5022, pp. 583-593, January 2003, (TR2003-29).

Yerazunis, W.S., "Sparse Binary Polynomial Hashing and the CRM114 Discriminator", *MIT Spam Conference*, January 2003.

Yedidia, J.S.; Bouchaud, J-P., "Renormalization Group Approach to Error-Correcting Codes", *Journal of Physics, A: Mathematical and Genera,* Vol. 36, pp. 1267-1288, January 2003, (TR2001-19).

2002

Guillamet, D.; Moghaddam, B., "Joint Distribution of Local Image Features for Appearance Modeling", *IAPR Workshop on Machine Vision Applications (MVA)*, December 2002, (TR2002-54).

* Moghaddam, B.; Shakhnarovich, G., "Boosted Dyadic Kernel Discriminants", *Advances in Neural Information Processing Systems (NIPS)*, Vol. 15, December 2002, (TR2002-55).

* Yin, P.; Vetro, A.; Liu, B.; Sun, H, "Drift Compensation for Reduced Spatial Resolution Transcoding", *IEEE Transactions on Circuits and Systems for Video Technology*, ISSN: 1051-8215, Vol. 12, Issue 11, pp.1009-1020, November 2002, (TR2002-47). [Best full-length paper of the year in a journal].

Papavassiliou, S.; Xu, S.; Orlik, P.; Snyder, M.; Sass, P., "Scalability in Global Mobile Information Systems (GloMo): Issues, Evaluation Methodology and Experiences", *ACM Wireless Networks*, ISSN: 1022-0038, Vol. 8, Issue 6, pp. 637-648, November 2002.

* Esenther, A.; Forlines, C.; Ryall, K.; Shipman, S., "DiamondTouch SDK: Support for Multi-User, Multi-Touch Applications", *ACM Conference on Computer Supported Cooperative Work (CSCW)*, November 2002, (TR2002-48).

* Shen, C.; Lesh, N.B.; Vernier, F.; Forlines, C.; Frost, J., "Building and Sharing Digital Group Histories", *ACM Conference on Computer Supported Cooperative Work (CSCW)*, pp. 3, November 2002, (TR2002-07).

Vetro, A.; Cai, J.; Chen, C.W., "Rate-Reduction Transcoding Design for Wireless Video Streaming", *Wireless Communications and Mobile Computing*, Vol. 2, Issue 6, pp.625-641, October 2002.

Frisken, S.F.; Perry, R.N., "Efficient Estimation of 3D Euclidean Distance Fields from 2D Range Images", *Volume Visualization Symposia (VolVis)*, ISBN: 0-7803-7641-2, pp. 81-88, October 2002, (TR2002-34).

Molisch, A.F.; Steinbauer, M.; Asplund, H.; Mehta, N.B., "Backward Compatibility of the COST 259 Directional Channel Model", *International Symposium on Wireless Personal Multimedia Communications*, ISSN: 1347-6890, Vol. 2, pp. 549-553, October 2002.

Frenzel, M.; Marks, J.; Ryall, K., "First UIST Interface-Design Contest", *ACM Interactions*, ISSN: 1072-5520, pp. 57-62, October 2002, (TR2001-54).

Schwartz, S.C.; Zhang, J.; Gu, D., "Iterative Channel Estimation for OFDM with Clipping", *International Symposium on Wireless Personal Multimedia Communications*, ISSN: 1347-6890, Vol. 3, pp. 1304-1308, October 2002.

Guo, J.; Matsubara, F.; Cukier, J.I.; Kong, H-S., "DTV for Personalized Mobile Access and Unified Home Control", *Personal Wireless Communications (PWC)*, pp. 71-78, October 2002.

Balter, O.; Sidner, C.L., "Bifrost Inbox Organizer: Giving Users Control Over the Inbox", *ACM Nordic Conference on Human-Computer Interaction (NordiCHI)*, ISBN: 1-58113-616-1, pp. 111-118, October 2002.

Sidner, C.L.; Dzikovska, M., "Engagement between Humans and Robots for Hosting Activities", *International Conference on Multimodal Interfaces (ICMI)*, October 2002, (TR2002-37).

Sidner, C.L.; Dzikovska, M., "Human-Robot Interaction: Engagement Between Humans and Robots for Hosting Activities", *IEEE International Conference on Multimodal Interfaces*, pp. 123-128, October 2002, (TR2002-37).

Raskar, R.; Low, K-L., "Blending Multiple Views", *Pacific Conference on Computer Graphics and Applications (PG)*, pp. 145-153, October 2002.

* Yedidia, J.S.,; Chen, J.; Fossorier, M., "Generating Code Representations Suitable for Belief Propagation Decoding", *Annual Conference on Communications, Control and Computing*, October 2002, (TR2002-40).

Raj, B.; Singh, R., "Classifier-Based Non-Linear Projection for Adaptive Endpointing of Continuous Speech", *Computer Speech and Language*, September 2002

* Zwicker, M.; Pfister, H.; van Baar, J.; Gross, M., "EWA Splatting", *IEEE Transactions on Visualization and Computer Graphics*, ISSN: 1077-2626, Vol. 8, Issue 3, pp. 223-238, July/September 2002, (TR2002-49).

Bimber, O.; Gatesy, S.M.; Witmer, L.M.; Raskar, R.; Encarnacao, L., "Merging Fossil Specimens with Computer-Generated Information", *IEEE Computer*, ISSN: 0018-9162, Vol. 35, Issue 9, pp. 25-30, September 2002.

* Wren, C.R.; Reynolds, C.J., "Parsimony & Transparency in Ubiquitous Interface Design", *International Conference on Ubiquitous Computing (UBICOMP)*, pp. 31-33, September 2002, (TR2002-22).

Horng, J.H.; Vannucci, G.; Zhang, J., "Capacity Enhancement for HSDPA in W-CDMA System", *IEEE Vehicular Technology Conference (VTC)*, ISSN: 1090-3038, Vol. 2, pp. 661-665, September 2002.

Porikli, F.M., "Automatic Threshold Determination of Centroid-Linkage Region Growing by MPEG-7 Dominant Color Descriptors", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, Vol. 1, pp. 793-796, September 2002.

Sun, X.; Manjunath, B.S.; Divakaran, A., "Representation of Motion Activity in Hierarchical Levels for Video Indexing and Filtering", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, Vol. 1, pp. 149-152, September 2002.

Divakaran, A.; Radhakrishnan, R.; Peker, K.A., "Motion Activity-Based Extraction of Key-Frames from Video Shots", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, Vol. 1, pp. 932-935, September 2002.

Lee, J.W.; Vetro, A.; Wang, Y.; Ho, W.S., "MPEG-4 Video Object-Based Rate Allocation with Variable Temporal Rates", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, Vol. 3, pp. 701-704, September 2002.

Vetro, A.; Chen, C.W., "Rate-Reduction Transcoding Design for Wireless Video Streaming", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, Vol. 1, pp. 29-32, September 2002.

Sidner, C.L.; Forlines, C., "Subset Languages for Conversing with Collaborative Interface Agents", *International Conference on Spoken Language Processing (ICSLP)*, September 2002, (TR2002-36).

Zhang, J.; Li, L.; Poon, T.C., "Turbo Coded HSDPA Systems with Transmit Diversity over Frequency Selective Fading Channels", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Vol. 1, pp. 90-94, September 2002.

* Almers, P.; Tufvesson, R.; Edfors, O.; Molisch, A.F., "Measured Capacity Gain Using Waterfilling in Frequency Selective MIMO Channels", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Vol. 3, pp.1347-1351, September 2002. [Best paper at the conference.]

Divakaran, A., "Video Summarization and Indexing Using Combinations of the MPEG-7 Motion Activity Descriptor and Other Audio-Visual Descriptors", *Thyrrhenian International Workshop on Digital Communications (IWDC)*, September 2002.

Ren, L.; Pfister, H.; Zwicker, M., "Object Space EWA Splatting: A Hardware-Accelerated Approach to High-Quality Point Rendering", *Conference of the European Association for Computer Graphics (Eurographics),* September 2002, (TR2002-31).

Karpenko, O.; Hughes, J.F.; Raskar, R., "Free-Form Sketching with Variation Implicit Surfaces", *Conference of the European Association for Computer Graphics (Eurographics)*, September 2002, (TR2002-27).

* Vetro, A.; Hata, T.; Kuwahara, N.; Kalva, H.; Sekiguchi, S., "Complexity-Quality Analysis of Transcoding Architectures for Reduced Spatial Resolution", *IEEE Transactions on Consumer Electronics*, ISSN: 0098-3063, Vol. 48, Issue 3, pp. 515-521, August 2002, (TR2002-46). [Second best full-length paper of the year in the journal.]

Porikli, F.M.; Wang, Y., "Constrained Video Object Segmentation by Color Masks and MPEG-7 Descriptors", *IEEE International Conference on Multimedia and Expo (ICME)*, Vol. 1, pp. 441-444, August 2002.

Wolf, P.P.; Raj, B., "The MERL SpokenQuery Information Retrieval System: A System for Retrieving Pertinent Documents from a Spoken Query", *IEEE International Conference on Multimedia and Expo (ICME)*, Vol. 2, pp. 317-320, August 2002.

Moghaddam, B.; Tian, Q.; Lesh, N.B.; Shen, C.; Huang, T.S., "PDH: A Human-Centric Interface for Image Libraries", *IEEE International Conference on Multimedia and Expo (ICME)*, Vol. 1, pp. 901-904, August 2002, (TR2002-52).

Shao, H.R.; Shen, C.; Kong, H-S., "Adaptive Pre-Stored Video Streaming with End System Multicast", *International Conference on Signal Processing (ICSP)*, Vol. 1, pp. 917-920, August 2002, (TR2002-44).

Kong, H-S.; Guan, L.; Kung, S.Y., "A Self-Organizing Tree Map Approach for Image Segmentation", *International Conference on Signal Processing (ICSP)*, Vol. 1, pp. 588-591, August 2002

Moghaddam, B.; Zhou, X., "Factorized Local Appearance Models", *IEEE International Conference on Pattern Recognition (ICPR)*, ISSN: 1051-4651, Vol. 3, pp. 553-556, August 2002, (TR2002-51).

Yedidia, J.S., Freeman, W.T.; Weiss, Y., "Constructing Free Energy Approximations and Generalized Belief Propagation Algorithms", August 2002, (TR2002-35).

Gonzalez, O.; Ramamritham, K.; Shen, C.; Fohler, G., "Building Reliable Component-Based Software Systems", *Artech House Publishers, ISBN 1-58053-327-2*, Chapter 15, July 2002.

Divakaran, A.; Asai, K., "Video Browsing System for Personal Video Recorders", *Multimedia Systems and Applications V*, Vol. 4861, pp. 22-25, July 2002.

Kong, H.S.; Vetro, A.; Kalva, H.; Fu, D.; Zhang, X.: Guo, J.; Sun, H.;, "A Platform for Real-Time Content Adaptive Video Transmission Over Heterogeneous Networks", *SPIE Multimedia Systems and Applications V*, Vol. 4861, pp. 43-49, July 2002.

Peker, K.A.; Divakaran, A., "Framework for Measurement of the Intensity of Motion Activity of Video Segments", *SPIE Internet Multimedia Management Systems III*, Vol. 4862, pp. 126-137, July 2002.

* Klau, G.W.; Lesh, N.B.; Marks, J.W.; Mitzenmacher, M., "Human-Guided Tabu Search", *National Conference on Artificial Intelligence (AAAI)*, ISBN: 0-262-51129-0, pp. 41-47, July 2002, (TR2002-09).

* Garland, A.; Lesh, N.B., "Plan Evaluation with Incomplete Action Descriptions", *National Conference on Artificial Intelligence (AAAI)*, ISBN: 0-262-51129-0, pp. 461-467, July 2002, (TR2002-05).

* Matusik, W.; Pfister, H.; Beardsley, P.A.; Ngan, A.; Ziegler, R.; McMillan, L., "Image-Based 3D Photography Using Opacity Hulls", *ACM SIGGRAPH*, ISSN:0730-0301, pp. 427-437, July 2002, (TR2002-29).

Wong, D., "Agent-Oriented Approach to Work Order Management for Circuit Breaker Maintenance", *IEEE Power Engineering Society Summer Meeting (PES)*, Vol. 1, pp. 553-558, July 2002.

Tian, Q.; Moghaddam, B.; Huang, T.S., "Visualization, Estimation and User-Modeling for Interactive Browsing of Image Libraries", *International Conference on Image and Video Retrieval (CIVR)*, July 2002, (TR2002-53).

* Rich, C.; Lesh, N.B.; Rickel, J.; Garland, A., "A Plug-in Architecture for Generating Collaborative Agent Responses", *International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, ISBN: 1-58113-480-0, pp. 782-789, July 2002, (TR2002-10).

You, C.; Gu, D.; Zhang, J., "Timing Synchronization for OFDM", *World Multi-Conference on Systemics, Cybernetics and Informatics (ISAS-SCI)*, July 2002.

Porikli, F.M.; Sahinoglu, Z., "An Online Renegotiation-Based Bandwidth Management with Circuit Assignment for VBR Traffic in Communications Networks", *World Multi-Conference on Systemics, Cybernetics and Informatics (ISAS-SCI)*, July 2002.

Cao, L.; Orlik, P.V.; Gu, D.; Zhang, J., "A Trellis Based Technique for Blind Channel Estimation and Signal Detection", *World Multi-Conference on Systemics, Cybernetics and Informatics (ISAS-SCI)*, July 2002.

Radhakrishnan, R.; Divakaran, A., "Automatic Detection of Talking Head Segments in News Video in the Compressed Domain", *World Multi-Conference on Systemics, Cybernetics and Informatics (ISAS-SCI)*, July 2002.

Li, L.; Zhang, J.; Horng, J.H., "Channel Estimation Error Impact on the Link Level Performance of Space-Time Coded Systems", *World Multi-Conference on Systemics, Cybernetics and Informatics (ISAS-SCI)*, July 2002.

Orlik, P.V.; Derand, T.; Horng, J.H.; Wang, D., "Study of Rate 1/4 HSDPA Turbo Codes and Scaling Effect", *World Multi-Conference on Systemics, Cybernetics and Informatics (ISAS-SCI)*, July 2002.

Radhakrishnan, R.; Divakaran, A., "A Date Reduction Procedure for 'Principal Cast' and other 'Talking Head' Detection", *World Multi-Conference on Systemics, Cybernetics and Informatics (ISAS-SCI)*, July 2002.

Kong, H-S.; Guan, L.; Kung, S.Y., "An Efficient Codebook Design Method of Vector Quantizer via Using Self-Organizing Tree Map", *World Multi-Conference on Systemics, Cybernetics and Informatics (ISAS-SCI)*, pp. 516-519, July 2002.

Kuwahara, N.; Hata, T.; Vetro, A., "Map-Top Video Surveillance with Adaptive Video Delivery", *World Multi-Conference on Systemics, Cybernetics and Informatics (ISAS-SCI)*, July 2002

* Moghaddam, B., "Principal Manifolds and Probabilistic Subspaces for Visual Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, ISSN: 0162-8828, Vol. 24, Issue 6, pp. 780-788, June 2002, (TR2002-13).

Matusik, W.; Pfister, H.; Ngan, A.; Ziegler, R.; McMillan, L., "Acquisition and Rendering of Transparent and Refractive Objects", *Eurographics Workshop on Rendering (EGRW)*, ISBN: 1-58113-434-3, pp. 267-278, June 2002, (TR2002-30).

Moghaddam, B.; Zhou, X., "Factorization for Probabilistic Local Appearance Models", *IASTED International Conference on Signal Processing, Pattern Recognition and Applications (SPPRA)*, June 2002, (TR2002-50).

Sparacino, F.; Wren, C.; Azarbayejani, A.; Pentland, A., "Browsing 3-D Spaces with 3-D Vision: Body-Driven Navigation Through the Internet City", *IEEE International Symposium on 3D Data Processing Visualization and Transmission*, pp. 224-231, June 2002.

* Vetro, A.; Hata, T.; Kuwahara, N.; Kalva, H., "Complexity-Quality Analysis of MPEG-2 to MPEG-4 Transcoding Architectures", *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 130-131, June 2002. [Second best paper at the conference.]

Lin, Y-C.; Wang, C-N.; Chiang, T.; Vetro, A.; Sun, H., "Efficient FGS to Single Layer Transcoding", *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 134-135, June 2002.

Horng, J.H.; Zhang, J., "Signal Circulation for Adaptive Equalization in Digital Communication Systems", *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 300-301, June 2002.

Sahinoglu, Z.; Peker, K.A.; Matsubara, F.; Cukier, J., "A Mobile Network Architecture with Personalized Instant Information Access", *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 34-35, June 2002.

Divakaran, A.; Cabasson, R., "Content-Based Browsing System for Personal Video Recorders", *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 114-115, June 2002.

Quigley, A.; Leigh, D.L.; Lesh, N.B.; Marks, J.W.; Ryall, K.; Wittenburg, K.B., "Semi-Automatic Antenna Design via Sampling and Visualization", *IEEE Antennas and Propagation Society International Symposium*, Vol. 2, pp. 342-345, June 2002, (TR2002-02).

Dietz, P.H.; Leigh, D.L.; Yerazunis, W.S., "Wireless Liquid Level Sensing for Restaurant Applications", *IEEE Sensors*, Vol. 1, pp. 715-720, June 2002, (TR2002-21).

Rickel, J.; Lesh, N.B.; Rich, C.; Sidner, C.L.; Gertner, A., "Collaborative Discourse Theory as a Foundation for Tutorial Dialogue", *International Conference on Intelligent Tutoring Systems (ITS)*, Vol. 2363, pp. 542-551, June 2002.

Raskar, R.; Ziegler, R.; Willwacher, T., "Cartoon Dioramas in Motion", *International Symposium on Non-Photorealistic Animation and Rendering (NPAR)*, ISBN: 1-58113-494-0, pp. 7-ff, June 2002, (TR2002-28).

Brand, M.E., "Subspace Mappings for Image Sequences", *Statistical Methods in Video Processing*, June 2002, (TR2002-25).

Singh, R.; Stern, R.M.; Raj, B., "Signal and Feature Compensation Methods for Robust Speech Recognition", *CRC Handbook on Noise Reduction in Speech Applications*, May 2002.

Singh, R.; Raj, B.; Stern, R.M., "Model Compensation and Matched Condition Methods for Robust Speech Recognition", *CRC Handbook on Noise Reduction in Speech Applications*, May 2002.

* Moghaddam, B.; Yang, M-H., "Learning Gender with Support Faces", *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, ISSN: 0162-8828, Vol. 24, Issue 5, pp. 707-711, May 2002, (TR2002-12).

Gu, D.; Zhang, J.; You, C.; Miyake, M.; Ma, X., "An OFDM Receiver Structure for Broadband Wireless LANs", *International Conference on Third Generation Wireless and Beyond (3Gwireless)*, May 2002.

Kung, S.Y.; Zhang, X.; Vannucci, G.; Zhang, J., "Sequential Bezout Equalizer for MIMO Systems", *International Conference on Third Generation Wireless and Beyond (3Gwireless)*, May 2002.

Cao, L.; Zhang, J.; Chen, C.W,; Orlik, P.V.; Miyake, M., "Study of Two 1/4 Turbo Code Receivers in Non-Frequency Selective Fast Fading Channels", *International Conference on Third Generation Wireless and Beyond (3Gwireless)*, May 2002.

Ma, X.; Kobayashi, H.; Schwartz, S.C.; Zhang, J.; Gu, D., "Time Domain Channel Estimation for OFDM Using EM Algorithm", *International Conference on Third Generation Wireless and Beyond (3Gwireless)*, May 2002.

Brand, M.E., "Incremental Singular Value Decomposition of Uncertain Data with Missing Values", *European Conference on Computer Vision (ECCV)*, Vol. 2350, pp. 707-720, May 2002, (TR2002-24).

Pang, B.; Shao, H-R.; Zhu, W.; Gao, W., "An Edge-Based Admission Control for Differentiated Services Network", *Workshop on High Performance Switching and Routing (HPSR)*, May 2002.

Vetro, A.; Yin, P.; Liu, B.; Sun, H., "Reduced Spatio-Temporal Transcoding Using an Intra Refresh Technique", *IEEE International Symposium on Circuits and Systems (ISCAS)*, Vol. 4, pp. 723-726, May 2002.

Zhang, X.M.; Shi, Y.Q.; Chen, H.; Haimovich, A.M.; Vetro, A.; Sun, H., "Successive Packing Based Interleaver for Design Turbo Codes", *IEEE International Symposium on Circuits and Systems (ISCAS)*, Vol. 1, pp. 17-20, May 2002.

* Vernier, F.; Lesh, N.B.; Shen, C., "Visualization Techniques for Circular Tabletop Interfaces", *ACM Advanced Visual Interfaces (AVI)*, pp. 257-263, May 2002, (TR2002-01).

Klau, G.W.; Lesh, N.B.; Marks, J.W.; Mitzenmacher, M.; Schafer, G.T., "The HuGS Platform: A Toolkit for Interactive Optimization", *Advanced Visual Interfaces (AVI)*, pp. 324-330, May 2002, (TR2002-08).

Shakhnarovich, G.; Viola, P.A.; Moghaddam, B., "A Unified Learning Framework for Real-Time Face Detection and Classification", *IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, pp. 14-21, May 2002, (TR2002-23).

Xie, L.; Chang, S-F.; Divakaran, A.; Sun, H., "Structure Analysis of Soccer Video with Hidden Markov Models", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-6149, Vol. 4, pp. 4096-4099, May 2002.

Vetro, A.; Wang, Y.; Sun, H., "A Probabilistic Approach for Rate-Distortion Modeling of Multi-scale Binary Shape", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-6149, Vol. 4, pp. 3353-3356, May 2002.

Seltzer, M.; Raj, B.; Stern, R., "Speech Recognizer Based Microphone Array Processing for Robust Hands-Free Speech Recognition", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-62149, Vol. 1, pp.897-900, May 2002.

Cao. L.; Chen, C.W.; Zhang, J.; Orlik, P.V.; Gu, D., "Blind Channel Estimation and Equalization Using Viterbi Algorithms", *IEEE Vehicular Technology Conference (VTC)*, Vol. 3, pp. 1532-1535, May 2002.

Jeannin, S.; Divakaran, B.; Mory, B., "Motion Descriptors", *Introduction to MPEG-7: Multimedia Content Description Interface*, Edited by Manjunath, B.S.; Salembier, P.; Sikora, T., ISBN: 0 471 48678 7, April 2002.

* Scott, S.D.; Lesh, N.; Klau, G.W., "Investigating Human-Computer Optimization", *ACM Conference on Human Factors in Computing Systems (CHI)*, ISBN: 1-58113-453-3, pp. 155-162, April 2002, (TR2001-39).

* Terry, M.; Mynatt, E.D.; Ryall, K.; Leigh, D.L., "Social Net: Using Patterns of Physical Proximity Over Time to Infer Shared Interests", *ACM Conference on Human Factors in Computing Systems (CHI)*, ISBN: 1-58113-454-1, pp. 816-817, April 2002.

Sahinoglu, Z.; Tekinay, S., "Efficient Parameter Selection for Use of Self-Similarity in Real Time Resource Management", *IEEE International Conference on Communications (ICC)*, Vol. 4, pp. 2212-2216, April 2002.

Pang, B.; Shao, H-R.; Zhu, W.; Gao, W., "An Admission Control Scheme to Provide End-to-End Statistical Qos Provision", *IEEE International Performance, Computing and Communications Conference (IPCCC),* April 2002.

Freeman, W.T.; Jones, T.R.; Pasztor, E.C., "Example-Based Super-Resolution", *IEEE Computer Graphics and Applications*, Vol. 22, Issue 2, pp.56-65, March/April 2002.

Wang, D., Orlik, P.V., Derand, T., Cao, L., Kobayashi, H.; Zhang, J., "Improved Low-Complexity Turbo Decoding for HSDPA Applications", *Annual Conference on Information Sciences and Systems (ISS)*, March 2002.

Vetro, A.; Devillers, S., "Position Paper: Delivery Context in MPEG-21", *W3C Delivery Context Workshop*, March 2002.

Yedidia, J.S., "Generalized Belief Propagation and Free Energy Minimization", *Information Theory Workshop at Mathematical Sciences Research Institute (MSRI)*, March 2002.

Divakaran, A.; Radhakrishnan, R., "Data Reduction Procedure for 'Principal-Cast' and Other 'Talking-Head' Detection", *SPIE Conference on Storage and Retrieval for Media Databases*, Vol. 4676, pp. 177-182, January 2002.

Peker, K.A.; Cabasson, R.; Divakaran, A., "Rapid Generation of Sports Video Highlights Using the MPEG-7 Motion Activity Descriptor", *SPIE Conference on Storage and Retrieval for Media Databases*, Vol 4676, pps 318-323, January 2002.

Raskar, R.; van Baar, J.; Chai, J.X., "A Low-Cost Projector Mosaic with Fast Registration", *Asian Conference on Computer Vision (ACCV)*, January 2002, (TR2002-14).

Yin, P.; Vetro, A.; Sun, H.; Liu, B., "Drift Compensation Architectures and Techniques for Reduced-Resolution Transcoding", *SPIE Conference on Visual Communications and Image Processing (VCIP)*, Vol. 4671, pp. 180-191, January 2002.

Zhang, X.M.; Vetro, A.; Shi, Y.Q.; Sun, H., "Constant-Quality Constrained-Rate Allocation for FGS Video Coded Bitstreams", *SPIE Conference on Visual Communications and Image Processing (VCIP)*, Vol. 4671, pp. 817-827, January 2002.

Lee, J.W.; Vetro, A.; Wang, Y.; Ho, Y.S., "Object-Based Rate Allocation with Spatio-Temporal Trade-Offs", *SPIE Conference on Visual Communications and Image Processing (VCIP)*, Vol. 4671, pp. 374-384, January 2002.

Bhatti, G.M.; Sahinoglu, Z.; Peker, K.A.; Guo, J.; Matsubara, F., "A TV-Centric Home Network to Provide a Unified Access to UPnP and PLC Domains", *IEEE International Workshop on Networked Appliances (IWNA)*, pp. 234-242, January 2002.

Eisenstein, J.; Rich, C., "Agents and GUI's from Task Models", *ACM International Conference on Intelligent User Interfaces*, ISBN: 1-58113-382-0 pp. 47-54, January 2002, (TR2002-16).

Sidner, C.L.; Dzikovska, M., "Hosting Activities: Experience with and Future Directions for a Robot Agent Host", *ACM International Conference on Intelligent User Interfaces*, ISBN: 1-58113-382-0, pp. 143-150, January 2002, (TR2002-03).

2001

Casey, M.A., "General Sound Similarity and Sound Recognition Tools", *Introduction to MPEG-7: Multimedia Content Description Language*, 2001.

Freeman, W.T.; Haddon, J.A.; Pasztor, E.C., "Learning Motion Analysis", *Statistical Theories of the Brain*, edited by R. Rao, B. Olshausen and M. Lewicki, 2001, (TR2000-32).

Rich, C.; Sidner, C.L.; Lesh, N.B., "COLLAGEN: Applying Collaborative Discourse Theory to Human-Computer Interaction", *Artificial Intelligence Magazine*, Vol. 22, Issue 4, pp. 15-25, Winter 2001, (TR2000-38).

* Brand, M.E., "Morphable 3D Models from Video", *IEEE Computer Vision and Pattern Recognition (CVPR)*, Vol. 2, pp. 456, December 2001, (TR2001-37). [Best paper at the conference.]

Raj, B.; Migdal, J.; Singh, R., "Distributed Speech Recognition with Codec Parameters", *IEEE Workshop on Automatic Speech Recognition and Understanding*, pp. 127-130, December 2001, (TR2001-45).

* Brand, M.E.; Bhotika, R., "Flexible Flow for 3D Nonrigid Tracking and Shape Recovery", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, ISSN: 1063-6919, Vol. 1, pp. 315-322, December 2001, (TR2001-37).

* Raskar, R.; Beardsley, P.A., "A Self-Correcting Projector", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, ISSN: 1063-6919, Vol. 2, pp. 504-508, December 2001, (TR2001-46).

Porikli, F.M., "Image Simplification by Robust Estimator Based Reconstruction Filter", *International Symposium on Computer and Information Sciences (ISCIS)*, November 2001.

Porikli, F.M., "Video Object Segmentation by Volume Growing Using Feature-Based Motion Estimator", *International Symposium on Computer and Information Sciences (ISCIS)*, November 2001.

Raskar, R.; Low, K-L, "Interacting with Spatially Augmented Reality", *ACM International Conference on Virtual Reality, Computer Graphics and Visualization in Africa (AFRIGRAPH)*, November 2001, (TR2001-51)

* Sahinoglu, Z., "Novel Adaptive Bandwidth Allocation Algorithm: Wavelet-Decomposed Signal Energy Approach", *IEEE Global Telecommunications Conference (Globecom)*, November 2001.

* Ye, D.; Shao, H-R.; Zhu, W.; Zhang, Y-Q., "Wavelet-Based Traffic Smoothing and Multiplexing for VBR Streams", *IEEE Global Telecommunications Conference (Globecom),* November 2001.

* Liu, J.; Shao, H-R.; Li, B.; Zhu, W.; Zhang, Y-Q., "A Scalable Bit-Stream Packetization and Adaptive Rate Control Framework for Object-Based Video Multicast", *IEEE Global Telecommunications Conference (Globecom),* November 2001.

Yerazunis, W.S.; Carbone, M., "Privacy-Enhanced Displays by Time-Masking Images", *Australian Conference on Computer-Human Interaction (OzCHI)*, November 2001, (TR2002-11).

* Dietz, P.H.; Leigh, D.L., "DiamondTouch: A Multi-User Touch Technology", *ACM Symposium on User Interface Software and Technology (UIST),* ISBN: 1-58113-438-X, pp. 219-226, November 2001.

* Dietz, P.H.; Yerazunis, W.S., "Real-Time Audio Buffering for Telephone Applications", *ACM Symposium on User Interface Software and Technology (UIST)*, ISBN: 1-58113-438-X, pp.193-194, November 2001.

* Divakaran, A.; Regunathan, R.; Peker, K.A., "Video Summarization Using Descriptors of Motion Activity: A Motion Activity Based Approach to Key-Frame Extraction from Video Shots", *Journal of Electronic Imaging*, Vol. 10, Issue 4, pp. 909-916, October 2001.

Bandyopadhyay, D.; Raskar, R.; Fuchs, H., "Dynamic Shader Lamps: Painting on Movable Objects", *IEEE and ACM International Symposium on Augmented Reality*, pp. 207-216, October 2001, (TR2001-50).

Wu, Y.; Vetro, A.; Sun, H.; Kung, S.Y., "Intelligent Multi-Hop Video Communications", *IEEE Pacific-Rim Conference On Multimedia (PCM)*, LNCS 2195, pp. 716, October 2001.

Sun, X.; Divakaran, A.; Manjunath, B.S., "A Motion Activity Descriptor and Its Extraction in the Compressed Domain", *IEEE Pacific-Rim Conference on Multimedia (PCM)*, LNCS 2195, pp. 450, October 2001.

Garland, A.; Ryall, K.; Rich, C., "Learning Hierarchical Task Models by Defining and Refining Examples", *ACM International Conference on Knowledge Capture (KCAP)*, ISBN: 1-58113-380-4, pp. 44-51, October 2001, (TR2001-26).

Zwicker, M.; Pfister, H.; van Baar, J.; Gross, M., "EWA Volume Splatting", *IEEE Visualization*, ISSN: 1070-2385, pp. 29-36, October 2001, (TR2001-27).

Porikli, F.M.; Sahinoglu, Z., "Dynamic Bandwidth Allocation with Optimal Number of Renegotiations in ATM Networks", *International Conference on Computer Communications and Networks*, pp. 290-295, October 2001.

Porikli, F.M., "Accurate Detection of Edge Orientation for Color and Multi-Spectral Imagery", *IEEE International Conference on Image Processing (ICIP)*, October 2001.

Porikli, F.M.; Wang, Y., "An Unsupervised Multi-Resolution Object Extraction Algorithm Using Video-Cubes", *IEEE International Conference on Image Processing (ICIP)*, October 2001.

Horng, J.H.; Nakache, Y-P., "Adaptive Diversity Combining for Interference Cancellation in CDMA Systems", *IEEE Vehicular Technology Conference (VTC)*, Vol. 3, pp. 1899-1902, October 2001.

Peker, K.A.; Divakaran, A.; Sun, H., "Constant Pace skimming and Temporal Sub-sampling of Video Using Motion Activity", *IEEE International Conference on Image Processing (ICIP)*, Vol. 3, pp. 414-417, October 2001.

Moghaddam, B.; Zhou, X.; Huang, T.S., "ICA-Based Probabilistic Local Appearance Models", *IEEE International Conference on Image Processing (ICIP)*, Vol. 1, pp. 161-164, October 2001, (TR2001-29).

Vetro, A.; Wang, Y.; Sun, H., "Rate-Distortion Optimized Video Coding Considering Frameskip", *IEEE International Conference on Image Processing (ICIP)*, Vol. 3, pp. 534-537, October 2001.

* Yedidia, J.S.; Sudderth, E.B.; Bouchaud, J-P., "Projection Algebra Analysis of Error-Correcting Codes", *Allerton Conference on Communication, Control, and Computing*, pp. 662-671, October, 2001, (TR2001-35).

* Moghaddam, B.; Nastar, C.; Pentland, A., "Bayesian Face Recognition with Deformable Image Models", *IEEE International Conference on Image Analysis & Processing (ICIAP)*, pp. 26-35, September 2001, (TR2001-05). [Winner of the 2001 Pierre Devijver Prixe from the International Association of Pattern Recognition (IAPR).]

Moghaddam, B.; Tian, Q.; Lesh, N.B.; Shen, C.; Huang, T.S., "Visualization and Layout for Personal Photo Libraries", *International Workshop on Content-Based Multimedia Indexing (CBMI)*, September 2001, (TR2001-28).

Tian, Q.; Moghaddam, B.; Huang, T.S., "Display Optimization for Image Browsing", *International Workshop Multimedia Databases and Image Communication*, Vol. 2184, pp. 167-178, September 2001.

Wang, D.; Kobayashi, H.; Zhang, J.; Gu, D., "On Bi-Directional SOVA Decoding of Turbo Codes", *Institute of Electronics Information and Communication Engineers Conference (IEICE)*, September 2001.

Divakaran, A., "An Overview of MPEG-7 Motion Descriptors and Their Applications", *International Conference on Computer Analysis of Images and Patterns (CAIP)*, LNCS 2124, pp. 29, September 2001.

Porikli, F.M., "Object Segmentation of Color Video Sequences", *International Conference on Computer Analysis of Images and Pattern (CAIP),* LNCS 2124, pp. 610, September 2001.

Casey, M.A., "Reduced-Rank Spectra and Entropic Priors as Consistent and Reliable Cues for General Sound Recognition", *Workshop for Consistent & Reliable Acoustic Cues (CRAC)*, September 2001.

Raj, B.; Seltzer, M.; Stern, R., "Robust Speech Recognition: The Case for Restoring Missing Features", *Workshop for Consistent and Reliable Acoustic Cues (CRAC)*, September 2001, (TR2001-40).

Kruja, E.; Marks, J.; Blair, A.; Waters, R.C., "A Short Note on the History of Graph Drawing", International Symposium on Graph Drawing, pp. 272-286, September 2001.

Casey, M.A., "General Sound Similarity and Recognition: Musical Applications of MPEG-7 Audio", *Organized Sound*, August 2001.

Xu, P.; Chang, S.F.; Divakaran, A.; Vetro, A.; Sun, H., "Algorithm and System for High-Level Structure Analysis and Event Detection in Soccer Video", *IEEE International Conference on Multimedia and Expo (ICME)*, August 2001.

Peker, K.A.; Divakaran, A., "A Novel Pair-Wise Comparison Based Analytical Framework for Automatic Measurement of Intensity of Motion Activity of Video", *IEEE International Conference on Multimedia and Expo (ICME)*, August 2001.

Vetro, A.; Wang, Y.; Sun, H., "Estimating Distortion for Coded and Non-Coded Frames for Frameskip-Optimized Video Coding", *IEEE International Conference on Multimedia and Expo (ICME)*, August 2001.

Qiu, J.; Shao, H-R.; Zhu, W.; Zhang, Y-Q., "An End-to-End Probing Based Admission Control Scheme in IP Networks", *IEEE International Conference on Multimedia and Expo (ICME),* August 2001.

Divakaran, A.; Bober, M.; Asai, K., "MPEG-7 Applications for Video Browsing and Analysis", *SPIE Multimedia Systems and Applications IV*, Vol. 4518, pp. 69-73, August 2001.

Frisken, S.F.; Perry, R.N., "A Computationally Efficient Framework for Modeling Soft Body Impact", *ACM SIGGRAPH*, August 2001, (TR2001-11).

Perry, R.N., "A New Method for Numerical Constrained Optimization", *ACM SIGGRAPH*, August 2001, (TR2001-14).

Pope, J.; Frisken, S.F.; Perry, R.N., "Dynamic Meshing Using Adaptively Sampled Distance Fields", *ACM SIGGRAPH*, August 2001, (TR2001-13).

Raskar, R., "Hardware Support for Non-Photorealistic Rendering: i.3.5", *SIGGRAPH/Eurographics Workshop on Graphics Hardware (HWWS)*, ISBN: 1-58113-407-X, pp. 410-46i.3, August 2001, (TR2001-23).

Jones, T.R.; Perry, R.N.; Callahan, M., "Shadermaps: A Method for Accelerating Procedural Shading", *ACM SIGGRAPH*, August 2001, (TR2000-25).

Frisken, S.F.; Perry, R.N., "Computing 3D Geometry Directly from Range Images", *ACM SIGGRAPH*, August 2001, (TR2001-10).

* Perry, R.N.; Frisken, S.F., "Kizamu: A System for Sculpting Digital Characters", *ACM SIGGRAPH*, ISBN: 1-58113-374-X, pp. 47-56, August 2001, (TR2001-08).

* Zwicker, M.; Pfister, H.; van Baar, J.; Gross, M., "Surface Splatting", *ACM SIGGRAPH*, ISBN: 1-58113-374-X, pp. 371-378, August 2001, (TR2001-20).

* Efros, A.A.; Freeman, W.T., "Image Quilting for Texture Synthesis and Transfer", *ACM SIGGRAPH*, ISBN: 1-58113-374-X, pp. 341-346, August 2001, (TR2001-17).

Rich, C.; Lesh, N.B.; Sidner, C.L., "Human-Computer Collaboration for Universal Access", *International Conference on Universal Access in Human-Computer Interaction (UAHCI)*, Invited Paper, August 2001.

Chen, P.; Horng, J.H.; Bao, J., Kobayashi, H., "A Joint Channel Estimation and Signal Detection Algorithm for OFDM System", *International Symposium on Signals, Systems and Electronics (ISSSE)*, July 2001.

Horng, J.; Nakache, Y-P., "Using Diversity Combining for Adaptive Interference Cancellation in CDMA Systems", *International Symposium on Signals, Systems and Electronics (ISSSE)*, July 2001.

* Lesh, N.B.; Rich, C.; Sidner, C.L., "Collaborating with Focused and Unfocused Users Under Imperfect Communication", *International Conference on User Modeling (UM)*, pp. 63-74, July 2001, (TR2000-39). [Outstanding Paper Award.]

* Bell, M.; Freeman, W.T., "Learning Local Evidence for Shading and Reflectance", *IEEE International Conference on Computer Vision (ICCV)*, Vol. 1, pp. 670-677, July 2001, (TR2001-04).

Cassell, J.; Nakano, Y.I.; Bickmore, T.W.; Sidner, C.L.; Rich, C., "Non-Verbal Cues for Discourse Structure", *Association for Computational Linguistics Annual Conference (ACL)*, pp. 106-115, July 2001, (TR2001-24).

Wu, H.; Shao, H-R.; Li, X., "Enhanced Communication Support for Multimedia Transmission Over the Differentiated Services Network", *International Conference on Networking (ICN),* July 2001

Project Reports

The body and soul of any research lab is the portfolio of projects it pursues. Therefore it is appropriate that the main body of this annual report consists of descriptions of the various projects being done at MERL. For ease of reference, the reports are grouped into eight topic areas.

- Computer Vision
- Digital Video
- Graphics Applications
- Interface Devices
- Optimization Algorithms
- Speech and Audio
- User Interface Software
- Wireless Communications and Networking

Each topical section begins with a short discussion of the topic area, highlighting MERL's major efforts. It then continues with a number of one-page project reports. These reports describe projects completed in the last twelve months and major milestones in continuing efforts.
The individual project reports begin with a brief summary at the top, followed by a more detailed discussion. The bottom of the report indicates the principal lab at MERL involved with the project and a contact person. Also included is a characterization of the type of project. The purpose of this is to indicate the kind of result that has been obtained.

- Initial Investigation – Work is underway on the project, but no firm results have been obtained yet. The project report is included to give a better understand of a direction in which MERL is heading.
- Research – The results obtained are in the form of papers, patents, and/or research prototypes. They represent valuable knowledge, but significant advanced development work will be required before this knowledge can be applied to products.
- Advanced Development – The results are (or will be) in forms that can be directly used in product development. The exact form of the result depends on what is being produced. For software projects, the results are typically code that can be directly used in products. For semiconductor chip projects, the results are typically in the form of detailed specifications for algorithms to be embedded in silicon.

Computer Vision

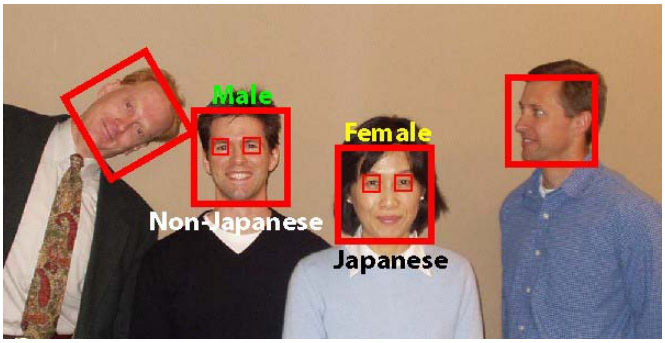
Computer Vision is the branch of computer science concerned with the analysis of images to extract information about the world. This is the same function that the human visual system provides (although perhaps accomplished through different mechanisms). As sensor and computer hardware drops in cost, these visual functions can become features in a wide range of products where they provide automatic, fast, convenient, and precise alternatives for tasks that were previously manual.

Much of the computer vision research at MERL is focused on the area of surveillance. For example, MERL has pioneered a state of the art approach to detecting object classes such as human faces in cluttered scenes. This approach uses a powerful machine-learning framework to automatically build very fast object detectors given a set of positive and negative examples of the object class. The same approach has been successfully applied to the problems of face detection, facial feature finding, face recognition, and gender and race classification. Another focus in the surveillance area is object tracking in video. Some of the work in tracking has used stereo cameras to track objects in 3-D. Other work has looked at the problem of tracking objects across different cameras in multi-camera systems.

The following project descriptions describe the many projects going on at MERL in the area of computer vision. They include work on biometric systems tracking systems, traffic analysis systems, intelligent video browsing systems and image matching systems. These systems are being applied to many areas of MELCO's businesses such as surveillance and security, consumer products (cell phones and DVD players) and elevators.

Project Descriptions

Face Detection/Gender & Race Classification



See Color Figure 1.

Automatic face detection is a critical component in the new domain of computer human observation and computer human interaction (HCI). There are many examples including: user-interfaces that can detect the presence and number of users; teleconference systems can automatically devote additional bandwidth to participant's faces; video security systems can record facial images of individuals after unauthorized entry; and indexing of image and video content can benefit from meta-data from face detection and classification. This is an extension of our previous work on fast face detection. Our new results include a system that can locate facial features such as the eyes, detect rotated and profile faces and determine the gender (male/female) and race (asian / non-asian) of the face.

Background and Objectives: Automatic face detection is a critical component in the new domain of computer human observation and computer human interaction (HCI). There are many examples including: user-interfaces that can detect the presence and number of users; teleconference systems can automatically devote additional bandwidth to participant's faces; video security systems can record facial images of individuals after unauthorized entry; and indexing of image and video content can benefit from meta-data from face detection and classification. This is an extension of our previous work on fast face detection. Our new results include a system that can locate facial features such as the eyes, detect rotated and profile faces and determine the gender (male/female) and race (asian / non-asian) of the face.

Technical Discussion: There are three main contributions of our face detection framework. First: a new image representation called an Integral Image that allows for very fast feature evaluation. Second: a method for constructing a classifier by selecting a small number of important features using AdaBoost. In order to ensure fast classification, the learning process must exclude a large majority of the available features, and focus on a small set of critical features. Third: a method for combining successively more complex classifiers in a cascade structure which dramatically increases the speed of the detector by focusing attention on promising regions of the image. The notion behind focus of attention approaches is that it is often possible to rapidly determine where in an image an object might occur. More complex processing is reserved only for these promising regions.

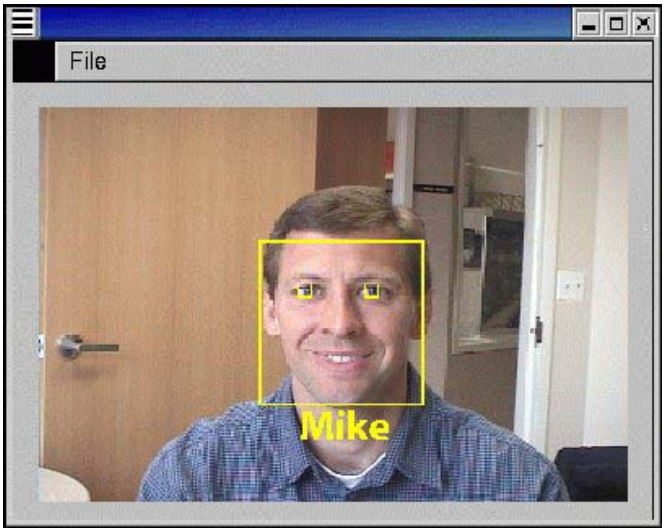
Collaboration: MTL, MRL, Sentansoken, Johosoken

Future Direction: We are currently extending this framework to use multiple video frames as input to the classifier. This extension will make it possible to detect pedestrians in a video sequence.

Contact: Michael Jones, Baback Moghaddam
http://www.merl.com/projects/FaceDetection/

Lab: MERL Technology Lab
Project Type: Research

Face Recognition



See Color Figure 2

The goal of this project is to have a computer recognize a person from an image of his or her face. There are many applications for face recognition. Some examples are: user-interfaces, summarizing surveillance video, indexing image and video databases, and access control. This is an extension of previous work at MERL on face detection and classification. The machine-learning framework used for face detection has been extended to solve the problem of face recognition.

Background and Objectives: This work builds on our previous work on face detection, which locates all faces in an image. We use the same machine learning framework developed for that problem and apply it to the problem of face recognition which determines who each face is. This requires extending the classifier to handle two input images at a time instead of just one. The problem of face recognition has been studied by a number of researchers. Our objective is to develop a state of the art system and give MELCO some intellectual property in this area.

Technical Discussion: The problem of face recognition can be solved using a similarity function that determines how similar two input faces are. In our approach we learn a similarity function from examples. The examples consist of pairs of images that are either of the same face or of different faces. Our similarity function consists of many simple features which we call rectangle features. These features are computationally efficient which makes our face recognizer very fast. The learning problem is to find the best set of features and feature parameters that separate the examples of same faces from the examples of different faces. The resulting features can then be used to judge the similarity of any two-face images.

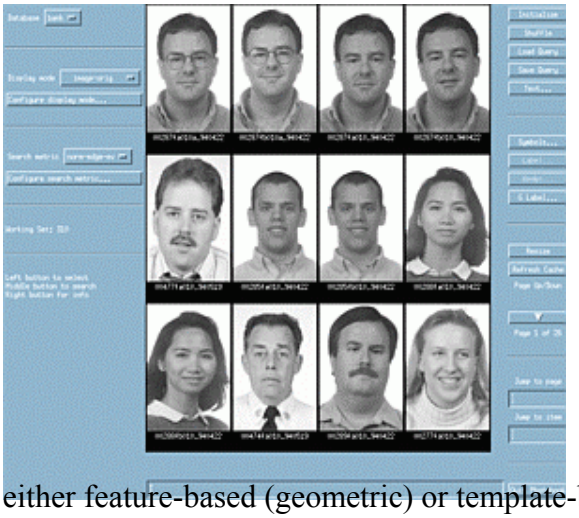
Collaboration: MTL, Sentansoken, Kamaden.

Future Direction: Future work will explore using 3-D head models to normalize the 2-Dface image in terms of its pose and illumination. Variations to pose and illumination are the biggest problems for 2-D face recognition

Contact: Michael Jones, Jay Thornton
http://www.merl.com/projects/FaceRecognition/

Lab: MERL Technology Lab
Project Type: Research

Probabilistic Modeling for Face Recognition



To assess the state-of-the-art in face recognition technology, the US Defense Advanced Research Project's Agency (DARPA) set up the "FERET" program and database to provide a common objective testbed for evaluating various algorithms. The system developed in this project was originally the winner of the 1996 FERET competition. It not only advanced the envelope of performance previously established by other techniques but also introduced a new modeling paradigm for face recognition ("intra/extra" learning) and was first to use probabilistic dual "eigenfaces."

Background and Objectives: Past techniques for face recognition can be categorized as either feature-based (geometric) or template-based (photometric), of which the latter have proved the more successful. Template-based methods use measures of facial similarity based on standard Euclidean error norms (e.g., template matching) or subspace-restricted error norms (e.g., eigenspace matching). The latter "eigenfaces" has in the past decade become the "golden standard" for comparison. The goal of this research was to improve on this standard benchmark while formulating a novel probabilistic similarity function for recognition.

Technical Discussion: We advocate a probabilistic measure of facial similarity, based on a Bayesian analysis of image differences: we model two mutually exclusive classes of variation between two facial images: intra-personal (variations in appearance of the same individual, due to different expressions or lighting, for example) and extra-personal (variations in appearance due to different identity). The high-dimensional probability density functions for each respective class are then obtained from training data and used to compute a similarity measure based on the a posteriori probability of membership in the intra-personal class. The performance advantage of of our technique over standard nearest-neighbor eigenspace matching was demonstrated using results from DARPA's 1996 "FERET" face recognition competition, in which this system was found to be the top performer. This framework is particularly advantageous in that the intra/extra classes learned explicitly characterize the type of appearance variations which are critical in formulating a meaningful measure of similarity (and a suitably nonlinear discriminant).

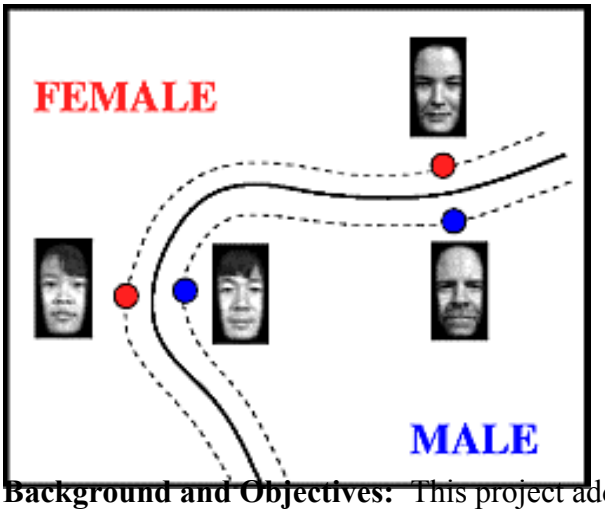
Collaboration: The "intra-extra" face modeling methodology originally developed in this project is being used in the new realtime 2D face recognition system of Mike Jones (MTL).

Future Direction: The core technology of this project can be easily extended to 3D face recognition (see the "3Dfacerec" project) as well as potentially other forms of biometrics.

Contact: Baback Moghaddam
http://www.merl.com/projects/face-rec/

Lab: MERL Research Lab
Project Type: Research

Support Vector Learning for Gender Classification



Computer vision systems for people monitoring will eventually play an important role in our lives by means of automated human (face) detection, body tracking, action (gesture) recognition, person identification and estimation of age and gender. We have developed a facial gender classifier using Support Vector Machine (SVM) learning with performance superior to existing gender classifiers. This technology can, for example, be used for passive surveillance and control in "smart buildings" as well as gender-mediated HCI.

Background and Objectives: This project addresses the problem of classifying gender from low-resolution images in which only the main facial regions appear (i.e., without hair information). We wanted to investigate the minimal amount of face information (resolution) required to learn male and female faces by various pattern classifiers. Previous studies on gender classification have relied on high-resolution images with hair information and used relatively small datasets for their experiments. In our study, we demonstrate that SVM classifiers are able to learn and classify gender from a large set of hairless low-resolution images with the highest accuracy.

Technical Discussion: A Support Vector Machine is a learning algorithm for pattern classification and regression. The basic principle behind SVMs is finding the optimal linear (or nonlinear) hyperplane (see above figure) such that the expected classification error for unseen test samples is minimized (i.e., good generalization performance). We investigated the utility of SVMs for visual gender classification with low-resolution "thumbnail" faces (21-by-12 pixels) processed from 1,755 images from the FERET face database. The performance of SVMs (3.4% error) is shown to be superior to traditional pattern classifiers (Linear, Quadratic, Fisher Linear Discriminant, Nearest-Neighbor) as well as more modern techniques such as Radial Basis Function (RBF) classifiers and large ensemble-RBF networks.

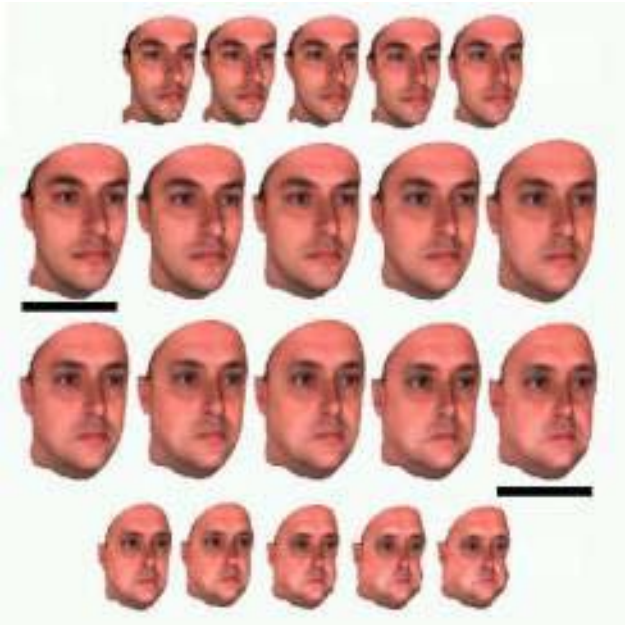
Collaboration: Joint work with Ming-Hsuan Yang (Honda Fundamental Research).

Future Direction: The SVM-based visual gender classifier developed is not invariant to head orientation and (like many other systems) requires fully frontal facial images. While it functions quite well with frontal views of the face, the inability to recognize gender from different viewpoints is a limitation which can affect its utility in unconstrained imaging environments. However, the system can be of use in application scenarios where frontal views are readily available (e.g., doorways, public kiosks, advertising billboards, etc.).

Contact: Baback Moghaddam
http://www.merl.com/projects/gender/

Lab: MERL Research Lab
Project Type: Research

Manifold of Faces



See Color Figure 3

The manifold of faces is a mathematical model of the set of all realistic-looking faces. We estimate this manifold from a few hundred high-resolution 3D face models, and impose a coordinate system on it so that every face has an "address" consisting of roughly 20 numbers. We can then compress and reconstruct faces with very high fidelity, and synthesize novel faces with very high realism. The image shows faces generated along a straight line on the manifold; underlined faces are real.

Background and Objectives: Good models of human faces are necessary for animation, security (recognition/ identification), and video coding applications. Previous attempts to model variability in human faces have depended on purely linear models. These can easily generate objects that are not face-like. We seek to estimate a compact function that efficiently codes (compresses) and decodes (reconstructs) realistic 3D graphical models of faces.

Technical Discussion: This project is a demonstration of charting - a new data reduction technology developed at MERL. Although it takes roughly one million numbers to specify a high-resolution 3D face model, it is very likely most of the visible differences between any two faces can be described with just a hundred numbers. In short, the set of all faces is a low-dimensional manifold that is embedded in a million-dimensional ambient space. We use charting to estimate the manifold from a small sample of faces. Charting constructs a low-dimensional coordinate system on the manifold, plus smooth functions that map between the ambient space and the new coordinate space. Operations that are difficult to do with raw face models become simple linear operations on the charted manifold: morphing (interpolation); caricature (extrapolation); comparison (distance).

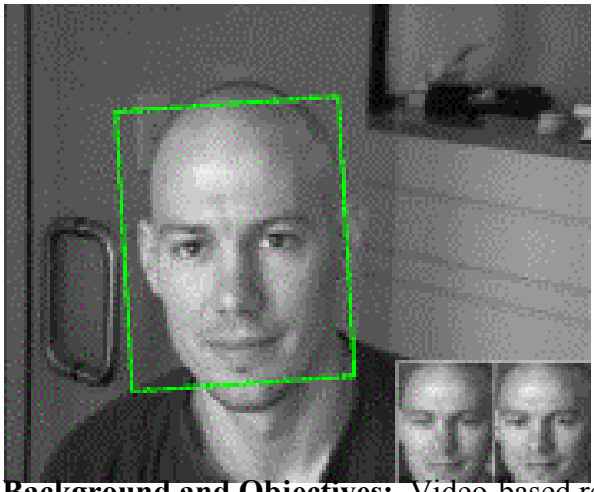
Collaboration: Yuechao Zhao, student, Harvard University

Future Direction: Charting can also model changes in appearance due to changes in expression and lighting. This will be evaluated for 3D facial recognition systems.

Contact: Matthew Brand
http://www.merl.com/projects/ManifoldFaces/

Lab: MERL Research Lab
Project Type: Research

Visual Tracking & Recognition with Particle Filters



We present a method for simultaneous tracking and recognition of visual objects from video using a time series model with stochastic diffusion. Specifically, by modeling the dynamics with a particle filter we are able to achieve a very stabilized tracker and an accurate recognizer when confronted by pose, scale and illumination variations. We have tested our tracker on real-time face detection (invariant to scale/rotation) as well as tracking vehicles from ground-level views as well as aerial surveillance video of tanks seen from oblique (affine) views.

Background and Objectives: Video-based recognition needs to handle uncertainties in both tracking and recognition. We have focused on face recognition for biometric or surveillance applications. We augment a time-series face tracker in the following ways: (i) Modeling the inter-frame motion and appearance changes within the video sequence; (ii) Modeling the appearance changes between the video frames and gallery images by constructing intra- and extra-personal spaces which can be treated as a 'generalized' version of discriminative analysis and (iii) Utilizing the fact that the gallery images are in frontal views

Technical Discussion: Tracking needs modeling inter-frame motion and appearance changes whereas recognition needs modeling appearance changes between frames and gallery images. In conventional tracking algorithms, the appearance model is either fixed or rapidly changing, and the motion model is simply a random walk with fixed noise variance (the number of particles is typically fixed). To stabilize the tracker, we propose the following features: an observation model arising from an adaptive appearance model, an adaptive velocity motions model with adaptive noise variance, and an adaptive number of particles. The adaptive-velocity model is derived using a first-order linear predictor based on the appearance difference between the incoming observation and the previous particle configuration. Occlusion analysis is implemented using robust statistics. Experimental results on tracking visual objects in long outdoor and indoor video sequences demonstrate the effectiveness and robustness of our tracking algorithm. We then perform simultaneous tracking and recognition by embedding them in one particle filter. For recognition purposes, we model the appearance changes between frames and gallery images by constructing the intra- and extra-personal spaces.

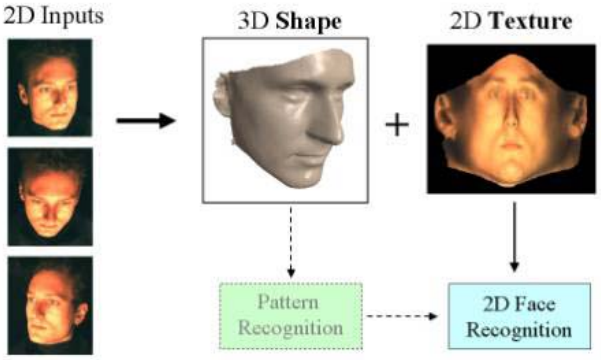
Collaboration: This project was joint work with Shaouha Zhou and Rama Chellappa (UMD).

Future Direction: We are planning further extensions to the adaptive velocity motion model as well as more complex PDFs for intra/extra similarity modeling (such as Gaussian Mixture Models). More experiments in the automotive domain are needed as well as an investigation into tracking pedestrians.

Contact: Baback Moghaddam
http://www.merl.com/projects/pftrack/

Lab: MERL Research Lab
Project Type: Research

3D Face Recognition



See Color Figure 4

The creation of realistic 3D face models is not only a fundamental problem in computer graphics but also the critical foundation for designing robust 3D face recognition systems in computer vision. We have developed a novel method to obtain the 3D shape (and texture) of an arbitrary human face using a sequence of 2D silhouette (binary) images as input. Our face model is a linear combination of "eigenheads" obtained by Principal Component Analysis (PCA) of a training set of laser-scanned 3D human faces. The coefficients of this linear decomposition are used as our model parameters. We have devised a near-automatic method for estimating these shape coefficients from simple 2D input images/video. We are currently applying our model to illumination- and pose-invariant face recognition.

Background and Objectives: A fundamental goal of this research is the application of 3D face models to the problem of robust face recognition in computer vision. In particular, we are concerned with addressing the two most critical and complicating factors affecting face recognition performance: illumination and pose variation, both of which can only be fully addressed with explicit use of 3D face models. We follow a data-driven approach to reconstruct accurate human face geometry from photographs. Our underlying face model is not synthetic but is based on real human faces measured by laser-based cylindrical scanners. Using our 3D face model we are able to obtain accurate and realistic 3D shape/reflectance models for computer graphics as well as the application of these 3D face models to the problem of robust face recognition in computer vision.

Technical Discussion: We focus on acquiring relatively accurate geometry of a face from multiple silhouette images at affordable cost and with no user interaction. Using silhouettes separates the geometric subtleties of the human head from the nuances of shading and texture. As a consequence, we do not require knowledge of external parameters such as light direction and light intensity. We developed a robust and efficient method to reconstruct human faces from silhouettes. We propose a novel algorithm for establishing correspondence between two faces and use a novel and efficient error metric (boundary weighted XOR) in our optimization procedures. The method is very robust even when presented with partial and noisy information.

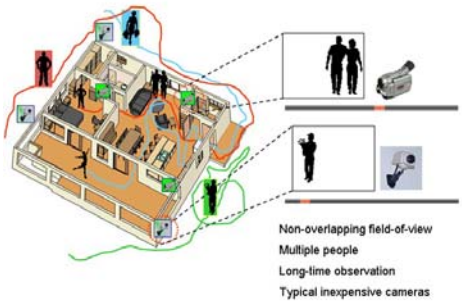
Collaboration: This is joint work between MERL and the Ohio-State University (OSU).

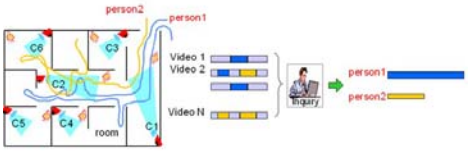
Future Direction: Having established and demonstrated superior performance in 3D scanning (shape and texture acquisition), our next phase will focus on robust 3D face recognition for various surveillance and biometric applications as part of MERL's "Computer-Human Observation" (CHO) project.

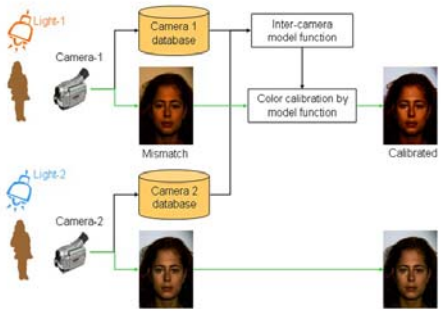
Contact: Hanspeter Pfister, Baback Moghaddam
http://www.merl.com/projects/3Dfacerec/

Lab: MERL Research Lab
Project Type: Initial Investigation

Multi-Camera Systems







The nature of single-camera single-room architecture multi-camera surveillance applications demands automatic and accurate calibration, detection of object of interest, tracking, fusion of multiple modalities to solve inter-camera correspondence problem, easy access and retrieving video data, capability to make semantic query, and effective abstraction of video content. Although several multi-camera setups have been adapted for 3D vision problems, the non-overlapping camera systems have not investigated thoroughly. Considering the huge amount of the video data a multi-camera system may produces over a short time period, more sophisticated tools for control, representation, and content analysis became an urgent need.

Background and Objectives: We designed a framework where we can extract the object-wise semantics from a non-overlapping field-of-view multi-camera system. This framework has four main components: camera calibration, automatic tracking, inter-camera data fusion, and query generation. We developed an object-based video content labeling method to restructure the camera-oriented videos into object-oriented results. We proposed a summarization technique using the motion activity characteristics of the encoded video segments to provide a solution to the storage and presentation of the immense video data. To solve the calibration problem, we developed a correlation matrix and dynamic programming based method.

Technical Discussion: An automatic object tracking and video summarization method for multi-camera systems with a large number of non-overlapping field-of-view cameras is developed. In this framework, video sequences are stored for each object as opposed to storing a sequence for each camera. Object-based representation enables annotation of video segments, and extraction of content semantics for further analysis. We also present a novel solution to the inter-camera color calibration problem. The transitive model function enables effective compensation for lighting changes and radiometric distortions for large-scale systems. After initial calibration, objects are tracked at each camera by background subtraction and mean-shift analysis. The correspondence of objects between different cameras is established by using a Bayesian Belief Network. This framework empowers the user to get a concise response to queries such as, “Which locations did an object visit on Monday and what did it do there?”

Collaboration: We are closely working with Johosoken and Sentansoken.

Future Direction: We are working on event detection for surveillance video. We will also integrate our face recognition technology into the multi-camera system.

Contact: Fatih Porikli **Lab:** MERL Technology Lab
http://www.merl.com/projects/MultiCamera/ **Project Type:** Research

Object Tracking & Understanding

See Color Figure 9

Object tracking is important because it enables several important applications such as: Security and surveillance - to recognize people, to provide better sense of security using visual information; Medical therapy - to improve the quality of life for physical therapy patients and disabled people; Retail space instrumentation - to analyze shopping behavior of customers, to enhance building and environment design; Video abstraction: to obtain automatic annotation of videos - to generate object-based summaries; Traffic management - to analyze flow, to detect accidents; Video editing - to eliminate cumbersome human-operator interaction, to design futuristic video effects; Interactive games - to provide natural ways of interaction with intelligent systems such as weightless remote control

Background and Objectives: The current object tracking project has four main components: 1) Adaptive background generation and shadow removal; 2) Single-camera tracking with mean-shift techniques; 3) Multi-camera radiometric calibration; 4) Gesture recognition, object-based summary generation using multi-camera tracking information. Accurate object segmentation and tracking under the constraint of low computational complexity presents a challenge. Our aim is to find solutions that are robust, simple, computationally feasible, modular, and easily adaptable to various applications.

Technical Discussion: We made the semi-automatic mean-shift tracker completely automatic using an improved GMM background subtraction method. We improved the adaptation performance of the original GMM by observing the amount of illumination change in the background. The performance of the background adaptation and mean-shift analysis based object tracking method is comparable with the state-of-art, and it is fully automatic. We invented solution to the inter-camera color calibration problem, which is very important for multi-camera systems. We also introduced a distance metric and a modeling function to evaluate the inter-camera radiometric properties.

Collaboration: Currently, we are closely working with Johosoken and Sentansoken

Future Direction: We are working on large-scale system management; behavior modeling, event detection issues.

Contact: Fatih Porikli
http://www.merl.com/projects/ObjectTracking/

Lab: MERL Technology Lab
Project Type: Research

Pedestrian Flow: Self-Configuring Sensor Networks

See Color Figure 5

This project focuses on capturing and exploiting the macroscopic patterns of behavior exhibited by the occupants of a building. By comparing readings from sensors scattered over the entirety of a building, it is possible to automatically extract information about the relationships between the sensors and the space and people that they observe. Even with very cheap sensors, such as motion detectors, and low computational overhead, it is possible to discover and exploit these important patterns to improve many building systems: elevators, heating and cooling systems, lighting, information networks, safety systems, and security systems.

Background and Objectives: The occupants of a building generate patterns as they move from place to place, stand at a corner talking, or loiter by the coffee machine. These patterns leave their mark on every object in a building. Even the lowly carpet will eventually be able to tell you quite a lot about these patterns by how it wears. A network of cheap sensors should be able to perceive these patterns and provide that context to relevant building systems. These sensor systems will only be practical if they are cheap to manufacture, install and operate. Because of their size, it is important that these networks a capable of configuring themselves and adapting to changes in the environment without special help from skilled operators. Different systems might take advantage of different kinds of contextual data. Building systems would benefit from being able to predict future room occupation levels based on current and past behavior. Elevator schedulers would benefit from more accurate predictions of passenger arrivals. Surveillance systems could leverage transitions probabilities between sensors to improve tracking or identification. The goal of this project is to explore these, and other possibilities.

Technical Discussion: By taking advantage of the normal movements of occupants in a building, we tap a large but noisy source of data. By analyzing the temporal relationships of motion events observed by 20 sensors observing a 175 square meter office area, we have been able to recover probabilistic descriptions of the relative positions of the test sensors that compare favorably to ground truth.

Collaboration: We are working with the vision group in the MERL Technology Laboratory.

Future Direction: In the near future we plan to improve on the fidelity of the geometric calibration results, and explore the issue of event prediction in the elevator arrival scenario.

Contact: Christopher R. Wren
http://www.merl.com/projects/PedestrianFlow/

Lab: MERL Research Lab
Project Type: Research

Video Mining

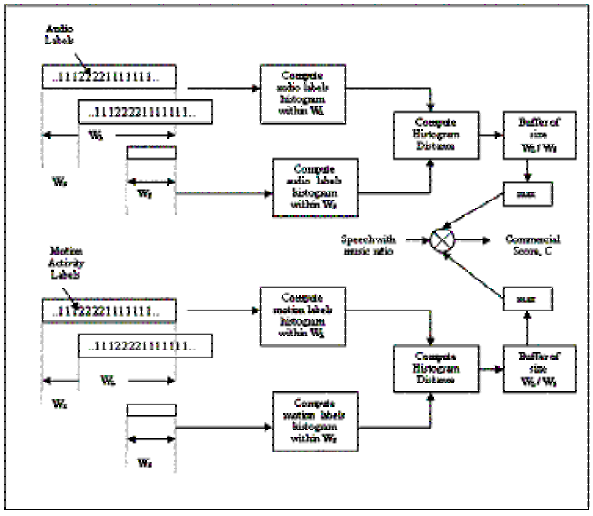


Illustration of Label Mining Framework for Commercial Detection

We consider the video mining problem in the light of existing data mining techniques. The fundamental aim of data mining is to discover patterns. In our video browsing work so far, we attempted to discover patterns in audio-visual content through principal cast detection, sports highlights detection, and location of "significant" parts of video sequences. Our approach to video mining is to think of it as "blind" or content-adaptive processing that relies as little as possible on a priori knowledge. In the figure at left, we illustrate commercial detection using discovery of unusual audio-visual patterns using adaptively constructed statistical models of usual events.

Background and Objectives: Detection of unusual video events is important for both consumer video applications such as sports highlights extraction and commercial message detection, as well as surveillance applications. Since content production styles vary widely even within genres, content-adaptive event detection is essential. Furthermore, the unusual events are both rare and diverse making application of conventional machine learning techniques difficult. In this project, we are developing techniques to adaptively detect unusual audio-visual events for both consumer applications such as video browsing systems for HDD-enabled DVD recorders, and surveillance video applications such as traffic video monitoring.

Technical Discussion: Since our target product platforms place severe constraints on computational complexity, we emphasize computational simplicity in our event discovery techniques. We thus also rely on feature extraction in the compressed domain. A typical example of our content adaptive approach is to first construct a "global" audio class histogram for an entire video program and then to compare the global histogram with histograms of smaller segments of the content so as to uncover local departures from the usual. We have obtained promising results for detection of Sports Highlights and Commercial Messages so far.

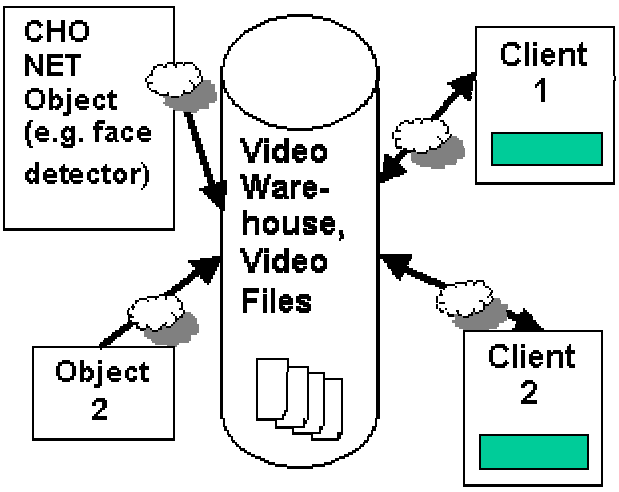
Collaboration: Johosoken, Sentansoken.

Future Direction: Our current research focus is on Consumer Video Applications. We are interested in broadening the class of detectable events, for example by targeting other genres of video such as films, drama, sitcoms etc, as well as surveillance video.

Contact: Kadir Peker, Ajay Divakaran
http://www.merl.com/projects/VideoMining/

Lab: MERL Technology Lab
Project Type: Initial Investigation

Video Warehousing and Face Classification Visualization



Applications such as surveillance or customer behavior analysis will require visualization and analysis of historical meta-data produced by various image processing algorithms (such as the face detection and classification algorithms). In order to allow for efficient data mining and visualization of such data, we proposed introducing a standard database layer responsible for managing the "video data-warehouse". As a first application of this approach, we created a framework to visually compare the results a face detection algorithm with ground truth data, as well as comparing the results of different algorithms.

Background and Objectives: A modular approach to software creation assures that applications are more flexible, robust, and are easier to maintain. By creating a database layer for applications that rely on metadata produced by video analysis (such as the face detection or face recognition algorithms), we make these modular. We often create large datasets for use in testing and improving the basic algorithms. At the same time these datasets can be used to test innovative approaches to visualization of large video and image data. Here we demonstrate this approach with an application that helps to verify and improve the effectiveness of the face detection algorithm. By allowing the user to compare the results of different versions of the algorithm, we create a regression testing framework. Comparing the algorithm with hand labeled ground truth helps determine which images are particularly challenging. This will make improving training set selection more efficient.

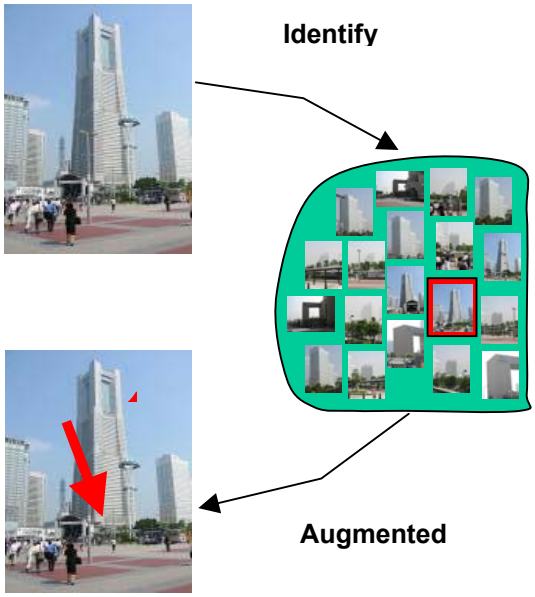
Technical Discussion: The database component of the application was constructed to interact with CHO NET objects that broadcast data over networks using a (MERL patent-pending) real time protocol. This allowed the detection component to be decoupled from other applications. Stored data can be rebroadcast imitating the original CHO NET stream, or can be accessed directly. Direct access to the underlying database (MySQL) can be made from Java via JDBC, or from C/C++ via ODBC/TCL.

Future Direction: As more data becomes stored in the database, other visualizations techniques can be applied to create applications that can be used in a high-end surveillance and other markets (e.g., "Show me all the video frames containing a person X" or "What fraction of the attendees were male?"). This technology is also being applied to the "People Prediction" project.

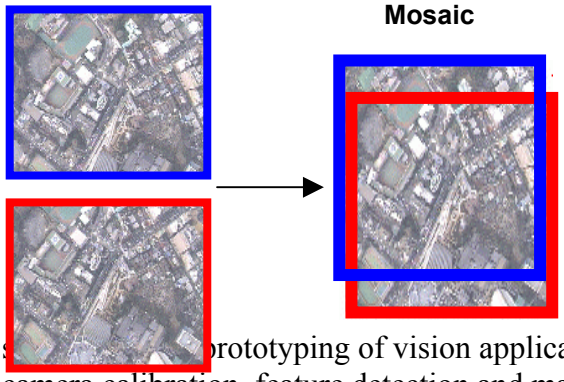
Contact: Sergei Makar-Limanov **Lab:** MERL Technology Lab
http://www.merl.com/projects/VideoWarehousing/ **Project Type:** Advanced Development

UrbanMatch and AerialMatch - Image Matching Applications

(1) UrbanMatch



(2) AerialMatch



UrbanMatch and AerialMatch are image-matching applications, built with the Diamond3D vision library. UrbanMatch is part of a proposed system that matches a live image of a building, taken with a cellphone camera, to a centralized database. GPS provides approximate cellphone location, and hence a small set of possible target images in the database, while image matching serves to identify a specific building. Finally augmentation data for the building (such as identity, prices and menus for shops and restaurants etc) is retrieved from the database and overlaid on the live image. The second application, AerialMatch, is for matching consecutive images in a sequence taken from a helicopter, to build a mosaic of the observed area on the ground. This is part of a proposed system that matches the current camera view from the helicopter to a geographical map of the area, again as a prerequisite for augmenting the live image.

Background and Objectives: As image-matching becomes more reliable, augmentation of live imagery will be increasingly common. The goal of the Diamond3D library is to provide flexible software to support multiple applications.

Technical Discussion: Diamond3D is a geometry-based computer vision library s rototyping of vision applications. The functionality of the library includes camera calibration, feature detection and matching, and 3D reconstruction. Vision components in the library utilize a Model-View-Control design, allowing them to be treated as "black-boxes" which are easily re-used between applications. The image matching is based on detection and matching of corner points. The matching process first uses gradient-similarity around corners to determine multiple putative matches for each corner, then uses geometric constraints - the fundamental matrix and the planar homography - to identify the true matches.

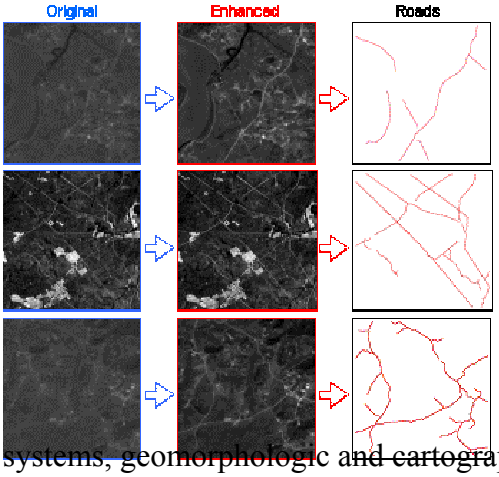
Collaboration: Dr Wakimoto, Internet Media Systems Dept, Johosoken

Future Direction: The current systems are prototypes to demonstrate the feasibility of the feature-matching approach. New work will address robustness for more varied situations

Contact: Paul Beardsley
http://www.merl.com/projects/UrbanMatch/

Lab: MERL Research Lab
Project Type: Advanced Development

Road Extraction for Satellite Imagery



Unsupervised extraction of roads eliminates the need for human operators to perform the time consuming and expensive process of mapping roads from satellite imagery. As increasing volumes of imagery become available, fully automatic methods are required to interpret the visible features such as roads, railroads, drainage, and other meaningful curvilinear structures in multi-spectral satellite imagery. There exists an even greater need for a mechanism that handles low-resolution images. The essence of detecting curvilinear elements is also related to the problem of deriving anatomical structures in medical imaging as well as locating material defects in product quality control systems, geomorphologic and cartographic applications.

Background and Objectives: Most of the proposed road detection algorithms require user assistance to mark both starting and ending points of road segments. Due to the noise sensitivity, asymmetry of the contrast at the both sides of the edges, and the difficulty of obtaining precise edge directions, edge based methods are inadequate for very low-resolution multi-spectral imagery. Our main objective is to develop automatic, robust, and computationally feasible road detection and satellite imagery analysis algorithms.

Technical Discussion: First, the input image is filtered to suppress the regions that the likelihood of existing a road pixel is low. Then, the road magnitude and orientation are computed by evaluating the responses from a quadruple orthogonal line filter set. A mapping from the line domain to the vector domain is used to determine the line strength and orientation for each point. A major problem of road extraction algorithms is disconnected road segments due to the poor visibility of the roads in the original image. Often roads are divided into several short segments, or completely missing from the image. To solve this problem, we fit Gaussian models to image points, which represent the likelihood of being road points. These models are evaluated recursively to determine the correlation between the neighboring points. The iterative process consists of finding the connected road points, fusing them with the previous image, passing them through the directional line filter set and computing new magnitudes and orientations. The road segments are updated, and the process continues until there are no further changes in the roads extracted.

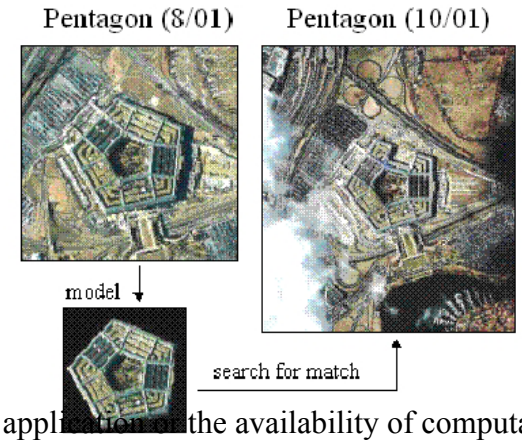
Collaboration: MERL is working closely with Johosoken on the road extraction project.

Future Direction: We plan to apply road extraction technology to automatically align airborne video camera imagery to the available road maps. In addition, we will be developing algorithms to detect other structures such as houses, overpasses, different types of vegetation, and water bodies.

Contact: Fatih Porikli
http://www.merl.com/projects/RoadExtraction/

Lab: MERL Technology Lab
Project Type: Research

Factorized Local Appearance Models



We propose a novel scheme for image-based object detection and localization by modeling the joint distribution of k-tuple salient point feature vectors which are factorized component wise after an independent component analysis (ICA). Furthermore, we use a distance-sensitive histograming technique for capturing spatial dependencies which enable us to model non-rigid objects as well as distortions caused by articulation.

Background and Objectives: For appearance based object modeling in images, the choice of method is usually a trade-off determined by the nature of the application and the availability of computational resources. Existing object representation schemes provide models either for global features or for local features and their spatial relationships. With increased complexity, the latter provides higher modeling power and accuracy. Among various local appearance and structure models, there are those that assume rigidity of appearance and viewing angle, thus adopting more explicit models while others employ stochastic models and use probabilistic distance/matching metrics. Our objective is to model the high-order dependencies of local image structure by estimating the complete joint distribution of multiple salient point feature vectors using a density factorization approach.

Technical Discussion: We construct a probabilistic appearance model with an emphasis on the representation of non-rigid and approximate local image structures. We use joint histograms on k-tuples (k salient points) to enhance the modeling power for local dependency, while reducing the complexity by histogram factorization along the feature components. Although, the gain in modeling power of joint densities can increase the computational complexity, we propose histogram factorization based on independent component analysis to reduce the dimensionality dramatically, thus reducing the computation to a level that can be easily handled by today's personal computers. For modeling local structures, we use a distance-sensitive histograming technique. A clear advantage of the proposed method is the flexibility in modeling spatial relationships. Experiments have yielded promising results on robust object localization in cluttered scenes as well as image retrieval. Most recently we have adopted parametric models using mixture of Gaussians with resulting enhancements in performance.

Collaboration: This project is joint collaboration Xiang Zhou and Thomas S. Huang (UIUC) and David Guillamet (University of Barcelona).

Future Direction: In the future, we plan to explore a more explicit way to incorporate spatial adjacency into the factorized local appearance model via graph matching. Also we are especially interested in applications of this technology to satellite image and photo-reconnaissance as shown in the example image above.

Contact: Baback Moghaddam
http://www.merl.com/projects/flam/

Lab: MERL Research Lab
Project Type: Research

Digital Video

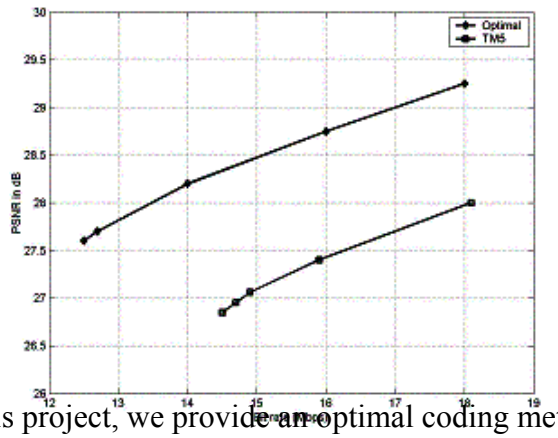
The field of Digital Video embraces techniques that span across several disciplines including traditional electrical engineering areas such as signal processing, communication and networking, as well as computer science areas, such as data analysis, content understanding, and database. Whether you realize it or not, Digital Video is an important part of our everyday lives – it is a media that supplies a primary source of information, it enables communication and trade, and it entertains us. Research in this field is moving at breakneck speeds, and MERL is clearly at the frontier setting the pace.

At MERL, we focus on technology that not only improves current video-centric systems, but also establishes a vision for next-generation systems for consumer, business, and government markets. One focus is on video compression, which targets DTV broadcast, DVD/HDD recording, and streaming applications. Within the video streaming and networking area, we investigate techniques to support video QoS and specifically conduct research on robust transmission, traffic management, and power-aware transmission. An important topic for next-generation systems is universal multimedia access. Video transcoding to specific bit-rates, formats, and spatio-temporal resolutions is a key technology to support this framework, and MERL has developed award-winning architectures and algorithms that achieve this adaptation with low complexity and high quality. We also focus on systems for multimedia storage and retrieval. Video indexing and summarization are active areas of research at MERL, and we apply this expertise to integrated systems such as DVD recorders, PVR devices, and surveillance systems. Such systems also make use of advanced multimedia content analysis, such as event detection, object tracking, content mining, and knowledge management. These are all areas that MERL has established a leading position within the research community. Active participation and contributions in MPEG standards has always been a major part of our activity and interoperability based on standards drives much of the work being done in this area.

In the pages that follow, we provide a brief overview of several recent projects that we have been working on during the past year. With regards to compression and transmission, this includes our work on high-compression encoding of video, object-based coding for efficient storage, and adaptive transmission on overlay networks. In the content analysis domain, we highlight work that has been done on video summarization. With a view towards universal multimedia access, transcoding, Pamlink-21 and our current efforts in MPEG-21 standardization are described.

Project Descriptions

MPEG Video High Compression



Recently, video high compression techniques have been applied to many consumer electronic products, such as digital video cameras and DVDs. These consumer products normally require high image quality and low bit rate. In MPEG video coding standards, the coding syntax and semantics are only specified for the decoder. The standard leaves the flexibilities to the users to develop their own MPEG encoder. High video compression can be achieved by using suitable encoder design method. In this project, we provide an optimal coding method to achieve high compression at low bit rate. The algorithm includes both picture level and macroblock level mode optimization techniques. Compared to the MPEG-2 TM5 encoder, at the same quality level, we can achieve about 25% lower bit rate.

Background and Objectives: As more and more applications, such as digital terrestrial broadcasting, digital satellite broadcasting, DVD and Internet streaming etc, become popular, the high compression techniques will become more important. Therefore, developing new high compression techniques will benefit to many MELCO's businesses. The aim of this project is to develop efficient algorithms for achieving high performance at low bit rate.

Technical Discussion: At high compression ratios, a number of visual artifacts because of the underlying block-based approach could be observed. For instance, block artifacts, ringing noise and mosquito noise etc. These artifacts severely degrade the image/video quality. In this project, we will use post-filtering techniques to suppress the artifacts caused by high compression coding and to improve the perceptual visual quality. The post-filter will be implemented in two operation modes: independent mode and embedded mode. In the independent mode, the post-filter is independent to the decoder. It takes the decoded YUV image as input, processes the YUV image and outputs a quality improved YUV image. In the embedded mode, the post-filter can be integrated into decoder to fully utilize the coding information in the video bitstreams.

Collaboration: This project is done in collaboration with the Multimedia Information Coding & Transmission Technology Department at Johosoken.

Future Direction: We will consider the following techniques: pre-filtering, motion vector quality improvement and post-filtering.

Contact: Anthony Vetro, Hao-Song Kong
http://www.merl.com/projects/MPEGVideoCompression/

Lab: MERL Technology Lab
Project Type: Research

Adaptive Video on Overlay Networks

Overlay network is regarded as a promising solution for providing high-level network services over today's Internet. This project investigates flexible rate adaptation and bit-stream partitioning mechanisms for the scalable video, and also enhanced overlay services such as multi-path transmission and end-system multicast. Our approach can improve the rate and error control capabilities in video streaming to heterogeneous user devices.

Background and Objectives: Multimedia services will proliferate in the future, particularly with the popularity of the Internet, wireless & mobile networks and cameras on handheld devices. QoS (Quality of Service) is one of the most important unsolved research issues in this area. Overlay network brings more flexibility to add higher layer functionalities for today's Internet to support QoS. This project aims to propose new bit-stream adaptation and partitioning mechanisms for scalable video and new enhanced overlay services to support adaptive video unicast and multicast.

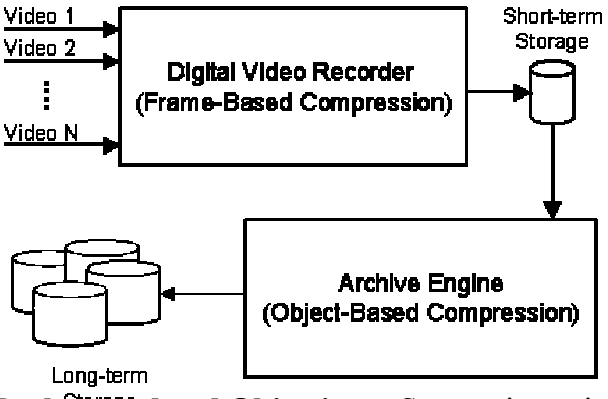
Technical Discussion: In this project we investigate how to improve rate adaptation and reliability control for scalable video streaming on overlay networks. First, we propose an enhancement layer truncation scheme for the Fine Granular Scalability (FGS) video coding. The target of our scheme is to minimize/reduce the quality variation of different parts within each one frame when only partial information of one enhancement layer can be kept after truncation according to the available network bandwidth. Secondly, we propose to partition some enhancement layers in the original FGS stream into multiple descriptions for effective transport over the multiple paths provided by overlay networks. Experimental result shows that our approach can improve the decoded visual quality.

Future Direction: We will propose QoS-enhanced multi-path and end-system multicast mechanisms for multimedia applications.

Contact: Huairong Shao, Chia Shen
http://www.merl.com/projects/AdaptiveVideo/

Lab: MERL Research Lab
Project Type: Research

MPEG-4 Object Coding for Time-Lapse Recorder



This project explores the use of using MPEG-4 object-based coding for the long-term archive of surveillance videos. The key processes needed to accomplish this are object segmentation and object-based coding. Through our experiments, and with the video sequences that we have tested, it has been found that 78% of the total storage can be saved, and that the quality of the reconstructed image depends heavily on the accuracy of the object segmentation.

Background and Objectives: Storage is a primary concern for time-lapse recorder systems used in surveillance applications. Typically, such systems allow input from many camera sources and the system is required to store several months of video data for each source. An archiving engine is useful to transfer the video data from short-term to long-term storage in which some degradation to the originally stored video is tolerable.

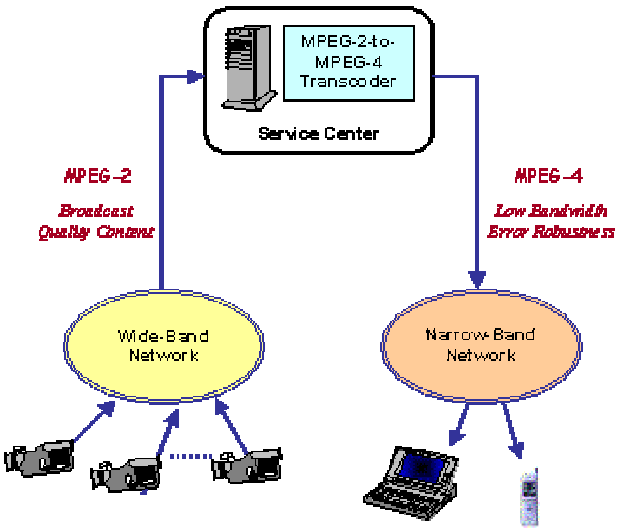
Technical Discussion: The proposed system relies on two important processes: segmentation and object-based coding. In this system, it is assumed that certain inaccuracies in the reconstructed scene can be tolerated, however subjective quality and semantics of the scene must be strictly maintained. Preliminary results indicate an overall bit savings of 78%.

Collaboration: Information Processing Department at Sentansoken.

Future Direction: The algorithms that have been developed need to be tested further on more generic scenes. Since the segmentation accuracy has the most impact on the reconstructed scene, we also consider improving the accuracy of the segmentation algorithm.

Contact: Anthony Vetro **Lab:** MERL Technology Lab
http://www.merl.com/projects/MPEG4ObjectCoding/ **Project Type:** Initial Investigation

MPEG Transcoding for Surveillance



In general, the purpose of a transcoder is to convert compressed content, such as an MPEG bitstream, into a format that satisfies transport over dynamic networks, as well as playback and recording of content with various devices. In this project, we have developed software for real-time MPEG-2 to MPEG-4 transcoding with a reduced bit-rate and spatio-temporal resolution. For surveillance applications, this enables MPEG-2 broadcast quality content to be received by a central service center and be distributed to remote clients over narrow-band networks. MPEG-4 is the preferred format such networks due to its coding efficiency and error robustness.

Background and Objectives: Recent advances in signal processing combined with an increase in network capacity are paving the way for users to enjoy services wherever they go and on a host of multimedia capable devices. Such devices include laptops and mobile handheld devices. Each of these terminals may support a variety of different formats. Furthermore, the access networks are often characterized by different bandwidth constraints, and the terminals themselves vary in display capabilities, processing power and memory capacity. Therefore, it is required to convert and deliver the content according to the network and terminal characteristics.

Technical Discussion: Our MPEG-2 to MPEG-4 transcoding software is able to achieve reduced bit-rates, spatial resolutions, and temporal resolutions. The transcoding is done in an efficient way such that multiple bitstreams can be transcoded with general-purpose processors. The brute force approach decodes the original bitstream to the spatial-domain, performs some intermediate processing, and then finally re-encodes to a new bitstream. Our proposed architectures simplify this process, while still maintaining the picture quality.

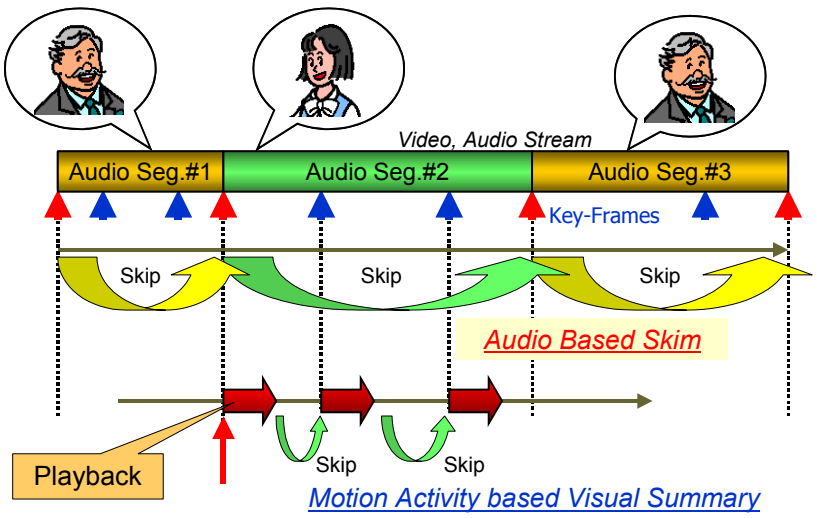
Collaboration: Technical aspects of this project are presently done in collaboration with the Multimedia Information Coding & Transmission Technology Department at Johosoken. Past collaborators include the Image Information Processing Department at Sentansoken and Princeton University.

Future Direction: We continue to add features, reduce the complexity improve robustness to errors and improve output quality. Also, we seek new applications that this technology, such as set-top box and digital still camera.

Contact: Anthony Vetro
http://www.merl.com/projects/MPEG-transcoding/

Lab: MERL Technology Lab
Project Type: Advanced Development

Video Summarization for the DVD Recorder



The Personal Video Recorder such as the Recordable-DVD Recorder and/or Hard Disk Recorder has become popular as a large volume storage device for video/audio content and a browsing function that would quickly provide a desired scene to the user is required as an essential part of such a large capacity system. We propose an intra-program content browsing system using not only a combination of motion based video summarization and topic-related metadata in the incoming video stream but also an audio-assisted video browsing feature that enables completely automatic topic-based browsing. As illustrated in the figure on the left, we use audio analysis based on generalized sound recognition to locate semantic segment boundaries and then summarize each segment using our motion activity based technique.

Background and Objectives: As more and more audio-visual content becomes available in digital form in various places around the world, the ability to locate desired content will become more and more important. Already text based search engines help retrieve textual data from the World Wide Web, but no equivalent identifying information exists for A/V content. The proposed MPEG-7 standard will standardize a multimedia content description interface that will enable efficient searching and browsing of worldwide multimedia content. In this project we emphasize the Personal Video Recorder application, that provides the user with the content he wants when he wants it by storing a large volume of content recorded from broadcast and then providing effective navigation of the stored content using summarization and indexing.

Technical Discussion: The system relies on extraction of compact descriptors in the compressed audio and video domain, which makes both the content preparation and the content access fast. It primarily relies on the MPEG-7 motion activity descriptor, and also makes use of simple color histograms and MPEG-7 generalized sound recognition. We have a unique motion activity based approach to video summarization.

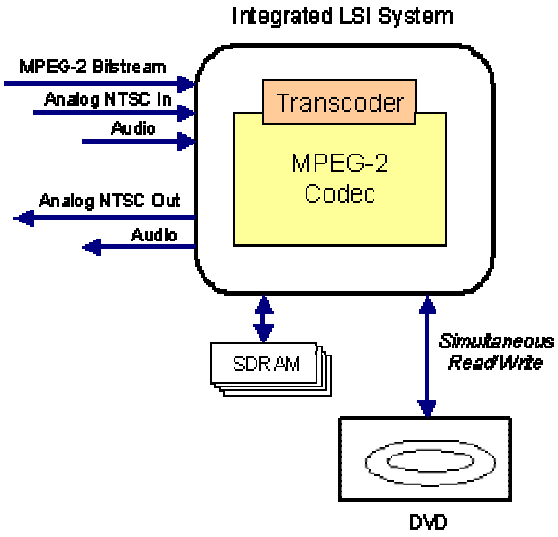
Collaboration: Johosoken, Sentansoken.

Future Direction: Our current research focus is on Content Summarization. We are still

Contact: Ajay Divakaran
http://www.merl.com/projects/VideoSummarization/

Lab: MERL Technology Lab
Project Type: Advanced Development

MPEG Transcoding for DVD Recording



In general, the purpose of a transcoder is to convert compressed content, such as an MPEG bitstream, into a format that satisfies transport over dynamic networks, as well as playback and recording of content with various devices. In this project, we are developing a software model for an MPEG-2 transcoder that adapts the bit-rate and spatial resolution. The transcoding logic should be integrated with an MPEG-2 Codec to make an efficient LSI design. For DVD recording, the MPEG-2 broadcast contents may be Main Profile @ Main Level (MP@ML, Standard-Definition Television) or Main Profile @ High Level (MP@HL, High-Definition Television), but the recorded DVD contents must be MP@ML. Also, the bits recorded to the disk have some limitations as well, and are usually variable bit-rate.

Background and Objectives: Digital television is now being broadcast around the world in the MPEG-2 format. Depending on the county, either MP@ML or MP@HL can be received. In the market today, you can find broadcast receivers, which are now being integrated into the newer digital television sets. Also, you can find a wide array of DVD players. The aim of this project is to be able to record the television broadcast onto the DVD medium in an efficient way. One important feature of the DVD recorder is simultaneous read/write.

Technical Discussion: An optimal mode decision and quantizer selection algorithm is implemented in the system. The transcoding achieves high performance. Due to the complexity of the optimal method, the current system is operated offline. We use the system as a benchmark for quality measurement comparison. A simplified algorithm is also implemented. It greatly reduces the complexity of the optimal method and still achieves higher performances than the conventional method, such as TM5.

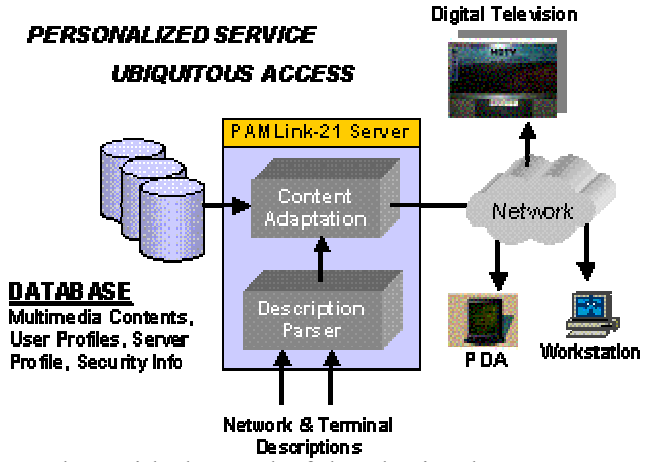
Collaboration: The aspects of this system were discussed with Sentansoken and Johosoken.

Future Direction: We continue to make effort to improve the video quality and reduce the complexity of the simplified system. The mode decision and quantizer selection algorithms will be further investigated and simplified.

Contact: Anthony Vetro, Hao-Song Kong
http://www.merl.com/projects/transcoding_dvd/

Lab: MERL Technology Lab
Project Type: Advanced Development

PAMLink-21



The primary aim of the PAMLink project was to develop an end-to-end connectivity solution for Internet Appliances and Internet services. With this in mind, PAMLink services are characterized by a highly personalized experience for the individual user, where the services and information are aggregated to meet individual needs at any given moment. This is very much in-line with the vision of the emerging MPEG-21 standard, which is to enable transparent access to multimedia content. PAMLink-21 merges these two initiatives together with the goal of developing key components for an MPEG-21 system and integrating them with a rich multimedia-enabled infrastructure.

Background and Objectives: This project concentrates on merging technology that has been developed to gain personal access from mobile devices with the emerging MPEG-21 standard.

Technical Discussion: The current thrust of this project is to use the PAMLink infrastructure and the MPEG-21 standard to stream multimedia content to various user's terminals each having different capabilities. Various descriptors, which are in the process of being standardized by MPEG-21, are used to describe the producer, the consumer, the network delivery system, as well as the expressions of various rights a user may have to the content. The server, whose responsibility is to manage the services and deliver content, is comprised of a content adaptation engine and various parsers to interpret the MPEG-21 descriptions. The client has an integrated HTML/XML browser and MPEG-4 Player, which has been implemented on a WinCE platform. At this current stage, the content delivery network is based on RTP/UDP/IP.

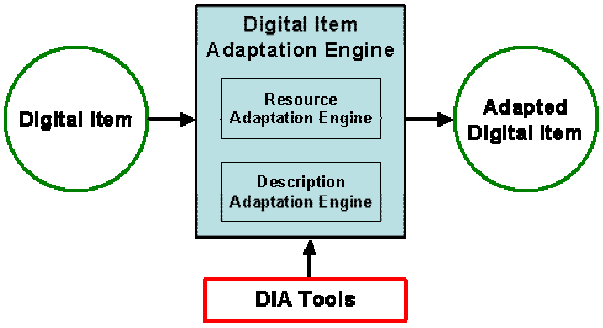
Collaboration: Certain part of this project are done in collaboration with the Multimedia Information Coding & Transmission Technology Department at Johosoken, such as the MPEG players, transcoding, video delivery, and MPEG-21 components.

Future Direction: This project ended in FY03, and will not continue in FY04.

Contact: Johnas Cukier
http://www.merl.com/projects/PAMLink21/

Lab: MERL Technology Lab
Project Type: Research

MPEG-21 Standards Activity



One of the goals for MPEG-21 is to enable transparent access to multimedia content across a wide range of networks and devices. The notion of a Digital Item has been introduced, which is the fundamental unit of transaction in the MPEG-21 multimedia framework. Simply stated, Digital Items contain a rich set of content with associated descriptions, and they are configurable to various conditions. Our activities in MPEG-21 are focused on the development of Part 7, Digital Item Adaptation.

Background and Objectives: Compression, transport, and multimedia description are examples of individual technologies that are already well defined. However, the lack of interoperable solutions across these spaces is stalling the deployment of advanced multimedia packaging and distribution applications. This problem has motivated the MPEG committee to start its work on defining normative tools to enable multimedia applications of the 21st century. Officially launched in June 2000, MPEG-21 is referred to as the Multimedia Framework. This framework applies across the entire consumption chain from content creators and rights holders to service providers and consumers.

Technical Discussion: To enable transparent access to content, it is essential to have available not only the description of the content but also a description of its format and of the usage environment in order that content adaptation may be performed to provide the User the best content experience for the content requested with the conditions available. While the content description problem has been addressed by MPEG-7, the description of content format and usage environments has not been addressed and it is now the target of MPEG-21 Digital Item Adaptation, which is Part 7 of MPEG-21 and will be finalized in December 2003. Among the tools targeted for standardization include usage environments descriptions, which include descriptions of the terminal, network, user and natural environment.

Collaboration: This project is done in collaboration with the Multimedia Information Coding & Transmission Technology Department at Johosoken.

Future Direction: Continue involvement in the development of MPEG-21, with particular emphasis on Part 7: Digital Item Adaptation.

Contact: Anthony Vetro
http://www.merl.com/projects/mpeg21/

Lab: MERL Technology Lab
Project Type: Research

Graphics Applications

Computer graphics research in general and at MERL has been undergoing some profound changes. During its formative years, computer graphics has focused largely on developing algorithms and systems for performing efficient simulations to produce images and animations. At present, this simulation framework for computer graphics is very mature. Now the field has been opening up to include high-quality imaging sensors, measurement devices, projectors, and other means of combining the real-world with synthetic, graphical models.

At MERL we explore new approaches to computer graphics that attempt to bridge the dichotomy between the classical and this data-driven modeling. These approaches differ from the simulation-based computational model, and instead depend more on the tools of interpolation and signal processing. To emphasize this new focus we changed the name of this research area to "graphics applications." MERL is in the unique position to make a huge impact in graphics applications because of our world-class researchers in the fields of computer vision, computer graphics, and machine learning.

This past year, graphics application research has produced some of the strongest technology transfer opportunities in the history of MERL. The projector related research by Raskar, van Baar, and Beardsley has already made its way into products and has won the Johosoken Management Award in 2003. The 3D reconstruction project of Ziegler and Pfister is in collaboration with Dr. Shiotani's Strategic Advanced Project entitled "Vision/Image based Modeling Subsystem for Facility Engineering Applications" (Sentansoken). The data-driven reconstruction and reflectance research by Pfister, Moghaddam, and Brand is being applied successfully to 3D face recognition, the next big step in computer human observation. And the ADF research by Frisken and Perry is bound to make a huge impact in the market for high-quality font representations for TVs, LCD displays, and portable devices, including cell phones and car navigation systems. MERL is very excited about the commercial potential of all its graphics application research, and we hope to have some positive impact on MELCO business in the near future.

Project Descriptions

Video Surveillance with NPR Image Fusion

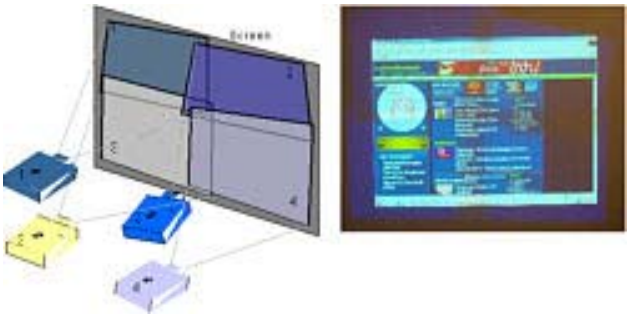
We have developed a class of techniques to enhance context in images and videos. The basic idea is to increase the information density in a set of low quality images by exploiting the context from a higher quality image captured under different conditions from the same view. For example, a nighttime surveillance video is enriched with information available in daytime images or video in the infrared frequency range is augmented with video in visible light spectrum.

Background and Objectives: An image is traditionally enhanced using information included within the same image. We exploit the idea that, for fixed cameras, the image of a scene can be captured under different conditions over time, e.g. illumination, wavelength, atmospheric conditions and containing static or dynamic objects. We propose a new image fusion approach to combine images with sufficiently different appearance into a seamless rendering. The method maintains fidelity of important features and robustly incorporates background contexts avoiding traditional problems such as aliasing, ghosting and haloing.

Technical Discussion: Our method first encodes the importance based on local variance in input images or videos. Then, instead of a convex combination of pixel intensities, we combine the intensity gradients scaled by the importance. The image reconstructed from the gradients achieves a smooth blend and at the same time preserves the detail in the input images. We have obtained results on indoor as well as outdoor scenes.

Contact: Ramesh Raskar
Lab: MERL Research Lab
http://www.merl.com/projects/NPRfusion/
Project Type: Research

Low Cost Projector Mosaic



We have developed a set of techniques to create large format displays using casually installed overlapping projectors. The image alignment and intensity blending is completely automatic. The time required to calibrate the system is less than 20 seconds allowing a very easy to use and flexible setup. Current multi-projector systems cost several times the price of the projectors due to the expensive infrastructure and maintenance. Our techniques focus on software methods to reduce the cost by eliminating the need of rigid support structures and manual alignment.

Background and Objectives: A photo-mosaic is a collection of images registered together to form one large image. Can we create a similar projector mosaic on a display surface by seamlessly merging output of overlapping projectors? For photo-mosaic, the images are captured by casually orienting and positioning the camera. Can we similarly create a large display using casually aligned projectors? Currently, large displays are generated by tiling together a two dimensional array of projectors by precisely aligning them. Due to such design constraints, the installation and operation of such systems is extremely expensive. If we allow casual approximate installation, the cost of multi-projector systems can be greatly reduced. Further, if the registration and blending of overlapping images is performed completely automatically, such systems can become very easy to use allowing wide spread use in homes, offices and shops.

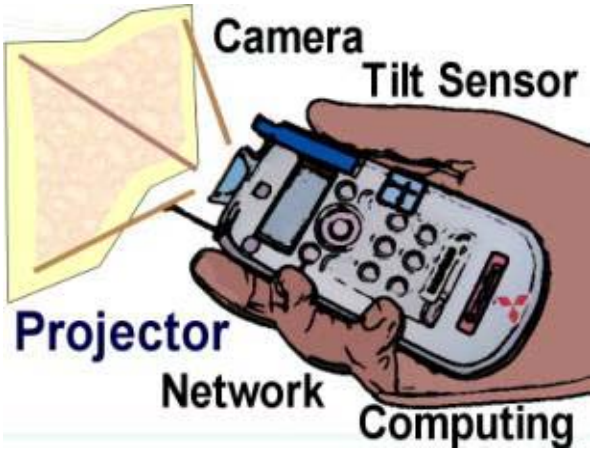
Technical Discussion: We use a camera in the loop to automatically compute the relative projector pose. Our algorithms solve two fundamental problems involved in multi-projector displays. First, we achieve sub-pixel alignment between overlapping projectors by exploiting the homography induced due to the display plane. Second, we use a fast technique to find intensity feathering weights to cross-fade the projected images in the overlapping region. We use 3D graphics hardware to achieve real-time pre-warping of the images so that they generate a single seamless image when projected. The techniques can be used for rear-projector or front-projection display.

Future Direction: We are investigating techniques to achieve color equivalence between overlapping projectors. We also are planning to reduce the calibration time to below 10 seconds.

Contact: Ramesh Raskar, Jeroen van Baar
http://www.merl.com/projects/ProjectorMosaic/

Lab: MERL Research Lab
Project Type: Advanced Development

iLamps: Intelligent, Locale-aware, Mobile Projectors



We have developed a concept of a smart, enhanced projector, iLamps. It can determine and respond to the geometry of the display surface, and can be used in an ad-hoc cluster to create a self-configuring display. In addition, we are exploring various algorithms for new display paradigms, their applications and solving underlying problems.

A projector is decoupled from the display and hence the size of the projector can be much smaller than the size of the image it produces. Hence, projectors look like natural fit with handheld devices such as cell phones and PDAs. Cell phones provide access to the large amounts of wireless data, which surround us, but their size dictates a small display area.

Background and Objectives: Projectors are currently undergoing a transformation as they evolve from static output devices to portable, environment-aware, communicating systems. An Information display is such a prevailing part of everyday life that new and more flexible ways to present data are likely to have significant impact. A hand-held networked projector (e.g. in a cell phone or PDA) can maintain compactness while still providing a reasonably sized display. A hand-held cell phone-projector becomes a portable and easily deployed information portal.

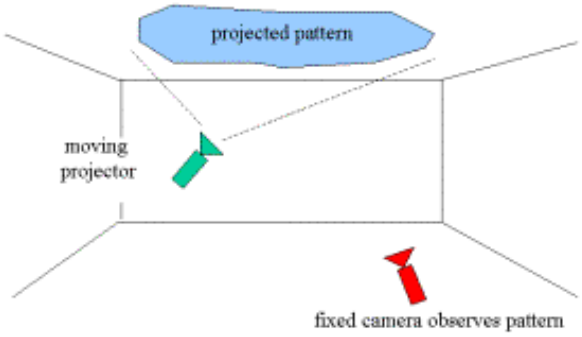
Technical Discussion: Our basic unit is a projector with attached camera, tilt-sensor and network component in addition to computing, storage and interface. Single units can recover 3D information about the surrounding environment, including the world vertical, allowing projection appropriate to display surface. Multiple, possibly heterogeneous, units are deployed in clusters, in which case the systems not only sense their external environment but also the cluster configuration, allowing self-configuring seamless large-area displays without the need additional sensors in the environment.

Collaboration: Mr Kameyama and Mr Ashizaki, MultiMedia Group, Johosoken; Dr Shiotani, Power & Public Utility System Dept, Sentansoken, Sanden.

Contact: Ramesh Raskar
http://www.merl.com/projects/iLamps/

Lab: MERL Research Lab
Project Type: Research

A Projector as a Novel Type of Motion Sensor



Motion sensors are useful for a large variety of applications including robot navigation, virtual reality, movie special effects, and 3D reconstruction. A camera is a cheap form of 6 DOF motion sensors, but automatic recovery of camera motion from images often fails in arbitrary environments. More robustness is obtained from installations using electromagnetic or ultrasound sensors but these are expensive, lack portability, and have a fixed-size workspace. Systems that involve instrumenting the environment with LEDs or distinctive visual markers also lack portability, and require significant calibration. This project describes a novel type of motion sensor that is cheap, portable, with relatively straightforward calibration, and the workspace can be arbitrarily large. It requires the presence of a planar surface for projection, and is most suitable for indoor use where the ceiling or sidewalls provide the required surfaces.

Background and Objectives: A cheap reliable motion sensor has many applications. Current work involves a motion sensor attached to a camera, to support 3D reconstruction.

Technical Discussion: A moving projector projects a (moving) pattern onto a planar surface, and this is observed by a fixed camera placed at an arbitrary position in the environment. The observed pattern allows recovery of the position and orientation of the projector - the projector can thus be attached to any object of interest, to serve as a motion sensor. The minimum requirement for the projected pattern is that it has three distinct points, so the projection device can be cheap, and total cost with camera could be sub-$100. For an extended workspace, multiple fixed cameras are placed around the environment.

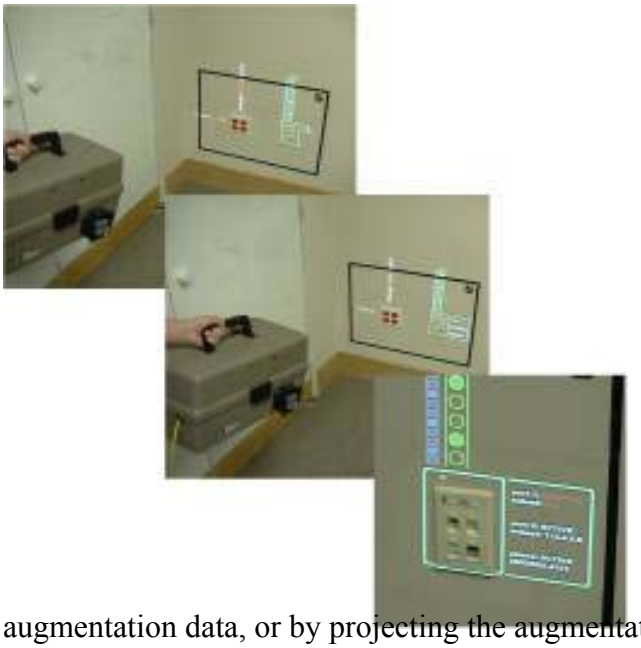
Collaboration: Dr. Shiotani, Power & Public Utility System Dept, Sentansoken, Sanden.

Future Direction: We are interested in real-world applications, such as 3D measurement for factory installations.

Contact: Paul Beardsley, Ramesh Raskar
http://www.merl.com/projects/ProjectorMotionSensor/

Lab: MERL Research Lab
Project Type: Research

Hand-Held Projectors for Augmented Reality



This work is about projecting augmentation data onto real world objects, using a portable hand-held projector. Augmentation is used (a) to provide data about the state of an object; (b) to provide training information about how to control a machine or device; (c) to guide a user to the location of a requested object e.g. by highlighting one bin in an array of storage bins; (d) to reveal the presence of electronic data (such as virtual post-it notes) in the environment.

Background and Objectives: Augmented reality is a well-established concept. It can be realized by displaying an image plus augmentation on a normal display, or by using see-thru eye-worn displays to supplement the user's direct view with augmentation data, or by projecting the augmentation directly onto the object of interest. Current systems for projecting the augmentation tend to be static, and often require significant calibration before use. This work demonstrates a much more flexible approach - portable projectors which can be used in hand-held mode, or can be easily deployed on surfaces around an object of interest.

Technical Discussion: The projector is augmented with a camera to support object recognition, and to allow computation of the pose of the projector. Common to previous approaches, we do object recognition by means of fiducials attached to the object of interest. Our fiducials are 'piecodes', which allow thousands of distinct color-codings. As well as providing identity, these fiducials are used to compute camera pose and hence projector pose. A hand-held projector can use various aspects of its context when projecting content onto a recognized object. We use proximity to the object to determine level-of-detail for the content. Other examples of context for content control will be gestural motion, and the presence of other devices during cooperative projection. We are currently investigating how to do mouse-style interactions with the projected augmentation data, by means of a projected cursor.

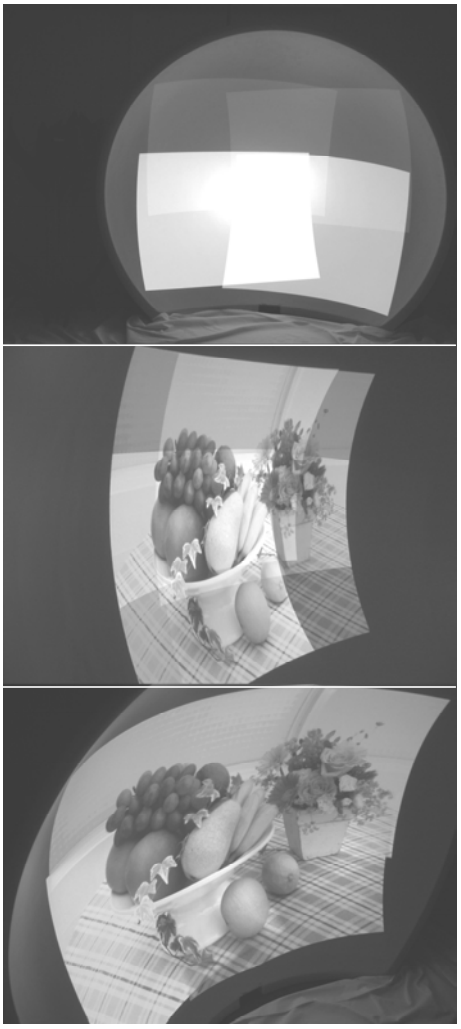
Collaboration: Dr Shiotani, Power & Public Utility System Dept, Sentansoken, Sanden.

Future Direction: Develop a real training application in collaboration with Sentansoken; multi-unit interaction.

Contact: Ramesh Raskar, Paul Beardsley
http://www.merl.com/projects/HandHeldProjectorAR/

Lab: MERL Research Lab
Project Type: Research

Multi-Projector Imagery on Curved Screens



We describe a new technique to display seamless images using overlapping projectors on curved quadric surfaces such as spherical or cylindrical shape. Current techniques for automatically registered seamless displays have focused mainly on planar displays. On the other hand, techniques for curved screens currently involve cumbersome manual alignment to make the installation conform to the intended design. We show a seamless real-time display system.

Background and Objectives: Large seamless displays using overlapping projectors is an emerging technology for constructing high-resolution semi-immersive visualization environments capable of presenting high-resolution images from scientific simulation, entertainment and instruction. General techniques that can handle setups where projectors have been casually installed and exploit geometric relationship between projectors and display surface eliminate cumbersome manual alignment and reduce maintenance costs.

Technical Discussion: We define a new quadric image transfer function and show how it can be used to achieve sub-pixel registration while interactively displaying two or three-dimensional datasets. Accurate estimation of geometric relationship between overlapping projectors is the key for achieving seamless displays. They influence the rendering algorithms and also determine soft edge blending efforts. Our technical contributions are as follows.

-- Simplification of quadric transfer
-- Calibration methods
-- Automatic sweet-spot detection and area of display
-- Software blending scheme using parametric approach
-- Fast rendering strategy to exploit hardware

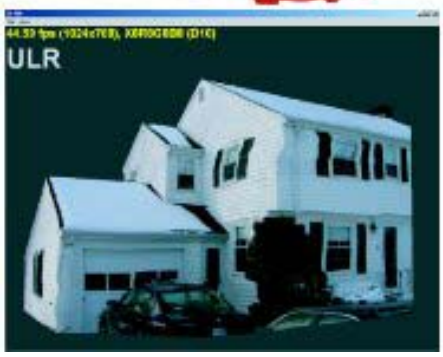
Collaboration: Mr Kameyama, Mr Ashizaki, MultiMedia Lab, Johosoken, Mr Ogata, MPC.

Future Direction: We would like to extend our work for surfaces with small deviations from quadric surfaces, and eventually general curved surfaces.

See Color Figure 10

Contact: Ramesh Raskar, Jeroen van Baar **Lab:** MERL Research Lab
http://www.merl.com/projects/CurvedScreenProjection/ **Project Type:** Advanced Development

3D Reconstruction from Photographs



See Color Figure 11

We created a novel algorithm for reconstructing 3D scenes from a set of images. The user defines a set of polygonal regions with corresponding labels in each image using familiar 2D photo-editing tools. Our reconstruction algorithm computes the 3D model with maximum volume that is consistent with the set of regions in the input images. The algorithm is fast, uses only 2D intersection operations, and directly computes a polygonal model. We also implemented a user-assisted system for 3D scene reconstruction.

Background and Objectives: Creating photorealistic 3D models of a scene from multiple photographs is a fundamental problem in computer vision and in image based modeling. The emphasis for most computer vision algorithms is on automatic reconstruction of the scene with little or no user interaction, making a priori assumptions about the geometry or reflectance. On the other hand many image-based modeling systems require that the user leads the construction of the 3D scene by specifying object characteristics in 3D. Our system requires only 2D interactions from the user and creates a 3D model, using this information, automatically.

Technical Discussion: The input to our system is a set of images of a scene from different viewpoints. The user selects corresponding regions in all images assisted by tools, e.g., intelligent scissors or intelligent paint. We are formulating the reconstruction problem as a computational geometry problem, which allows our algorithm to construct a 3D model of a scene with arbitrary number of objects. The reconstruction works well for all kinds of materials and is not dependent on surface properties. The 3D model is then rendered in real-time with Unstructured Lumigraph Rendering.

Collaboration: This project is a collaboration between MERL, MIT, Dr. Kameyama (Johosoken), and Dr. Shiotani (Sentansoken).

Future Direction: We plan to continually improve the tools and user interface for image segmentation and region matching.

Contact: Hanspeter Pfister
http://www.merl.com/projects/3DReconstruction/

Lab: MERL Research Lab
Project Type: Initial Investigation

Data-Driven Reflectance Model

We have developed a generative model for isotropic bi-directional reflectance distribution functions (BRDFs) based on measured reflectance data. We believe that our model is much easier to use and control than previous analytic models in which the meaning of parameters is often non-intuitive. Our model is also more realistic, which should improve applications in computer vision and computer graphics.

Background and Objectives: A fundamental problem in computer graphics and computer vision is modeling how light is reflected from surfaces. A class of functions called Bidirectional Reflectance Distribution Functions (BRDFs) characterizes the process where light transport occurs at an idealized surface point. Traditionally, physically or empirically inspired analytic reflection models are used. These BRDFmodels are only approximations of reflectance of real materials. Furthermore, most analytic reflection models are limited to describing only particular subclasses of materials - a given model can represent only the phenomena for which it is designed. Our data-driven reflectance model is based on real data and the produced BRDFs look very realistic. Furthermore, we provide a set of intuitive parameters that allow users to change the properties of the output BRDF. Our model has many applications in computer graphics and computer vision.

Technical Discussion: Instead of using analytical reflectance models, we represent each BRDF as a dense set of measurements. This allows us to interpolate and extrapolate in the space of acquired BRDFs to create new BRDFs. We treat each acquired BRDF as a single high-dimensional vector taken from a space of all possible BRDFs. We apply both linear (subspace) and non-linear (manifold) dimensionality reduction tools in an effort to discover a lower-dimensional representation that characterizes our measurements. We let users define perceptually meaningful parameterization directions to navigate in the reduced-dimension BRDF space. On the low-dimensional manifold, movement along these directions produces novel but valid BRDFs.

Collaboration: This project is a collaboration between MERL and MIT.

Future Direction: Analysis of more complex reflectance functions (4D BRDF, BTF, BSSRDF), specialized reflectance functions, real-time (hardware) rendering, inverse rendering, mapping between digital material mixtures and real-mixtures.

Contact: Hanspeter Pfister, Matthew Brand
http://www.merl.com/projects/brdf/

Lab: MERL Research Lab
Project Type: Research

Adaptively Sampled Distance Fields (ADFs)

ADFs are a new digital representation of shape with several advantages over existing approaches. They provide efficient and accurate representation of both smoothly curved surfaces and surfaces with fine detail. ADFs have the potential to impact many diverse industries including CAD/CAM (simulation, path planning, and verification for milling precision parts), Entertainment (building models for games and movies), Fonts (high-quality display of letterforms for PDAs), Visualization (volumetric visualization of molecular structure), 3D Scanning (3D models from image or range data), 3D Printing (rapid prototyping), and Color Management (projectors, PDAs, monitors, and printers).

Background and Objectives: Our objectives include fundamental research, incorporation of this research into a product-worthy C library ready for commercialization, development of a comprehensive patent portfolio, and collaboration with key industrial players to refine and expand the vision for ADFs.

Technical Discussion: A distance field is a scalar field that specifies a distance to a shape, where the distance may be signed to distinguish between the inside and outside of the shape. ADFs consist of adaptively sampled distance values, organized in a spatial data structure, with a method for reconstructing the distance field from the sampled distance values. This approach permits the accurate and compact representation of fine detail and smooth surfaces, together with efficient processing. ADFs allow: the representation of more than the surface (interiors and exteriors); the compact representation of sharp features and organic shapes; smooth surface reconstruction; trivial inside/outside and proximity testing; fast and simple CSG operations; fast geometric queries such as closest point; and efficient computation of surface offsetting, blending and filleting, collision detection, morphing, and rough cutting.

Collaboration: PTC (ProEngineer CAD/CAM software), Think3 (PTC competitor), Industrial Light and Magic (movies: Star Wars, Jurassic Park, ...), WetaFX (Lord of the Rings movie trilogy), NOVA documentary series, and MELCOÆs Factory Automation group.

Future Direction: Fundamental research and technology transfer.

Contact: Sarah Frisken, Ron Perry
http://www.merl.com/projects/adfs/

Lab: MERL Research Lab
Project Type: Research

Interface Devices

The goal of research and development in the field of interface devices is to provide new means of interaction with pervasive computing systems. New interface devices provide improvements - smaller, less expensive, more unobtrusive, more "natural." With technologies being developed at MERL, we are taking steps toward Mark Weiser's famous "ubiquitous computing" vision. However, MERL's drive is to create enabling technology that provides solutions to real-world problems.

Our work on interface devices is continuing, with major efforts in the capacitive and LED based sensing domains. DiamondTouch (a multi-user, multi-touch projective sensing display) and Digital Merchandising (retail displays that react to the actions of the customer) use very simple, inexpensive capacitance-based sensing systems. "LED Comm" and LED-based sensors use an LED's parasitic capacitance to give a new, very inexpensive way to communicate between devices and to sense very small amounts of certain materials or pollutants in an environment. MERL is also developing technology that adds significant new functionality to existing products (projectors, mobile phones, etc.) without adding significant cost and complexity.

Project Descriptions

DiamondTouch Hardware

DiamondTouch is a simultaneous, multi-user, touch sensitive input device developed at MERL. Not only can it detect multiple, simultaneous touch events, but it can also identify which user is touching where. This unique ability has made DiamondTouch a very useful device in the human-computer interface research community. Work on the DiamondTouch hardware has now produced two near production quality prototypes, the DT88 and the DT107. Both are meant to be used with any of Mitsubishi's line of video or computer data projectors. This document covers the hardware aspects of this effort.

Background and Objectives: DiamondTouch was first created in 2001 as an experimental multiuser interface device. We have been recognized for this technology, and are creating commercially viable products by seeding select university groups with prototype units.

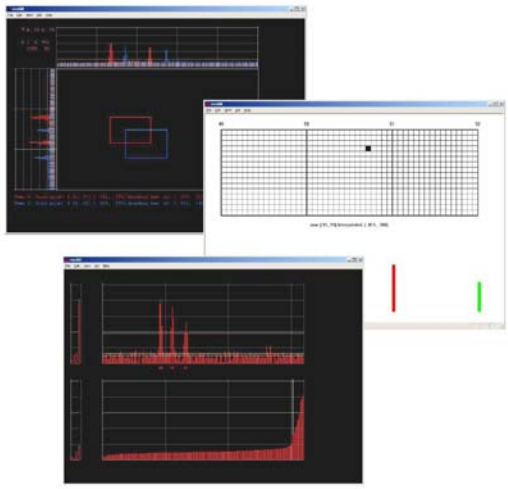
Technical Discussion: DiamondTouch uses an array of rows and columns of antennas embedded in the touch surface. Each antenna transmits a unique signal. Each user has their own receiver, usually connected to the user capacitively through a conductive part of the user's chair. When a user touches the touch surface, the antennas near the user's touch point couple an extremely small amount of signal into the user's body, the user's receiver picks up the signal and determines which of the antennas the user was touching. The current pre-production DiamondTouch units are available in two sizes - the DT88 (79cm diagonal), and the DT107 (107 cm diagonal), large enough for a full-size A1 drawing. Both units have a 4:3 aspect ratio, with full support for up to 4 users. The antenna pitch on both units is 5mm horizontally and vertically, and with signal interpolation we can achieve 0.1mm accuracy. The cases are cast urethane, and are gasketed to prevent minor spills from damaging the electronics. We have completed the first run of high-quality DT88 units, and are now starting the first run of high-quality DT107 units for demonstration and research use.

Collaboration: The DiamondTouch hardware project is a collaboration of MERL MRL and MTL; we are working with Johosoken, Kamaden and IDken to make best use of the hardware for securing future contracts for MELCO. Additionally, we are collaborating with many universities in researching further applications.

Future Direction: In research, we are working toward transparent or translucent DiamondTouch antennas, and a wireless DT system. In development, we are working toward productizing the DiamondTouch technology for multiuser interactions, using it as a selling point for MELCO systems and sales, and selling DiamondTouch units to third-party system builders.

Contact: William Yerazunis, Darren Leigh, David Wong **Lab:** MERL Research Lab
http://www.merl.com/projects/DiamondTouch/ **Project Type:** Advanced Development

DiamondTouch Software Development Kit (SDK)



DiamondTouch is a touch input device that distinguishes between multiple simultaneous users and tracks multiple simultaneous touch points for each user. (See DiamondTouch Hardware for more details on the hardware.) The DiamondTouch Software Development Kit (SDK) provides support for the development of Microsoft Windows and Linux applications that utilize DiamondTouch's capabilities to implement computer-supported collaboration and rich input modalities (such as gestures). When projected upon, the touch surface facilitates direct manipulation of user interface elements and provides a shared focus of attention for collaborating users. Possible applications include disaster-control command posts, power plant control rooms, business or technical meetings, and a variety of casual applications (e.g., musical instrument control, home coffee table, games etc). (See DiamondTouch Applications for more details on applications.)

Background and Objectives: The SDK implements key features of the technology, provides a platform for further exploration of its possibilities and applications, and is the vehicle whereby we support our collaborators (internal and external).

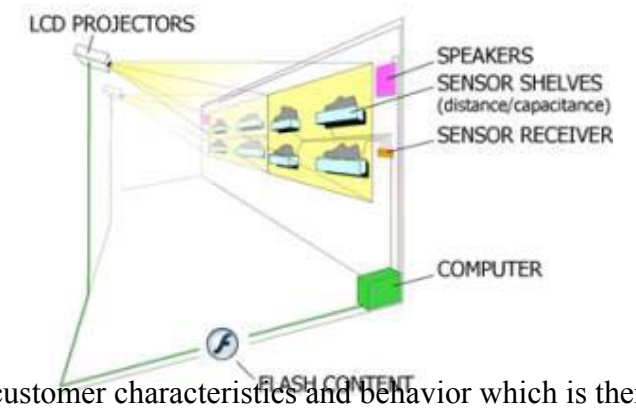
Technical Discussion: The SDK provides two libraries to support DT application development. The DiamondTouch hardware periodically produces frames of data indicating the proximity of the user's finger(s) to each antenna. The lower level library (dtio) reads these data frames from the device and affords access to the raw data and to various abstractions and interpretations of that data, such as the location of the touch point and the bounding box of the area touched. A weighted interpolation algorithm increases the effective resolution to subpixel resolution. Adaptive touch thresholding and other techniques improve robustness in the face of RF interference. The higher level library (dtlib) provides a more programmer-friendly API, providing access to more semantically oriented events. The SDK consists of dtio and dtlib (ANSI C), jdt (a Java interface layer), merldt (a Windows application providing mouse emulation, projector calibration, and diagnostic displays), and a simple multi-user application example.

Collaboration: DiamondTouch is a joint project of MERL's Technology and Research Laboratories. MERL is collaborating with MELCO partners from Kamaden, Johosoken and Sentansoken. We also have active collaborations with universities who will explore DiamondTouch as a collaborative input technology.

Future Direction: We will provide ongoing support for our collaborators and incorporate their feedback into future releases. We plan to investigate other input modalities (such as gestures). Overall, our main focus will be on adding value to existing or future MELCO products.

Contact: Kathy Ryall, Samuel Shipman, David Wong **Lab:** MERL Technology Lab
http://www.merl.com/projects/dtsdk/ **Project Type:** Advanced Development

Digital Merchandising



The goal of the Digital Merchandising project is to apply projection, sensing and real-time data analysis technologies to meet the needs of modern retailers. Initially, our focus has been on creating interactive, front-projected displays that automatically respond to normal customer behavior. Unlike traditional store displays, front projection allows new capabilities, such as non-flat displays, imagery placed directly on products, and seamless multi-device displays. A mix of capacitive, IR, and camera-based sensors give a rich view of customer characteristics and behavior which is then used to tune the displays to current customer needs.

Background and Objectives: MERL has a long history of research excellence in vision and data analysis, but the thrust into retail applications began fairly recently as an effort to identify new markets for video projectors. While various video display devices are used extensively in retailing, front projection has been almost entirely absent. Declining costs and improved projector brightness have recently made the use of video projectors in retail practical. Unlike traditional display devices, projectors naturally lend themselves to huge, immersive displays. By adding sensors and data analysis, we are able to extend these displays to create real-time, interactive environments. We believe that such systems will soon become the norm in retail establishments, and our task is to capture as much of this business for MELCO as possible.

Technical Discussion: Our first interactive display is a mock shoe store, which uses a single video projectors and capacitive proximity sensing, shelves. Normal customer interaction with the shoes triggers the system to provide more detailed information on the selected shoe. A new second-generation display in development uses three projectors, seamlessly blended to create a much larger display. Content for this multi-projector display can be created in Macromedia Flash. It is then rendered into OpenGL for the multiple projectors. The interaction is scripted in Python, which allows for complex system behavior.

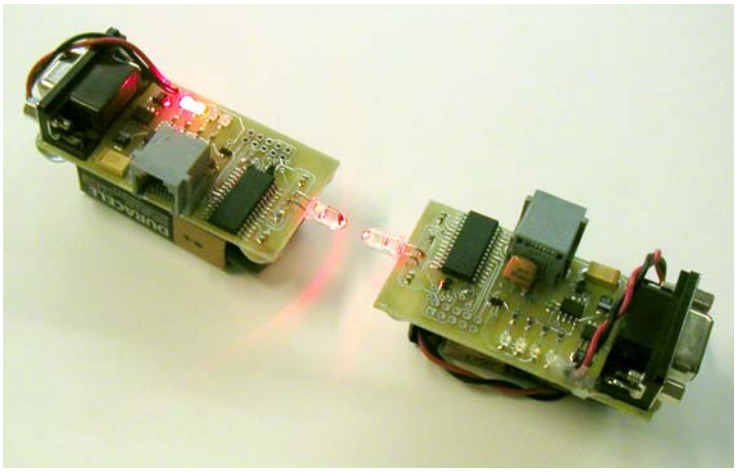
Collaboration: We are currently in talks with a potential retail partner to host an experimental display to be fielded in an actual store.

Future Direction: We hope to fully integrate MERL vision technologies in the near future. In addition, we are focusing on creating tools to make these systems more accessible.

Contact: Paul Dietz
http://www.merl.com/projects/digitalmerchandising/

Lab: MERL Technology Lab
Project Type: Research

Sensing and Communication Using Bi-Directional LEDs



Light Emitting Diodes (LEDs) are inexpensive and widely used as light sources. What is less well known is that LEDs are fundamentally photodiodes and as such are also light detectors. We have developed a novel microprocessor interface circuit that can alternately emit and detect light using only an LED, a resistor and two digital I/O pins.

Background and Objectives: This project began as an effort to create a smart backlighting system for television remote controls. A low-power capacitive proximity sensor detects active handling which in turn controls the backlight. To save battery life, the backlight should not be turned on in bright conditions. But adding a separate light sensor would require a new mechanical design for the remote, adding considerable cost. Our solution was to use the backlight LED itself as the light sensor. We developed a simple microprocessor interface technique that uses one additional digital I/O pin, but no other additional components compared to those need to simply light the LED. Since the circuit draws only microwatts of power, it has a minimal impact on battery life.

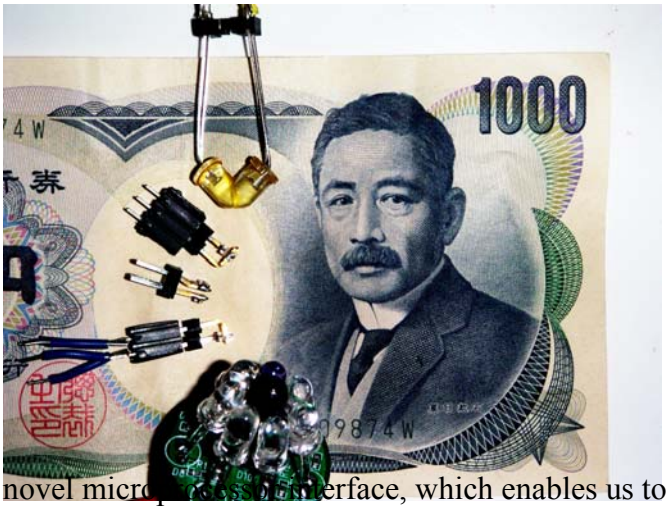
Technical Discussion: The LED microprocessor interface technique we have developed has far broader implications than simply controlling a backlight. By pointing two LEDs at each other, we can transmit data back and forth. This is trivially easy to do for distances on the order of a centimeter. The result is that almost any LED connected to a microprocessor can be thought of as a potential two-way communications port. We think of this technology as solving "the last centimeter problem" - you have two devices right next to each other, but they have no way to communicate. With LED Comm (LED-based communications), there is a link available, almost for free.

We have been looking into a variety of applications for this very fundamental technology. One opportunity is to turn the power light on appliances into a service port - download status and upload new firmware. Cell phones could swap phones numbers via their backlights. LED-based devices could replace RFID tags in many applications with benefits such as true peer-to-peer communications and a much smaller, far less expensive reader.

Contact: Paul Dietz, Darren Leigh, William Yerazunis
http://www.merl.com/projects/LEDcomm/

Lab: MERL Technology Lab
Project Type: Research

LED Chemical Sensors



Light emitting diodes (LEDs) have been used widely as light sources in low-cost chemical sensors, however, to date the fact that LEDs also behave as photodiodes for light detection has not been put to use. Detection in the past has been achieved most commonly with photodiodes, and although effective, this approach requires the added cost of an analog to digital converter. We have developed an ultra-cheap LED based sensing system that employs LEDs for both emission and detection. These devices are used in combination with a novel microprocessor interface, which enables us to measure changes in back-reflected light using a single digital I/O pin. The system holds great promise for following color changes in chemochromic sensing materials.

Background and Objectives: This project is an extension of work carried out on the development of sensing and communication using bi-directional LEDs. The device configuration has been modified to produce a single component incorporating two LEDs, one for emission and the other detection. The exterior of this is then dip-coated with a polymer membrane containing a chemochromic reagent. The color of this sensing film relative to the peak wavelength of the emitter LED affects the amount of back-reflected light received at the detector LED. The objective of this project is to develop colorimetric chemical sensors for a wide range of applications where sensor cost is an issue. Examples include environmental monitoring, clinical testing and ubiquitous sensing networks.

Technical Discussion: A wide range of chemical species can be detected using colorimetric reagents. Often these reactions are carried out in solution, but many can be translated into the solid-state by processing in an appropriate polymer matrix. Therefore, almost any colorimetric reaction could be immobilized onto the surface of a dual-LED to create a solid-state sensor. Uncoated Dual-LED sensors have also been used successfully to follow colorimetric reactions in solution, with detection possible at very low indicator dye concentrations. In addition to the sensing capabilities of this system, it holds the added bonus of offering short-range wireless communication of data.

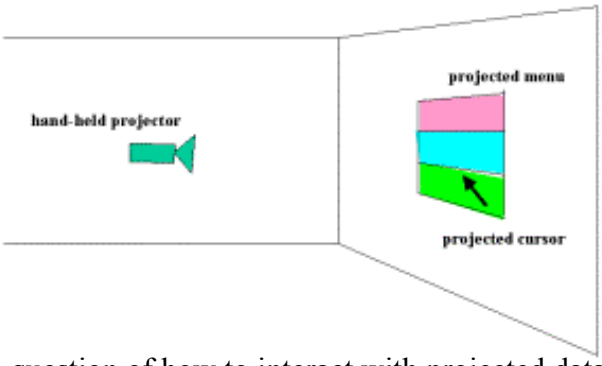
Collaboration: This is a joint effort of MERL and Dublin City University, Ireland.

Future Direction: Select specific applications. Incorporate sensors into MELCO systems.

Contact: William Yerazunis, Paul Dietz, Darren Leigh
http://www.merl.com/projects/LED_chemical_sensors/

Lab: MERL Research Lab
Project Type: Research

Mouse-Style Interaction with a Hand-held Projector



The decreasing size of projectors is opening the way for hand-held use. Furthermore, there is an obvious niche for hand-held projectors - an ordinary hand-held device with physical display cannot drop below a size where the display is readable; but for a device with projector, device size is decoupled from display size, so the device can become more and more compact while still providing an acceptable display.

Projecting from a hand-held device raises a question of how to interact with projected data in an analogous way to current mouse interactions. One could have a touch-pad on the device to control a cursor that moves independently across the projected data, but this puts a constraint on minimum device size. This project is investigating how to do mouse-style interaction with a projector plus mouse buttons, but with no need for a touch-pad.

Background and Objectives: A cellphone is a prime example of a device that could benefit from using a projector instead of (or in addition to) a physical display - the cellphone provides access to large amounts of data (say through web browsing), yet the small display size constrains what can be presented.

Technical Discussion: The projector has an attached motion-sensor. As an example interaction, say the user is projecting a menu onto a display surface and wishes to select a menu item using a cursor. The projected data conceptually has two parts: (i) the menu is projected so that it is motion-stabilized i.e. the menu appears fixed on the display surface even under projector motion; (ii) the cursor is not motion-stabilized - it may be fixed at, say, the center pixel of the projection - so its motion naturally reflects any pointing motion of the device by the user. Thus the cursor can be tracked across the stable menu until the desired item is reached, with a button click to select. This mechanism supports a large number of interactions apart from mouse interaction with projected data - for example cut-and-paste of textures from real surfaces; guided vision processing on a selected part of a scene; attachment, retrieval, and removal of virtual data in the environment.

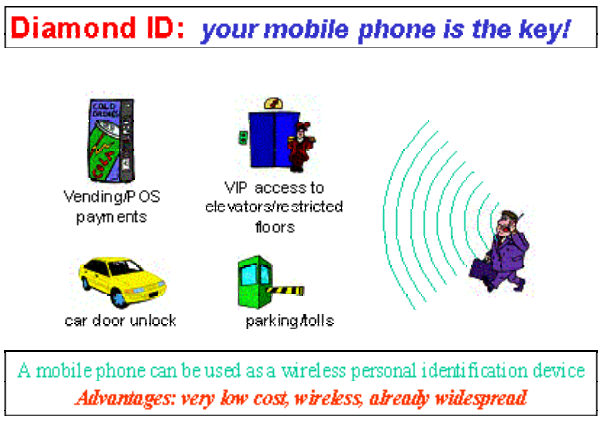
Collaboration: Dr. Shiotani, Power & Public Utility System Dept, Sentansoken, Sanden.

Future Direction: Projected augmentation for task guidance e.g. in factory maintenance tasks.

Contact: Paul Beardsley, Ramesh Raskar
http://www.merl.com/projects/ProjectorMouse/

Lab: MERL Research Lab
Project Type: Research

Diamond ID



Each mobile phone is assigned an unique telephone number. It can be used to identify the person using the phone. If a vending machine has the capability to receive the mobile phone number of the person trying to make a purchase, it can debit the charges electronically through the phone's account, enabling a new kind of convenient e-commerce payment method. Similarly, any kind of toll or fee collection devices may also take advantage of this new electronic transaction to collect payment.

Background and Objectives: The mobile phone has become extremely popular worldwide. At the same time, it also created fierce competition among mobile phone companies in order to attract subscribers. One of the ways to differentiate their service is to provide additional e-commerce features on the mobile phone itself. At the same time generate additional revenue beyond just the air-time charges. Diamond ID technology enables this e-commerce function on any existing modern mobile phone without any modifications to the phone itself.

Technical Discussion: An inexpensive ID receiver is placed inside the vending machine which decodes a special pattern of radio signal emitted from the mobile phone after the user initiates the electronic transaction. Since the transaction is wireless, it is very convenient to the user and reduce the chance of vandalism on the vending machine itself.

Collaboration: Johosoken - Multi-Media Department

Future Direction: MERL is actively searching for a mobile phone operator or e-commerce partner to develop this business opportunity.

Contact: George Fang
http://www.merl.com/projects/DiamondID/

Lab: MERL Technology Lab
Project Type: Research

Optimization Algorithms

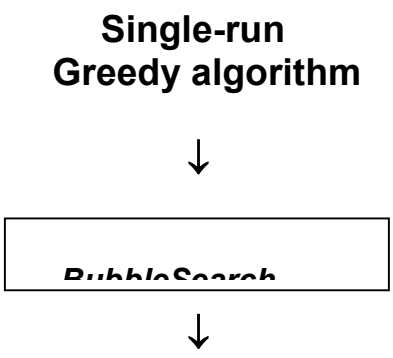
"Optimization Algorithms" cover a wide range of technologies for producing the best solution possible for a given problem according to predefined criteria, such as to maximize efficiency or minimize cost. Optimization algorithms play an essential role in many industries and scientific endeavors, from manufacturing to biotechnology. The airline industry, for example, uses optimization algorithms to produce flight schedules and crew schedules, attempting to minimize wait times for both equipment and staff members. While the techniques vary, all optimization algorithms are essentially methods for performing a series of calculations until an ideal solution is constructed or a termination condition is reached. Today's optimization algorithms produce higher quality solutions with smaller amounts of computation than their predecessors. Furthermore, as the speed of computers increases, optimization algorithms become more powerful tools for streamlining business, reducing operating costs, and scientific exploration.

MERL has growing expertise and accomplishments in many areas of optimization. Researchers at MERL work on both discrete and continuous optimization and contribute new frameworks for optimization as well as algorithm advances on particular problems. Recently, MERL has produced novel solutions to optimization problems that arise within machine learning, data mining, operations research, control theory, decision making, and inference tasks. A common trait among all projects is the focus on application areas of practical importance and on techniques with principled theoretical underpinnings.

Year 2002 projects include optimal control of groups of elevators in tall buildings; data-mining extremely large data-sets to predict consumer behavior, a new class of algorithms for inference problems such as correcting errors in data transmitted over a noisy channel; training binary classifiers for recognition and identification tasks, and generic techniques for improving domain-specific greedy algorithms.

Project Descriptions

BubbleSearch: Generic Methods for Improving Greedy Algorithms



We have developed generic methods for extending a popular class of greedy algorithms called priority algorithms. Our techniques produce anytime algorithms that can find better solutions than the original priority algorithm, given more computation time. Our techniques modify the order in which the priority algorithm considers the problem elements based on the Kendall-tau distance between orderings. The Kendall-tau distance is also known as the BubbleSort distance, which is why we call our approach BubbleSearch. We present both exhaustive and randomized variations of BubbleSearch. Our experiments in four domains show that even 100 iterations of BubbleSearch substantially improves the result compared to the original priority algorithm.

Background and Objectives: Priority algorithms have historically been used and are especially effective at solving many scheduling and packing problems. They are attractive because they are simple to design and implement, execute quickly, and often provide very good heuristic solutions. For some problems, they are the best-known heuristic. We have developed generic algorithms and software, which can extend any priority algorithm that, fits our formulation. Our approach is essentially an easy way to augment a good priority algorithm to make it better.

Technical Discussion: Priority algorithms consist of an ordering function and a placement function. A priority algorithm sequentially assigns values to elements according to the ordering of elements using its placement function. For example, the First Fit Decreasing algorithm for bin packing sequentially places items in order of decreasing size, where each item is placed into the lowest numbered bin in which it will fit. This paper explores our hypothesis that, while the solutions produced by priority algorithms are typically quite good, there are often better solutions "nearby" that can be found with small computational effort. Our hypothesis yields a natural and generic approach to extend priority algorithms to yield better solutions given additional running time. This approach uses the priority algorithm's ordering as a hint for how to explore the space of possible orderings. In particular, our algorithm uses the priority algorithm's placement rule to evaluate orderings that are close to the original ordering based on the Kendall-tau distance, also known as the BubbleSort distance.

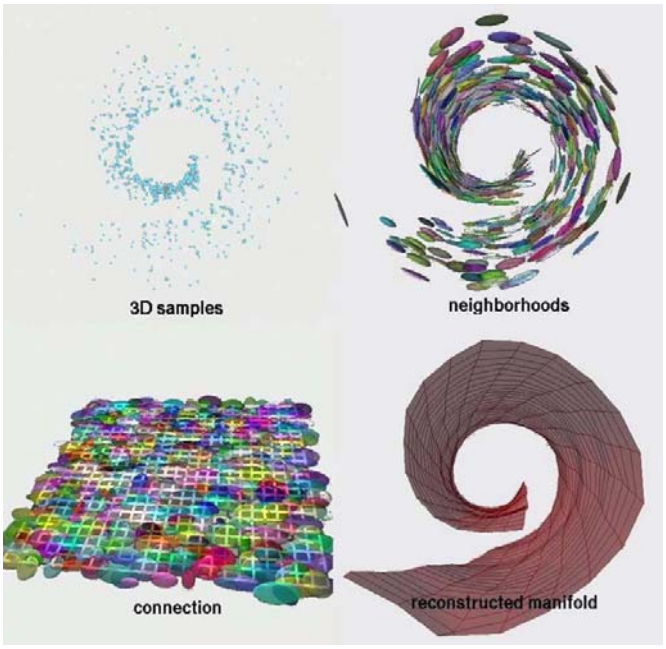
Collaboration: This work is done in collaboration with Harvard University.

Future Direction: We are interested in investigating methods to automatically tune the (two) parameters of BubbleSearch and to adapting data mining techniques to discover good orderings of elements given a placement function and a problem instance.

Contact: Neal Lesh
http://www.merl.com/projects/BubbleSearch/

Lab: MERL Research Lab
Project Type: Research

Nonlinear Dimensionality Reduction by Charting



Charting is a new method for estimating a manifold from samples. For example, the manifold of all human faces can be estimated from pictures, which are points in a 10^5-dimensional image space. Charting imposes a global low-dimensional coordinate system on the manifold and provides functions that can convert between high-dimensional samples and low-dimensional coordinates. The illustration shows a how a 2D manifold curled in 3-space is reconstructed from samples. Charting has also been used for face synthesis, modeling of physical phenomena, and discovery of word meanings from usage patterns. The chart is a highly effective means for compression, noise-reduction, decompression, and synthesis of samples.

See Color Figure 6

Background and Objectives: Dimensionality reduction of high-dimensional signals such as images, video, and audio is necessary first step for most signal processing applications. The dominant method is principal components analysis, a linear transform of the data that is most appropriate when the manifold is strictly flat. Charting is a generalization of PCA to handle data sampled from manifolds that are arbitrarily curved. It gives much more efficient encodings than PCA, compressing to fewer dimensions yet decompressing to more accurate reconstructions.

Technical Discussion: Charting is analogous to a book of street maps: Each map covers one neighborhood, and there is a way of pasting together the maps to cover the whole city. There may be distortion, because the combined map is 2D while the city has 3D shape due to its hills and valleys and the curvature of the planet. Charting first identifies a collection of low-dimensional spaces that individually cover local neighborhoods on the manifold, such that adjoining spaces are similarly oriented. It then computes a soft affine mixture of these subspaces that minimizes distortion over the whole manifold. The result is two functions that offer smooth continuous mappings between the global chart and the original data space.

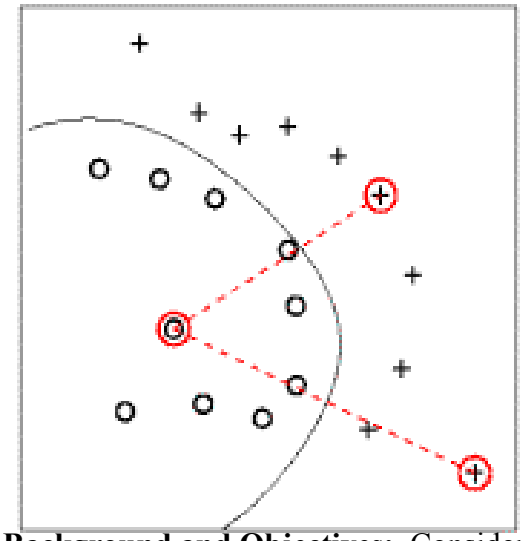
Collaboration: This technology is begin developed for MELCO's new data mining business.

Future Direction: We are now exploring methods for charting manifolds that are not isomorphic to Euclidean spaces.

Contact: Matthew Brand
http://www.merl.com/projects/charting/

Lab: MERL Research Lab
Project Type: Research

Hypercuts: Boosted Dyadic Kernel Discriminants



We present a novel machine learning (ML) algorithm for training complex binary classifiers by superposition of simpler hyper plane discriminants formed by projection of a test point onto a "dyad" (an oppositely labeled pair of training points). The learning is based on a real-valued variant of AdaBoost and the constituent hyperplane discriminants use kernels of the type used by Support Vector Machines (SVM). The resulting kernel classifier has an accuracy / complexity tradeoff which is more flexible than that of a SVM, is amenable to on-line adaptive learning (unlike SVMs) and has a classification performance which is as good (if not better) than comparable SVMs.

Background and Objectives: Consider a collection of linear discriminants, each formed by means of a hyperplane orthogonal to the vector connecting a "dyad" (ie. a pair of data points with opposite class labels). By applying the method of linear hypercuts to a nonlinearly transformed feature space (using Mercer kernels) we obtain nonlinear discriminants similar to SVMs. Using AdaBoost, the search for each new candidate hypercut explores the space of all possible dyadic classifiers. For both linear and nonlinear classification problems, boosted dyadic hypercuts form an efficient search strategy for exploring the space of all classifier hypotheses.

Technical Discussion: The ensemble of dyadic hypercuts is learned incrementally by means of a real-valued or confidence-rated version of AdaBoost, which provides a sound strategy for searching through the finite (but possibly large) set of hypercut hypotheses. In experiments with real-world datasets from the UCI ML repository, the generalization performance of the hypercut classifiers was found to be comparable to that of SVMs and k-NN classifiers. Furthermore, the computational cost of classification (at run time) was found to be similar to, or better than, a comparable SVM. In contrast to SVMs we offer an on-line and incremental learning machine for building kernel discriminants whose complexity (number of kernel evaluations) can be directly controlled (traded off for accuracy).

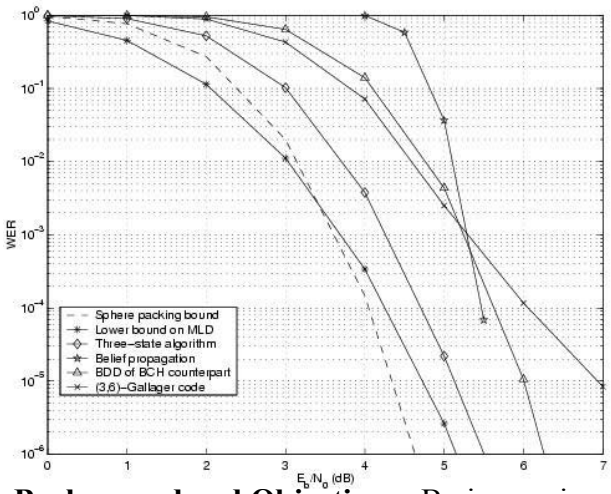
Collaboration: This MERL project was joint work with Gregory Shakhnarovich (MIT).

Future Direction: We are exploring optimal strategies for sampling the hypothesis space based on the AdaBoost weight distribution and also investigating dyads that are not necessarily based on training samples (eg. using cluster centroids instead). Also, this distribution may help devise search strategies that can adapt with each round of boosting in the online learning paradigm.

Contact: Baback Moghaddam
http://www.merl.com/projects/hypercuts/

Lab: MERL Research Lab
Project Type: Research

Iterative Decoding of Classical Codes



We have developed a new method to decode intermediate length multi-step majority logic decodable (MSMLD) codes. The most famous in this class of codes are the Reed-Muller codes, which were introduced in the 1950's, and found many applications in the 1960's. MSMLD codes have typically been decoded using a very simple (sub-optimal) decoding algorithm, but they have inferior minimum distance properties when compared with BCH codes, which have now largely eclipsed them.

Background and Objectives: By improving the MSMLD decoding algorithms to more closely resemble the iterative decoding algorithms that are successful with turbo-codes and low-density parity check (LDPC) codes, we have been able to improve their performance to the point where they surpass BCH codes (decoded in the standard way) on the binary symmetric channel, at least for low and moderate signal to noise ratios. In fact, in the intermediate blocklength regime, our decoding algorithms give the best known performance for rate 1/2 codes on the binary symmetric channel.

Technical Discussion: Our algorithm is a simple three-state bit-flipping algorithm. We use highly redundant parity check matrices to represent MSMLD codes. Bits decide whether they should agree with their received value, disagree, or abstain from making a decision, based on the number of parity checks that they belong to which "vote" in the various directions. For an MSMLD code of blocklength 255 and rate 1/2, we are able to decode within a few tenths of a decibel of optimal maximum likelihood decoding, and we can demonstrate that we outperform standard decoding of BCH codes down to word error rates of about one in ten trillion.

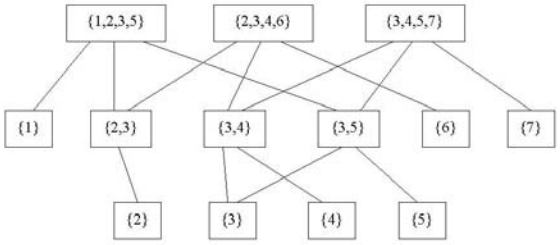
Collaboration: Marc Fossorier (University of Hawaii) and Ravi Palanki (Caltech)

Future Direction: Our decoding algorithm does not scale very well to MSMLD codes with blocklengths more than about 1000, but one can build longer codes in a variety of ways out of short ones. We are investigating such methods to try to obtain similar performance for longer codes.

Contact: Jonathan Yedidia
http://www.merl.com/projects/msmld/

Lab: MERL Research Lab
Project Type: Research

Representing Codes for Belief Propagation Decoding



The best error-correcting codes of short and intermediate block length that have so far been discovered are usually defined in ways that do not immediately suggest a sparse parity check matrix representation. For this reason, large classes of classical textbook codes, which would give excellent performance under optimal decoding, have been mostly ignored as candidates for the belief propagation (BP) decoding algorithm. Since BP decoders have proven so successful in decoding turbo-codes and low-density parity check codes and their relatives, it is natural to ask whether classical codes can also be represented in a way that allows them to be decoded using BP.

Background and Objectives: Any linear block error-correcting code can be represented in many equivalent ways by a parity check matrix, which defines the constraints that the different bits in the code must obey. Some parity check matrices are more suitable for BP decoders, namely those such that 1) the number of ones in each row is small, 2) the number of ones in each column is large, and 3) the number of pairs of rows that share ones in more than one column is small. We developed an algorithm that given an input parity check matrix, outputs a new parity check matrix that is improved in all of these characteristics, though it also has the drawback of introducing "auxiliary" bits which get no evidence from the channel.

Technical Discussion: When applied to Euclidean Geometry codes, this algorithm significantly improves the performance of BP decoding on the Binary Erasure Channel, but it does not help on the additive white Gaussian noise (AWGN) channel. The problem seems to be caused by the large number of auxiliary bits, which can send incorrect internal messages on this channel. Nevertheless, BP decoding of EG codes on the AWGN channel performs relatively well (within 0.8 dB of optimal maximum likelihood decoding) if one uses highly redundant parity check matrices.

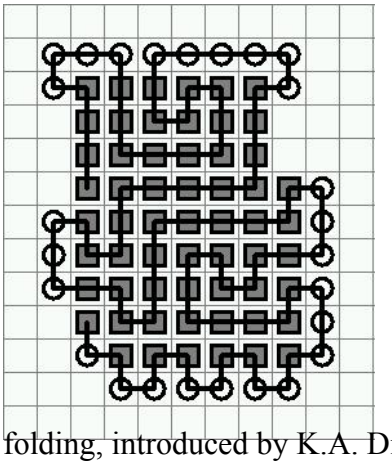
Collaboration: Marc Fossorier and Jinghu Chen (University of Hawaii)

Future Direction: We are currently investigating other related techniques to decode classical codes using iterative decoding algorithms.

Contact: Jonathan Yedidia
http://www.merl.com/projects/representing_codes/

Lab: MERL Research Lab
Project Type: Research

A Complete and Effective Move Set for Simplified Protein Folding



We have developed a new approach for solving a simplified version of the protein-folding problem. Our techniques were able to generate new best solutions for several of the largest benchmarks available in the literature. Our algorithm uses standard search techniques. Our key contribution is a novel set of transformations on protein configurations, which we call "pull moves". Theoretically, we have shown that pull moves have desirable properties for use in search algorithms. Experimentally, we have shown that a generic implementation of tabu search using pull moves performs extremely well on benchmark problems.

Background and Objectives: We have worked on the two-dimensional hydrophobic-hydrophilic (2D HP) model for protein folding, introduced by K.A. Dill in 1989. Although the model is simple, it is thought to capture many of the main features of the protein-folding problem. Thus, techniques that work well on the simple problem may be useful for the real protein-folding problem. This research has been performed in the context of the ongoing Human-Guided Search (HuGS) project on interactive optimization. While we have developed a fully-automatic system, the interactive visualizations of HuGS were critical to allowing the researchers on the project to understand the problem and develop solutions. Pull moves were originally designed to help a user quickly manipulate the current protein configuration but turned out to be very effective for computer search as well.

Technical Discussion: Solving a 2D HP problem instance involves placing a sequence of amino acids, each labeled as either hydrophobic (a square in the example) or hydrophilic (a circle in the example). Each amino acid must be placed horizontally or vertically adjacent on the grid to its neighbors in the sequence. The goal is to find the configuration with minimal energy, which typically means maximizing the number of adjacent hydrophobic pairs.

Pull moves typically begin by repositioning two amino acids that are neighbors in the given sequence to new, adjacent grid locations that are currently free. This may introduce a gap between two neighbors in the given sequence, but a valid configuration can be obtained by repositioning additional amino acids into grid locations that were occupied before the pull move began. In the worst case, every amino acid might be repositioned, but often a valid configuration is reached after only repositioning a few amino acids.

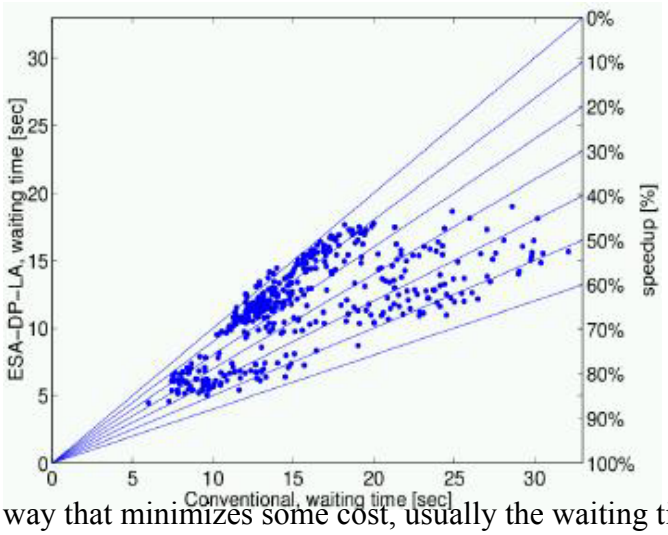
Collaboration: We are collaborating with professors in Harvard and McGill University.

Future Direction: We are interested in applying our techniques to more challenging protein-folding problems, starting with the 3-dimensional HP model and perhaps working toward 3-dimensional problems without an underlying unit grid. We would also like to further develop our interactive system for protein folding, built with the Human-Guided Search software, which would allow people to better understand particular protein configurations as well as apply their real-world knowledge to the problem.

Contact: Neal Lesh
http://www.merl.com/projects/protein-folding/

Lab: MERL Research Lab
Project Type: Research

Look-ahead for Group Elevator Control



In up-peak elevator traffic, performance measures such as average waiting time are dominated by new passengers arrivals at the lobby. Optimal scheduling requires balancing the waits of known passengers against those of passengers who have not yet arrived in the building. We do so in a decision-theoretic manner, realizing gains in system efficiency so sharp that some buildings can now be properly serviced with fewer elevator shafts.

Background and Objectives: Group elevator control is the problem of assigning elevators to pick up waiting passengers in a way that minimizes some cost, usually the waiting times of all passengers. The efficiency of the controller determines the capacity and throughput of a bank of elevators, and ultimately the cost of installing and running an elevator system.

We seek to put elevator scheduling on an optimal decision-theoretic basis. This requires modeling the effects of three kinds of passengers: those in cars, those who have signaled for service, and those who have not yet signaled their presence. The look-ahead project is concerned with minimizing waits of these future passengers.

Technical Discussion: Group elevator scheduling is an NP-hard sequential decision-making problem with unbounded state spaces and substantial uncertainty. We recently developed a decision-theoretic solution for the expected waiting times of all passengers in the building, marginalized over all possible passenger itineraries. Though commercially competitive, this solution does not contemplate future passengers. To that end, we developed a probabilistic model of how the pattern of elevator lobby landings affects the waits of future arrivals. This is discounted to reflect growing uncertainty in the remoter future, then balanced against the expected waiting times of known passengers to give a measure of the future cost of any scheduling decision.

Collaboration: Kouichi Sasakawa and the building systems group of the Industry Solutions Technology department, Sanken.

Future Direction: For further evaluation, this scheduler will be adapted for different elevator physics in MELCO's detailed elevator simulator and an industry-standard simulator. We will combine the look-ahead method with a new solution for bounding the maximum wait of any individual passenger. We are also developing solutions for optimal estimation of current and near-future traffic patterns.

Contact: Matthew Brand, Daniel Nikovski
http://www.merl.com/projects/ElevatorControl/

Lab: MERL Research Lab
Project Type: Research

Data-mining and Recommending

The Instant Movie Recommender (IMR) predicts how you would rate over 1000 movies on the basis of your ratings of a few familiar ones. As you vary the rating of one movie, all the other movies are instantly re-ranked and re-sorted. Predictions are made on the basis of correlations between ratings by previous users -- people who like "X" tend not to like "Y", etc. The IMR "learns" from your ratings -- updating the correlational model "on the fly". This update is a form of data mining, but with substantially better economics and usability than the standard practice of warehousing and batch-processing data.

The IMR showcases a new technology for fast updating of a "thin" Singular Value Decomposition -- a decomposition of tabular data into simple factors. The technology is distinguished both by its speed -- it is the first linear-time single-pass algorithm -- and by its ability to handle tables with many missing elements -- a common problem in data mining.

Background and Objectives: The SVD forms the core of many data mining algorithms and thousands of algorithms in signal processing. The "thin" SVD decomposes tabular data into the product of two small matrices, and is very useful for data compression, noise suppression, and prediction. For very large tables, computing an SVD is impractical because the compute time grows quadratically with the size of the table. Worse, the SVD is not uniquely defined if the table is missing some entries. We set out to develop an SVD algorithm that is suitable for data sets that are far too large to fit into the computer's memory and that are missing many elements.

Technical Discussion: The new Incremental Imputative SVD (IISVD) allows an SVD to be computed from streaming data. The data table arrives one row or column at a time, and the SVD is updated to reflect the newly arrived information. The data need not be stored. If the row is missing entries, the algorithm chooses values that are most consistent with the correlational structure exhibited in previous updates. This is how the IMR predicts ratings of unseen movies. Equally fast rank-1 SVD updates allow users to change or retract their ratings. Experiments with a dataset used by the data mining community for benchmarking indicate that the IMR is quite accurate, making predictions within 2 rating points of the "true rating" more than 99% of the time.

Collaboration: An IISVD library and IMR demo is being developed for Johosoken and MDIT to demonstrate the effectiveness of our "thin" SVD. A retail menswear vendor has completed a test of the system for targeted marketing and FCS is scheduled soon for a travel agency in Japan.

Future Direction: Future plans include expanding the application of this algorithm to time-series forecasting as well as generalizing the algorithm itself for nonlinear relations.

Contact: Matthew Brand, Frederick J. Igo, Jr., David Wong **Lab:** MERL Research Lab
http://www.merl.com/projects/DataMining/ **Project Type:** Advanced Development

Speech and Audio

Spoken interaction with devices is becoming increasingly important. There has been a dramatic increase in the complexity of consumer electronics, cellular telephones, and mobile and in-vehcle computing devices. This creates a market need for Spoken Language Interfaces in order to simplify the user interface and free-up the hands and eyes. MERL is working to provide technologies to enable the production of high-quality, easy-to-use Spoken Interfaces. In addition, we are using Sound Recognition to enable machines to listen and understand their surrounding sonic environment.

Speech Application Infrastructure. MERL's SPIEL Toolkit is middleware for Speech applications. It enables the rapid construction of new applications and reduces risk by providing a common API that separates applications from the specific APIs of the speech engines. DiamondTalk is an application-independent Java architecture for building conversational, multimodal Spoken Language Interfaces.

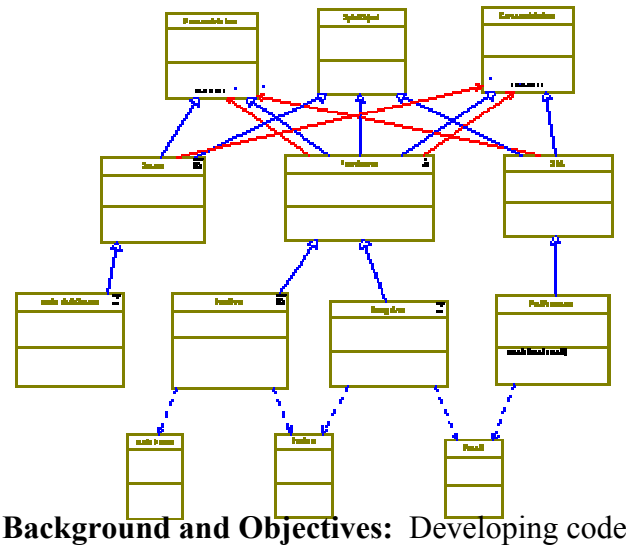
Novel Spoken Interface Technologies. MERL's SpokenQuery technology can enable Information Retrieval using only spoken queries. MediaFinder, which is built on SpokenQuery, is a series of prototype products that explore the problem of navigating large amounts of music and information with a hands and eyes free interface. MERL is also exploring speech-centric devices.

Conversational Speech Interfaces. With MERL's COLLAGEN technology, we are applying principles from human collaborative discourse theory to build a conversational framework on top of which we can layer speech interfaces. We are exploring mutimodal web interaction in our FormsTalk project. MERL's Voice Programming is a new approach to making it easy for people set up, customize, and operate the increasingly complex devices found in modern homes.

MERL's Sound Recognition technology will provide the basis for various types of audio sensing which can find applications in surveillance, factory automation, entertainment media analysis among others.

Project Descriptions

SPIEL Toolkit



The SPIEL Toolkit is middleware for Speech applications. It enables the rapid construction of new applications and reduces risk by providing a common API that separates applications from the specific APIs of the speech engines. This makes the application independent of the speech engine thereby enabling reuse of application code and fair performance comparisons of speech recognition engines. SPIEL is specified in UML and is platform and programming language independent. It is especially suitable for embedded, server, and distributed applications.

Background and Objectives: Developing code that controls the speech recognition process is costly, as it must concurrently control the capture of audio data, signal processing, pattern matching, and natural language processing in a limited memory and processor footprint. Further, the probabilistic behavior of speech recognition engines requires special logging support for reproducible testing and tuning of applications. Much of this logic is common between applications. Our objective is to identify and abstract this shared behavior as a common architecture. This greatly simplifies the production of robust and efficient applications. Since applications are not tied to a particular engine, this also enables a "best of breed" approach in choosing speech engines.

Technical Discussion: For a number of reasons, current industry standard speech middleware, such as SAPI from Microsoft and JSAPI from Sun, are not suitable for embedded, distributed or server applications. Unlike SAPI and JSAPI, SPIEL does not tie the application to an OS or language platform. Since SPIEL is specified in UML, it can be produced as C, C++ or Java, and on Windows, Linux, WINCE, and many embedded platforms.

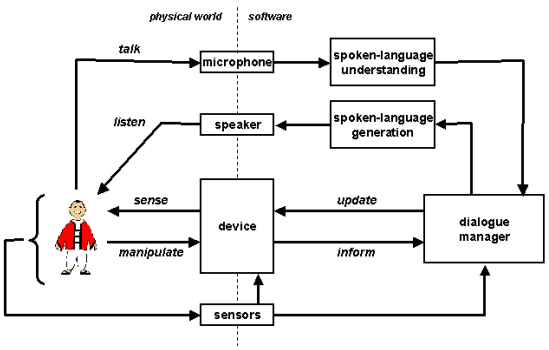
Collaboration: SPIEL forms the core of several speech projects including Speech Engine Evaluation, SpokenQuery, FormsTalk, MediaFinder, SpeechServer, DiamondTalk, One Button Phone, ComBadge, "The Cellphone of the Future" CeBit demo and the "Automobile of the Future" demo. Other MELCO business units have expressed interest in SPIEL. We are currently aware of projects for which SPIEL would be suitable in the Communications, Entertainment, Appliance and Automotive Business Units.

Future Direction: In order to make SPIEL more useful for prototyping and evaluating speech applications and engines, we plan to extend it by implementing interfaces for additional speech recognition, compression and text to speech engines.

Contact: Peter Wolf
http://www.merl.com/projects/SPIEL/

Lab: MERL Technology Lab
Project Type: Advanced Development

DiamondTalk: A Java Architecture for Spoken-Language Interfaces



DiamondTalk is an application-independent Java architecture for building conversational, multimodal spoken-language interfaces, especially for embedded applications, such as in automobiles, cell phones, home appliances, and robots. The top level components of DiamondTalk are shown in the diagram at the left. The first use of DiamondTalk has been as the platform for implementing a multimodal form-filling application (see FormsTalk project).

Background and Objectives: Building spoken-language interfaces, especially multimodal (e.g., involving both speech and touch or manipulation) and conversational ones, currently requires a large amount of specialized development for each application, and tends to be closely intertwined with the choice of speech recognition and generation engines. The goals of the DiamondTalk architecture are to reduce the amount of specialized development required for each new application by increasing the amount of code reuse (including testing and data collection tools) and to make it easy to substitute different speech engines in an existing application to take advantage of improved technology.

Technical Discussion: DiamondTalk is based on the Java Beans component architecture; components in other programming languages can easily be integrated using the Java Native Interface facilities. DiamondTalk also takes full advantage of the internationalization facilities of Java, so that it can be used, for example, to build interfaces in English or Japanese. DiamondTalk components communicate by sending events (shown by the arrows in the figure above) and by querying each other's state (not shown). The dialogue manager is at the center of the architecture. It receives information about device state changes and about the user's actions, utterances, and state changes. Based on this information, it can update the device state and produce representations of utterance meaning, which are sent to the spoken-language generation component. Collagen (see project description) is the default implementation of the dialogue manager. The spoken-language understanding component is typically decomposed into a speech recognizer and a semantic analyzer; the spoken-language generation component is typically decomposed into a language generator and a speech synthesizer.

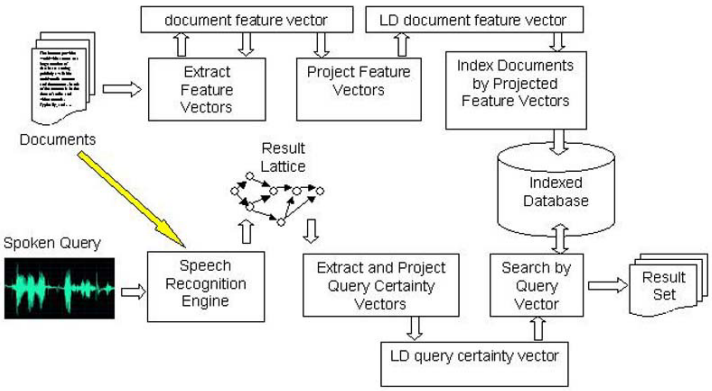
Collaboration: This is a joint project between MERL Research and MERL Technology.

Future Direction: We are beginning to retrofit some of our existing spoken-language interface prototypes to the DiamondTalk architecture, as was well as planning to use DiamondTalk for new applications. We expect to extend and improve DiamondTalk based on these experiences.

Contact: Charles Rich
http://www.merl.com/projects/DiamondTalk/

Lab: MERL Research Lab
Project Type: Research

SpokenQuery



SpokenQuery is technology for accessing databases using a verbal description of the desired information. It can be used to retrieve information such as web documents, music, government forms and industrial documentation using only speech. It is particularly useful in applications where hand and eyes free operation is desired. These include: call centers, information kiosks, automotive entertainment systems, Telematics, home entertainment systems, cellphone information systems, and hand held industrial systems.

Background & Objective: For many users, search engines such as Inktomi, Google and AllTheWeb have become the primary method of locating information on the Internet. In consequence, Information Retrieval (IR) has become an important and very lucrative technology. However, the current set of search engines all require typed input, and there are many situations where a keyboard is not acceptable. Clearly, typing queries would not be acceptable while driving an automobile. The ability to access information in the automobile, on the cellphone and on PDAs combined with an interface that is hands and eyes free will enable large new markets. The objective SpokenQuery is to enable Information Retrieval using only spoken queries.

Instead of typing the query, the SpokenQuery user verbally describes the desired information. The result is a list of items that are judged to be "pertinent" to the query. Similar to current IR systems, the list is not exact but should contain a significant number of useful items.

Technical Discussion: One naïve way to implement a SpokenQuery system would be to take the text output of an open-recognition dictation system (e.g. ViaVoice, NaturallySpeaking) and feed it as input to a current IR system (e.g. Inktomi, Google). However, this implementation would not perform very well in applications that are noisy, have a far field microphone, or are speaker independent. Unfortunately, applications such as the automobile, cellphone or PDA usually have all of these problems. Therefore MERL has developed a technique that uses more information than the simple text output of a recognizer - SpokenQuery considers all the words that might have been spoken and combines this information with knowledge of the database contents (i.e. what makes sense). Our algorithm takes advantage of the low correlation between acoustics and semantics and produces good retrieval results even in extremely challenging circumstances.

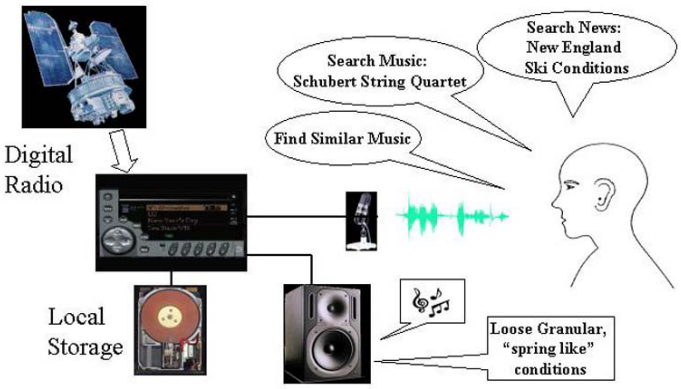
Collaboration: AdvisoryAgent, KnowledgeProvider, AdaptiveAgent, MediaFinder, FormsTalk.

Future Directions: Improve performance; Reduce memory and processor footprint; Produce prototype products and services for MELCO Business Units

Contact: Peter Wolf, Bhiksha Raj
http://www.merl.com/projects/SpokenQuery/

Lab: MERL Technology Lab
Project Type: Research

MediaFinder



The MediaFinder project is a series of prototype products that explore the problem of navigating large amounts of music and information with and hands and eyes free interface. The prototypes are aimed at future automotive, home entertainment, handheld entertainment and cellphone markets where the combination of inexpensive data storage, digital broadcasting, and wireless access to the internet enable music and information on demand services.

Background and Objectives: MediaFinder addresses an issue that is already a huge problem with current consumer electronics products. Current handheld MP3 players (e.g iPod) allow storage personal of personal collections of up to 2,000 songs and automotive units allow up to 16,000 songs. Unfortunately, no current product offers any acceptable way to navigate such a large collection. Neither a large graphical user interface, nor a pointing device (e.g. mouse or touch screen), nor a keyboard is possible on a handheld unit or in an automobile, so all that most current products provide is a tiny display of up to about 10 songs, and "next" and "previous" buttons. If current interfaces are poor for static collections of a few thousand songs, consider the limitations they would impose on music and information on demand services. The objective of the MediaFinder project is to explore methods of requesting and navigating music and information that would be acceptable in the automobile, on a cellphone and in the living room. Each of these three interfaces have very different requirements, however they all share the same problem - a traditional GUI with a keyboard and mouse is not acceptable; and they all share the same need - hands and eyes free operation with a minimum of button and display.

Technical Discussion: MediaFinder is an integration of three MERL technologies: SpokenQuery, SoundSpotter and MusicSkimmer. SpokenQuery allows the user to retrieve the desired music or information by describing it by voice; SoundSpotter allows the user to retrieve other music that is similar to a piece of music; and MusicSkimmer allows the user to browse a collection of music by listening to it very quickly. Combined with a minimal physical interface of a few buttons and a small display, these technologies are used to create a user interfaces that could be used in the automobile, living room or cellphone.

Collaboration: AdvisoryAgent, AdaptiveAgent, MediaFinder, FormsTalk.

Future Direction: Produce prototype products and services for MELCO Business Units.

Contact: Peter Wolf
http://www.merl.com/projects/MediaFinder/

Lab: MERL Technology Lab
Project Type: Advanced Development

ComBadge: A Voice-Operated Communications Device

The ComBadge is a two-way voice pager with a simple spoken user interface. The project encompasses the hardware, software, and user interface designs. A primary design goal has been to reduce the users' cognitive load, thus creating a communications device that is very simple and natural to use. We aim to appeal to those segments of the market where cell phone penetration is lowest, including children, the elderly, and the less-wealthy in the world.

Background and Objectives: Device costs are kept low by eliminating the display and keypad. Infrastructure costs are reduced by allowing more devices to share the available bandwidth because the messages are relatively short and the communication is asynchronous. The spoken command set is small, so that it can be easily learned and remembered, and recognized with few errors. Familiar names are used to contact other users by having each user add customized voice name tags for other ComBadges.

Technical Discussion: Speech recognition, audio compression, and radio transmission do not overlap, thereby reducing the peak power demand and extending battery life. Compression need not occur in real-time, which permits the use of a slower processor and/or a better compression algorithm. Inexpensive bandwidth intended for data, rather than voice, can be used at all stages of the network. Message delivery could be accomplished over the Internet. Asynchronous messaging also has advantages for users. The device can be very small, since it does not need to reach from mouth to ear. Users are less aware of dead spots in network coverage and are less irritated by network outages due to overloading, since these conditions produce delays rather than dropped calls. Furthermore, the ComBadge is less intrusive because users determine when they want to listen and respond to messages.

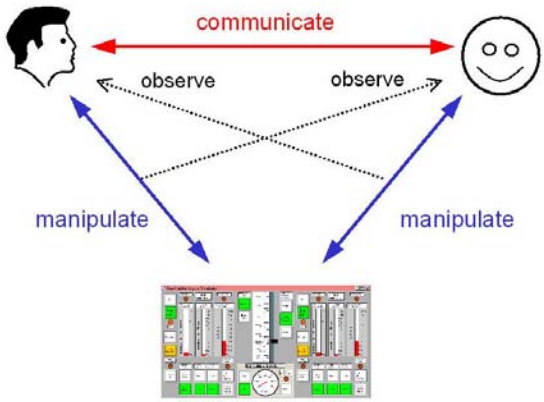
Collaboration: We are working very closely with the speech applications group at MERL Technology Lab.

Future Direction: We are exploring using the ComBadge for group messaging -- a spoken distribution list allows the same message to be delivered to several users. Delivering voice messages to machines as well as people is an open-ended opportunity. Initially, we are looking at using ComBadge to control household devices and as a voice portal to a wide variety of Internet services.

Contact: Joe Marks
http://www.merl.com/projects/ComBadge/

Lab: MERL Research Lab
Project Type: Research

COLLAGEN: Java Middleware for Collaborative Agents



COLLAGEN (for COLLaborative AGENt) is Java middleware for building collaborative agents. A collaborative agent is a software program that helps users solve problems, especially in complex or unfamiliar domains, by correcting errors, suggesting what to do next, and taking care of low-level details. A collaborative agent can be added to an existing graphical user interface, such as a software simulator, or integrated into the design of a new hardware device, such as a home appliance.

Background and Objectives: The theoretical foundations of COLLAGEN derive from the study of naturally occurring human collaboration, such as two people assembling a complex mechanical device or two computer users working on a spreadsheet together. The practical objective of the project is to maximize the software reuse in building collaborative agents for many different applications.

Technical Discussion: A collaborative agent in general communicates with the user (using either natural or artificial language), manipulates some shared hardware or software artifact, and observes the user's manipulation of the shared artifact. The key to COLLAGEN's application-independence is an abstract, hierarchical representation, called the "task model," of the sequences of actions typically performed to achieve goals in a particular domain. The task model captures all of the knowledge that is specific to a particular application. In essence, COLLAGEN is an "interpreter" for task models. COLLAGEN's representation of the current state of a collaborative dialogue consists of a plan tree, which tracks the status of steps in the task model, and a focus stack, which tracks the current focus of attention. COLLAGEN automatically updates these data structures whenever either the user or the agent speaks or performs a manipulation. The agent then uses these data structures to determine what to do or say in response.

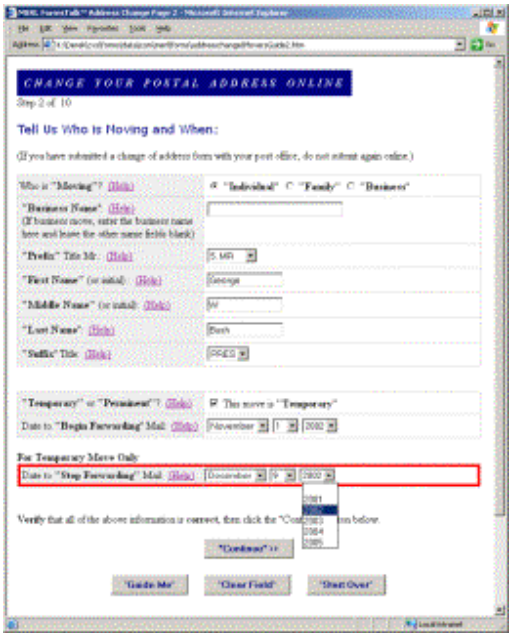
Collaboration: COLLAGEN is currently being used to build prototype systems for a range of applications, including power plant operator training (see Intelligent Agents for Operator Training project description), a programmable home thermostat (see Voice Programming project description), multimodal web form filling (see FormsTalk project description) and a hosting robot (see Human-Robot Interaction project description).

Future Direction: In addition to continuing to seek new applications (especially involving speech), we plan improvements in the basic operation of COLLAGEN in the areas of turn taking, causal knowledge, and negotiation.

Contact: Charles Rich, Neal Lesh, Candace Sidner
http://www.merl.com/projects/collagen/

Lab: MERL Research Lab
Project Type: Research

FormsTalk: Multimodal Mixed-Initiative Web Form Filling



FormsTalk is middleware for building web form-filling applications that support speech and touch, as well as conventional keyboard and mouse interaction. Our initial focus is "e-government" applications, such as filling out a postal change-of-address form (see left). Since our target user population includes occasional users with little or no prior technology experience, FormsTalk supports flexible, mixed-initiative interaction, in which either the user or the computer can take the lead, depending on circumstances.

Background and Objectives: Form-filling is a common framework for many different kinds of web applications. A multimodal approach improves the accessibility of these applications by allowing users to choose whichever mode (speech, touch, etc.) is the best match for their capabilities and the current task. To date, developing such interfaces has tended to be very labor-intensive, with a lot of application-specific code. The goal of FormsTalk is to reduce the amount of application-specific code, so that most of the labor for a new application is involved in authoring the form content, and deployment on different platforms (e.g., PCs, phones, kiosks) requires a minimum of additional effort.

Technical Discussion: FormsTalk is built on top of DiamondTalk (see project), which is an application-independent Java architecture for building conversational, multimodal spoken-language interfaces. DiamondTalk allows us to easily substitute different speech recognition and generation engines (e.g., from different vendors), as technologies and applications change. FormsTalk also uses Collagen (see project) as its dialogue manager, which provides its mixed-initiative capabilities We currently access speech recognition and generation components from a web browser using a Java plugin; in the future we may use VXML or a similar protocol to access speech components directly from inside the browser. A key part of our work is a careful generic interaction design for multimodal, mixed-initiative form-filling, including the optional use of a telephone handset, touch screen, and other features.

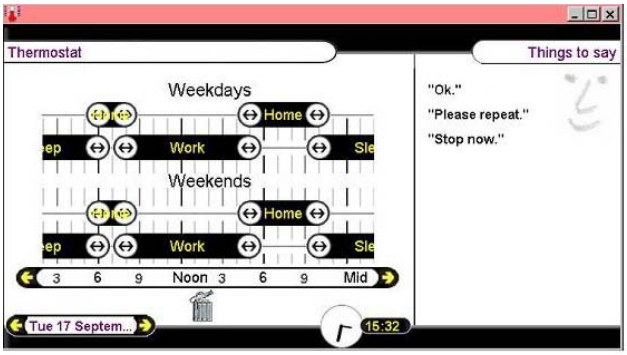
Collaboration: FormsTalk is the user-interface portion of the "Broadband WebServices" project led by the System Technology Dept. of Sentansoken. We are also members of the W3C Multimodal Interaction and W3C Voice Browser working groups.

Future Direction: Over the next year, our main focus is on how to gracefully recover from speech recognition and user errors. We also expect to discover ways to improve the basic architecture of FormsTalk as we apply it to new demonstration applications.

Contact: David Wong
http://www.merl.com/projects/FormsTalk/

Lab: MERL Technology Lab
Project Type: Research

Voice Programming for Home Products



Voice programming is a new approach to making it easy for people set up, customize, and operate the increasingly complex devices found in modern homes. The basic idea is for the device to provide interactive spoken guidance about how to use it. We have developed generic software for building such interfaces and have demonstrated it by building prototypes for a simulated personal video recorder and a simulated programmable home thermostat.

Background and Objectives: Digital processors are becoming a commonplace component of home products for entertainment, security, and housekeeping, leading to a veritable explosion of features and flexibility. The dark side of this trend, however, is that user interface technology has not kept pace. It is a now commonly observed fact that the overwhelming majority of the features of most new digitally-based devices are not used by the overwhelming majority customers. Our objective is to address this problem by taking advantage of recent advances in collaborative agents (see COLLAGEN project), intelligent tutoring (see Intelligent Agents for Operator Training and Task Guidance project) and spoken-language understanding (see DiamondTalk project).

Technical Discussion: Our most recent demonstration involves a home thermostat, in which the temperature of each room of a house is separately programmed for home, work, and sleep periods on weekdays and weekends. Similar features and issues arise in underground sprinkler control systems, personal video recorders, and home security systems. COLLAGEN provides us with a powerful and very general architecture for conversational, multimodal interaction. Another key element of our approach is the "Things to Say" list (see right side of figure above), which is dynamically generated from the task model and the current dialogue state and may include variables, such as "Today is <day of the week>." The Things to Say list addresses a key difficulty in spoken-language interfaces, namely how does the user know what the system will understand? Using this list also greatly improves the accuracy of speech recognition.

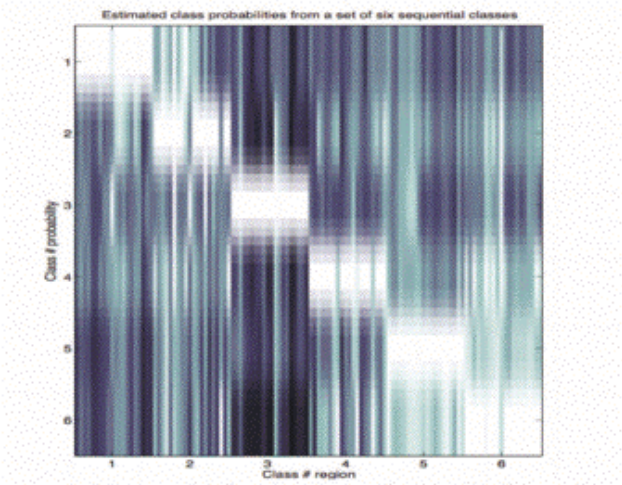
Collaboration: This work is in collaboration with the Intelligent Products group at the Delft University of Technology, Subfaculty of Industrial Design Engineering.

Future Direction: Adopting a consistent interface based on voice programming across a broad line of home products could yield benefits in brand identification and customer loyalty, as well as a reduction in product development costs. We also expect the mechanisms underlying the Things to Say list to migrate into the DiamondTalk.

Contact: Neal Lesh, Charles Rich, Candace Sidner
http://www.merl.com/projects/VoiceProgramming/

Lab: MERL Research Lab
Project Type: Research

Sound Recognition



Recognition results for six classes from a database of television program sounds. Perfect results would be depicted as six white squares along the diagonal. We achieve a recognition accuracy of 92%.

Sound recognition is a project with the goal to enable machines to listen and understand their surrounding sonic environment. We anticipate that this technology will provide the basis for various types of audio sensing which can find applications in surveillance, factory automation, entertainment media analysis, and other domains where audio feedback can be critical. Parts of our technology have also been included in the MPEG-7 standard for indexing and extraction of audio content.

Background and Objectives: Sound recognition has long been a neglected field of research. So far it has been dominated by speech recognition work that has found limited success for anything other than speech sounds. Our work is focused on generalized sound recognition, a framework equally capable of working with any type of sound regardless of its nature and its recording conditions. We have developed a system that can work in either of two modes. It can learn multiple sound classes by example and then recognize them in a real-time setting; thus providing a basis for applications such as audio surveillance. Or it can learn a single sound class and then provide a running estimate of how much the audio input deviates from it; a process that can be invaluable for machine fault detection.

Technical Discussion: The sound recognition project is an umbrella for various computational frameworks. Their crucial common step is a dimension-reducing preprocessing stage of input sound magnitude spectra. This step is used to provide an optimal representation for the subsequent classification, and is implemented by either Principal or Independent Components Analysis, or Reduced-rank Linear Discriminant Analysis transforms. The learning and classification steps are done using Hidden Markov Models or Gaussian Mixture Models. In special cases where the data set is dense we can obtain comparable performance using a much more efficient system using k-means.

Collaboration: Sentansoken and Sanden.

Future Direction: Generalized time-series recognition for ultra/infra-sound and vibrations analysis. We are also expecting to extend this work for joint audio-visual analysis, where camera input can provide additional information for classification.

Contact: Paris Smaragdis, Bret Harsham
http://www.merl.com/projects/SoundRecognition/

Lab: MERL Research Lab
Project Type: Research

User Interface Software

User interface is the most visible part of computers, computational applications, and devices. The acceptance of a new technology by the masses is largely dependent on how the user perceives the usefulness of the technology – a perception through the interaction with the affordances offered by the user interface. User interface research encompasses the full spectrum from user interface design, to the study and development of the scientific underpinnings of algorithms, to the understanding of the application and user experience with a given technology

At MERL, we are exploring new user interface paradigms, interaction metaphors and modalities, and interaction techniques that go beyond today's conventional single-user, mouse- and keyboard-based desktop user interface. In particular, our research investigates three areas of newly emerging user interfaces – (1) shared multi-user interfaces for human-human collaboration, (2) interactive interfaces for multimedia content browsing and retrieval, and (3) human-computer collaboration interfaces.

MERL plays a leading role in the international research community in the area of shared multi-user interfaces on direct-manipulation touch surfaces. DiamondTouch Applications, DiamondSpin, and UbiTable are three projects that enable co-present and collaborative group work on the tabletop. Personal Digital Historian, Timetunnel Interface for Video Browsing, and Content Management System provide innovative and interactive methods for the management and browsing of digital photos and video data. Human Robot Interaction for Hosting Activities, Intelligent Agents for Operator Training and Task Guidance, and Human-Guided Search are projects that explore human-computer and robot-human interactions.

Project Descriptions

DiamondTouch Applications

DiamondTouch is a novel multi-user input device. (See DiamondTouch Hardware and DiamondTouch SDK for more details.) DT applications provide a shared focus of attention for collaborating users. Possible applications include command-and-control rooms, business or technical meetings, and a variety of casual applications (e.g., musical instrument control, home coffee table, games, etc).

Background and Objectives: To date, software applications are intended for single users and designed to utilize traditional input devices (a single keyboard and mouse). In contrast, DiamondTouch is well-suited for shared-display groupware applications; it enables many people to simultaneously interact with the surface without interfering with each other. Our multi-user applications are developed with the DiamondTouch SDK.

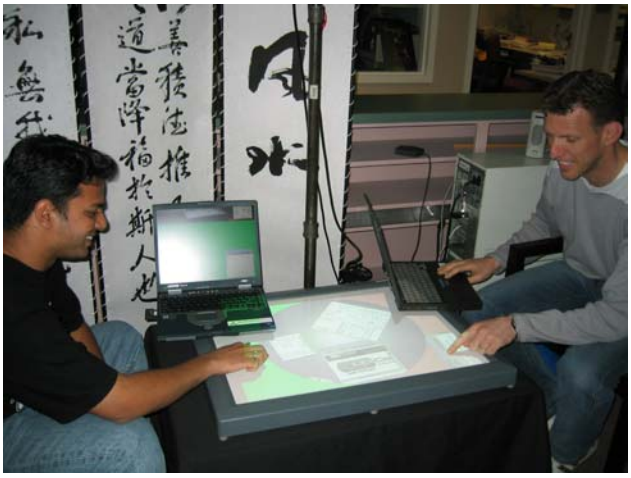
Technical Discussion: DiamondTouch applications come in a variety of styles. Some exploit DT's multi-user nature, while others may utilize its multi-touch capability -- each user may touch the unit in more than one place. Furthermore, DiamondTouch's ability to provide identification information for each touch (which users are touching where) is critical in many applications. DTMap, a multi-user map application, highlights DiamondTouch's ability to support input from multiple, simultaneous users and exploits its ability to identify the owner of each touch. In this application the display contains a satellite map image. Different views can be overlaid onto the map through a series of magic lenses. More generally, any layered information can be displayed in this application. DTMouse is an application to provide mouse emulation capabilities for DiamondTouch. Our mouse emulator works with traditional software, but it expects only one user to be using the mouse at any given time. We are experimenting to determine how best to implement a fully-functioning mouse with DiamondTouch. We have also implemented a number of entertainment applications to showcase DiamondTouch.

Collaboration: DiamondTouch is a joint project of MERL's Technology and Research Laboratories. MERL is collaborating with MELCO partners from Kamaden, Johosoken and Sentansoken. We also have active collaborations with universities who will explore DiamondTouch as a collaborative input technology.

Future Direction: our MELCO collaborators will use some of the applications we develop. We will work with them to refine the applications, and develop new functionality. We also plan to evaluate our design decisions through more formal user testing, and investigate the use of DiamondTouch for remote collaboration.

Contact: Kathy Ryall **Lab:** MERL Technology Lab
http://www.merl.com/projects/DTApplications/ **Project Type:** Advanced Development

UbiTable



Despite the mobility enabled by the plethora of technological tools such as laptops, PDA and cell phones, horizontal flat surfaces are still extensively used and much preferred for on-the-move face-to-face collaboration. The UbiTable project examines the design space of tabletops used as scrap displays. Scrap displays support kiosk-style walk-up interaction for impromptu face-to-face meetings without prior preparation. UbiTable affords easy and efficient set up of shared horizontal workspaces such that users can conveniently move contents among their laptops and the tabletop.

Background and Objectives: The support for serendipitous meetings will play a critical role in future ubiquitous computing spaces. The goal of UbiTable is to offer the affordances of a physical table, providing the flexibility by allowing users to layout shared digital documents with desired orientation and position. At the same time UbiTable augments traditional paper-based interactions by providing a flexible gradient or shades of sharing semantics. UbiTable addresses visual accessibility vs. electronic accessibility of documents, an issue which is critical to ubiquitous environments.

Technical Discussion: UbiTable separates the notion of visibility from privacy, and the notion of modifiability from ownership of documents on a shared horizontal surface. Based on observations of people collaborating around a table with physical artifacts (such as paper documents), UbiTable provides a gradient of Private, Personal and Public sharing models to afford well understood social protocols. Private documents are neither visible nor electronically accessible by others. Personal documents are semi-private documents; they are visible but not accessible by others. Public documents are both visible and electronically accessible by others. Electronic accessibility is defined by readability and modifiability, but not shared ownership. The owner of a document maintains the explicit control of the electronic distribution of the document.

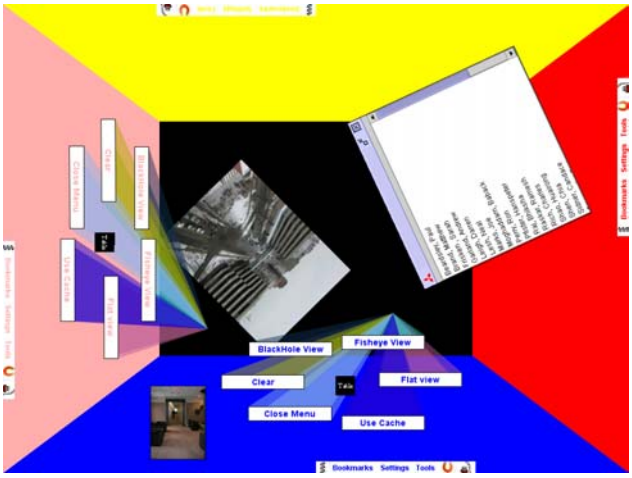
Collaboration: UbiTable is a collaborative research project between MERL Research and MERL Technology labs.

Future Direction: UbiTable is a new research project at MERL. We are undertaking a concurrent design and prototype process: (1) carrying out user studies of physical document sharing tasks in a face-to-face setting in order to understand the design space, and (2) constructing an initial prototype of the tabletop and laptop user interface and interactions. Meanwhile, the initial prototype will be used to study user preference, sharing semantics and visual layout strategies.

Contact: Chia Shen, Kathy Ryall
http://www.merl.com/projects/UbiTable/

Lab: MERL Research Lab
Project Type: Research

DiamondSpin



See Color Figure 7

DiamondSpin is an interactive tabletop environment that preserves the simplicity and informality of around-the-table interaction, while at the same time provides a rich set of novel UI functions for interactive and collaborative document browsing, visualization, manipulation and navigation by small groups of people.

Background and Objectives: Tables are a familiar piece of furniture commonly found in homes, offices, cafΘs, design centers, show rooms, waiting areas, and entertainment centers. Tables provide a familiar, casual and convenient physical setting for people to meet, chat, look over documents, and carry out tasks that require face-to-face collaboration. Digital documents, on the other hand, are commonly used only on desktop computers and handheld devices, due to a lack of a physical media that contain the necessary computational support for face-to-face around the table applications. Our objectives are to research and study input techniques, interaction techniques and visualization techniques which can enable tabletop applications where small groups of people collaborate.

Technical Discussion: User interfaces that are on a horizontal tabletop, and support simultaneous multi-user input actions have four characteristics: (1) they must handle the arbitrary location and orientation of documents on the table in real-time, (2) they must manage both rotation sensitive and rotation insensitive user interface components, (3) they must support the manipulation and display of large quantities of pixels from many overlapping documents, and (4) they must manage multi-user collaborative, and potentially conflicting, activities. We are experimenting with the construction of a real-time Cartesian to Polar coordinate transformation system to afford continuous orientation and arbitrary viewing angles among multiple people. The first version of DiamondSpin tool kit enables the construction of (a) arbitrary 2D geometric shapes of digital tabletops including rectangular, octagonal and circular interfaces, (b) multiple digital virtual tabletops, and (c) multiple digital work areas within the same display space. DiamondSpin also provides a set of tabletop document visualization methods.

Collaboration: We are currently distributing the first version of DiamondSpin Java Toolkit to universities with a free license agreement. Currently DiamondSpin license holders include academic partners from Europe, Japan and the U.S. We are developing a variety of applications using DiamondSpin at MERL and with our university partners.

Future Direction: Our next step is first to expand the multi-user functionality of DiamondSpin. Bimanual and compound gestures and multi-user conflict resolution policies are two of the key research areas we are pursuing.

Contact: Chia Shen
Lab: MERL Research Lab
http://www.merl.com/projects/diamondspin/
Project Type: Research

Personal Digital Historian (PDH)

PDH is a new digital content user interface and management system. Unlike conventional desktop user interfaces, PDH is intended for multi-user collaborative applications on single display groupware. PDH enables casual, interactive and exploratory retrieval, interaction with and visualization of digital contents. PDH is built on top of our DiamondSpin circular tabletop environment. Our current project includes research in the areas of content annotation, retrieval and presentation, visualization of and user interaction with images, audio, video and data, as well as the study of how people collaboratively use the single display interface.

Background and Objectives: As part of people's daily life at work, on the go and at home, their computers, PDAs and digital cameras generate larger and larger amounts of digital contents. However, technologies that allow people to easily utilize this digital data in a face-to-face conversational or group setting are lagging far behind. Applications are limited by the user interface potentials of current desktop computers and handheld devices. The objective of the PDH project is to take a step beyond.

Technical Discussion: Creating a new type of interface requires addressing many issues. One of our primary focuses is on developing content organization and retrieval methods that are easy and understandable for the users, and can be used without distracting them from their conversation. Rather than the folder & file mechanisms used by conventional document systems, PDH organizes the contents along the four W's of storytelling (Who, When, Where, and What) and allows users to design new contexts for organizing their structures. A second issue we have focused on is affording casual and exploratory interaction with data by combining a multiplicity of user interaction mechanisms including in-place query and in-place pop-up menus, direct manipulation, natural visual query formulation with minimal menu-driven interaction and freeform strokes. Finally, in order to support the multi-threaded and non-linear progression of group conversation, PDH provides tools to help people navigate a conversation as well as their content.

Future Direction: We are developing semantics and mechanisms for simultaneous multi-user input and developing new application scenarios for PDH along with continued user studies.

Contact: Chia Shen, Neal Lesh
http://www.merl.com/projects/PDH/

Lab: MERL Research Lab
Project Type: Research

Timetunnel Interface for Video Browsing

TimeTunnel is a captivating interactive technology for browsing video content. It takes advantage of innate human capabilities to track objects moving towards or away from the viewer

Background and Objectives: Mitsubishi Electric manufactures consumer audio-visual devices such as televisions and digital video recorders and also creates software that records and manages video content. These businesses have an interest in creating a new generation of human- the computer interfaces that boost effectiveness and attractiveness of their products and differentiate them from the competition. The Timetunnel Interface for Video Browsing has immediate visceral appeal and can provide a distinctive feature to video hardware and software products from MELCO.

From a research perspective, the objective of this project is to advance the state of the art in human-computer interfaces to multimedia content. The current generation of multimedia interfaces is based on the antiquated constraints of analog content and storage media.

Technical Discussion: The Timetunnel Interface for Video Browsing is in the family of interaction presentation techniques called Rapid Serial Visual Presentation. Rapid presentation of visual information has not largely been exploited in human-computer interfaces for TVs and other consumer electronics devices. However, there are indications that people are drawn to such interfaces because our visual system is capable of very rapid processing of visual information.

Using key frames extracted from a set of broadcast channels or else from analyzed scenes in recorded video, the Timetunnel Interface for Video Browsing presents the frames in a 3D space and offers the user controls for moving the frames towards or away from the view position. Users can adjust the speed and direction of movement. As they "surf," they can fine-tune the speed and move the sequence forwards or backwards. If they stop on a frame, the application either tunes to that channel or starts playing the video from that frame position.

We conducted an experiment designed to test a Timetunnel interface for video browsing on a television using a remote control. We found that subjects were significantly more accurate at fast-forwarding to a desired location in a recorded video clip while using Timetunnel than while using the traditional method of fast-forwarding, $t(59)= -12.93$, $p < 1E-18$.

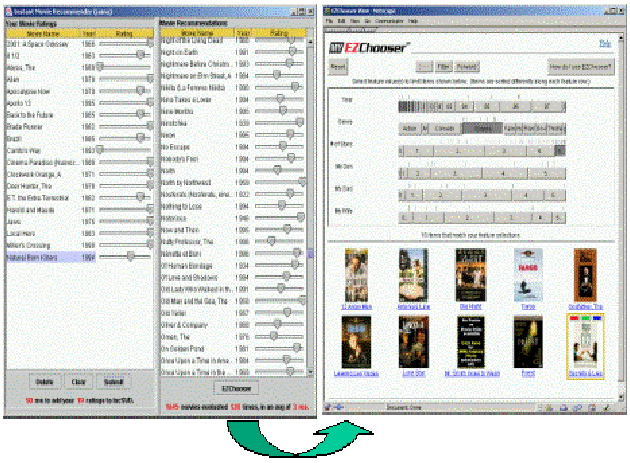
Collaboration: IDken, Sentansoken-Kyoto.

Future Direction: We expect to work with business units interested in incorporating the technology into their digital video product lines.

Contact: Kent Wittenburg, Alan Esenther
http://www.merl.com/projects/timetunnel/

Lab: MERL Technology Lab
Project Type: Research

Multi-Parametric Visualization



Movie Recommender with Data Mining and Visualization Techniques

Multi-Parametric Visualization is a set of methods and tools for visualization and querying of multidimensional information. It is designed to be easy to use and quick to yield insights into the data.

Background and Objectives: Information visualization tools, while an industry unto itself, have by and large remained difficult for non-specialists to use. Particularly when the information to be visualized comes in the form of a relational table, it is rare for an executive or a designer to be able to quickly and flexibly explore the data and derive insight to aid decision-making. The Multi-Parametric Visualization project at MERL has the objective of developing new techniques and tools that can improve engineering design, operations, and business decision-making at Mitsubishi Electric. In the longer run, we expect that these tools will also be included in service offerings to external customers in the power and utilities business among others.

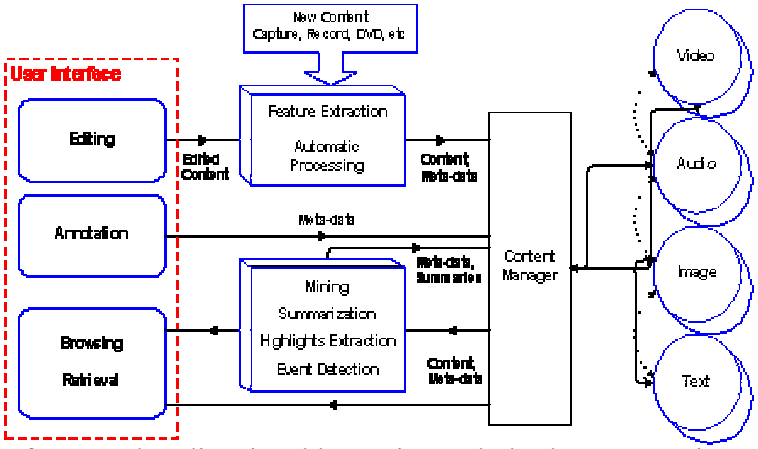
Technical Discussion: The multi-parametric visualization project began with the development of parallel visualizations called bargrams. There is typically one bargram for each column of information in a table. Bargrams are similar to histograms, revealing distribution of counts across an ordered set of "bins" or categories. However, the bargram saves visual real estate and offers additional visualization features by "tipping over" the histogram to create a simple one-dimensional visualization for each attribute. Users can quickly create queries and preview their results by selecting subsets of attribute values. For research purposes, MERL has licensed software from Verizon, created by members of the current staff at MTL when they were at Verizon Laboratories. This has allowed the project to quickly prototype new application domains and creates requirements for future work. MERL has also developed a new implementation of MPV for the Human-Guided Antenna Design project that allows designers to quickly evaluate and winnow down a large set of generated antenna designs based on their collective features. More recently, we have started to conduct research on other types of visualization techniques beyond the bargram concept to support time-series analysis.

Collaboration: Real-time Recommender Systems and Human-Guided Antenna Design projects at MERL and communications and power systems business units at Mitsubishi Electric.

Future Direction: The project expects to develop a comprehensive set of middleware that will allow for rapid development of multi-parametric visualization components for a variety of applications.

Contact: David Wong, Frederick J. Igo, Jr., Kent Wittenburg **Lab:** MERL Technology Lab
http://www.merl.com/projects/mpv/ **Project Type:** Research

Content Management System



The amount of available information in the form of text, audio, video, speech and images grows exponentially every day. While storage capacities have kept pace with this increase, such growth in information poses the challenge of managing it so that the desired information is retrieved quickly and conveniently. In our work so far, we have emphasized browsing of audio-visual content. In this project, we emphasize the seamless combination of text and audio-visual browsing to help the user navigate large volumes of heterogeneous content. In the figure at left, we illustrate our proposed framework. It integrates our previously developed video analysis techniques as well as anticipates further developments such as combinations of representations of content semantics with audio-visual browsing.

Background and Objectives: While storage technology makes it possible to store extremely large volumes of heterogeneous data (audio, video, speech, text, images), it is a challenge to use it effectively. The challenge lies in combination of summarization and indexing techniques for each of the individual modes so as to enable the user to get the desired content in any desired mode. In this project, we are developing content management techniques for electronic learning applications such as production and browsing systems for training and educational video.

Technical Discussion: Our system relies on fast computation so as to ensure rapid content preparation as well as browsing. There are broadly two technical challenges. The first is to integrate and enhance existing indexing and summarization techniques for each of the modes, in a single framework. The second is to devise content representation techniques that enable seamless merging of content semantics with audio-visual browsing, so as to enable multi-modal hyper-navigation of the content. We have obtained promising results with application of our video summarization techniques to training video obtained from MELCO, Japan, as well as other training video.

Collaboration: Johosoken, Sentansoken.

Future Direction: Our current research focus is on combining representations of content semantics with purely audio-visual indexing and summarization. In addition to our current set of technologies such as Video Indexing and Summarization, Video Object Segmentation and Tracking, we are also working on new techniques such as video mining.

Contact: Ajay Divakaran
http://www.merl.com/projects/ContentMgmt/

Lab: MERL Technology Lab
Project Type: Research

Human-Robot Interaction for Hosting Activities

We developed a stationary robot that collaborates with a person on a hosting task. It demonstrates a MERL technology called iGlassware to a visitor and engages the visitor in a discussion of their interests in MERL.

Background and Objectives: Our objective is to develop collaborative robots that not only speak to humans, but use physical gestures and movement to interact with them, thereby "engaging" the human in the interaction. One situation in which this kind of interaction is particularly appropriate we call "hosting." In a hosting activity, the "host" (which may be a human or a robot), provides the "visitor" (typically a human) with various kinds of information, entertainment, education, etc., related to their shared physical environment. Examples of human roles that include hosting activities are docents, sales clerks, receptionists, and real-estate agents.

Technical Discussion: We have developed a prototype hosting robot, embodied as a penguin. The robot demonstrates the iGlassware technology in collaboration with a human visitor. The robot invites the visitor to participate in the demo, convinces the reluctant visitor to participate, explains the technology and induces the visitor to use the technology as they interact. The robot also explains the purpose of iGlassware to an interested visitor. During the interaction the robot gestures appropriately at demonstration objects, gazes at the visitor during the visitor's turn, and gazes at visitors, onlookers and demonstration objects during its own turn. The robot also observes the visitor's gestures and determines the visitor's interest in the interaction. Our robot makes use of face detection, object detection and sound localization algorithms to find and track its interlocutor, the onlookers and the objects in the environment. It uses rules for interaction engagement to determine its gestural and spoken response during a demonstration and to assess the human collaborator's utterances and interest in the conversation. It uses the COLLAGEN middleware for collaborative conversation to manage the conversation and commercial speech recognition software to interpret spoken utterances.

Collaboration: This work has been done in collaboration with Chris Lee, MERL Technology Lab, Neal Lesh and Chuck Rich of MERL Research Lab.

Future Direction: We will evaluate our robot with human users to judge the effects and usefulness of engagement behaviors during hosting interactions. We will also use new learning techniques so that the robot can better estimate human intentions reflected in human gestures.

Contact: Candace Sidner
http://www.merl.com/projects/hosting/

Lab: MERL Research Lab
Project Type: Research

Intelligent Agents for Operator Training and Task Guidance

Intelligent software agents can help the operators of complex industrial equipment both with training and in the performance of their tasks. During training, the operators practice their tasks in a simulated environment. The intelligent agent guides operators through the steps of complex procedures, giving positive and negative feedback on their actions, and adapting to their skill level. We have developed generic software for building such software agents and have demonstrated it by building an agent for training operators of a simulated gas turbine engine.

Background and Objectives: This work extends the COLLAGEN middleware for building collaborative agents. Our goals are to enhance existing training systems by adding more sophisticated tutorial strategies and also to build new training and task guidance systems. Potential application areas include supervisory control and data acquisition (SCADA) systems, such as industrial plant control centers, equipment maintenance tasks, such as elevator repair, and the use of complex software interfaces, such as computer-aided design tools.

Technical Discussion: Our approach to operator training is based on "learning by doing." The software agent guides the operator through a sequence of example scenarios that incrementally expose the operator to the full complexity of the task to be learned. During the training process, the agent maintains a model of the operator's proficiency in each part of the task, so it can appropriately introduce new subtasks as well as give the operator the opportunity to practice previously taught knowledge.

Using COLLAGEN as our implementation base gives the training system developer two major advantages. First, we have an application-independent architecture in which pedagogical strategies are encoded in an application-independent manner; this allows the developer to reuse a large amount of code when creating a new training agent. Second, the same task model created to support a training agent can be reused to produce an agent which will act as an advisor or intelligent assistant to an operator during the actual performance of his job, if desired.

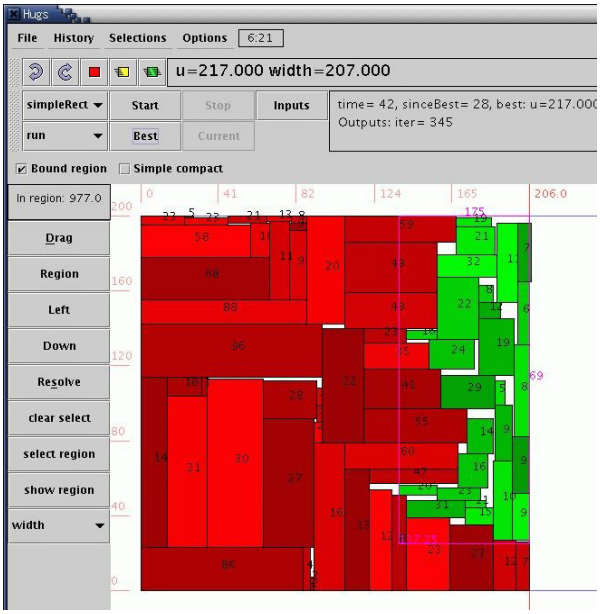
Collaboration: We are currently working very closely with Mitsubishi Electric's Advanced Technology R&D Center, and Energy and Industrial Systems Center. During this past year, we produced a prototype agent to train thermal power plant operators using a software simulator, which will be demonstrated in the customer showroom at the Energy and Industrial Systems Center. We are also collaborating with the University of Southern California, Information Sciences Institute and the MITRE Corporation on embedded training applications in general.

Future Direction: We are planning to expand the range of pedagogical strategies used by our training agents, explore the use of pedagogical agents to assist in human-human e-learning, and to develop tools and special-purpose task model languages for rapidly authoring new training and task guidance applications.

Contact: Neal Lesh, Charles Rich, David Wong
http://www.merl.com/projects/training/

Lab: MERL Research Lab
Project Type: Advanced Development

Human-Guided Search for Packing



We designed both automatic and interactive systems for tightly packing a given set of rectangles, without overlap, into a larger rectangle of fixed width. This is important for industrial problems in which a large number of variable-sized rectangles must be cut from a sheet of material, such as metal, glass, or wood. The goal is to minimize waste by packing the rectangles as tightly as possible. Our automatic method significantly improves on the best automatic method we found in the literature. Furthermore, we found that people can reason about packing problems extremely well. Our interactive system allows people to apply packing algorithms to quickly produce much tighter packings than the packing algorithms can produce without guidance.

See Color Figure 8

Background and Objectives: We have worked on what is referred to in the literature as the two-dimensional rectangular strip packing problem. The goal is to pack a given set of rectangles as tightly as possible, to minimize wasted space. This problem appears unamenable to standard local search techniques, such as simulated annealing or genetic algorithms. A standard simple heuristic, Bottom-Left-Decreasing (BLD), has been shown to handily beat these more sophisticated search techniques.

Technical Discussion: We have developed and demonstrated the effectiveness of BLD*, a stochastic search variation of the best prior heuristic, BLD. While BLD places the rectangles in decreasing order of height, width, area, and perimeter, BLD* successively tries random orderings, chosen from a distribution determined by their Kendall-tau distance from one of these fixed orderings. Our experiments on benchmark and randomly generated problems show that BLD* produces significantly better packings than BLD after only 1 minute of computation.

Furthermore, we observe that people seem able to reason about packing problems extremely well. We incorporate our new algorithms in an interactive system that combines the advantages of computer speed and human reasoning. Using the interactive system, we are able to quickly produce significantly better solutions than previous methods on benchmark problems.

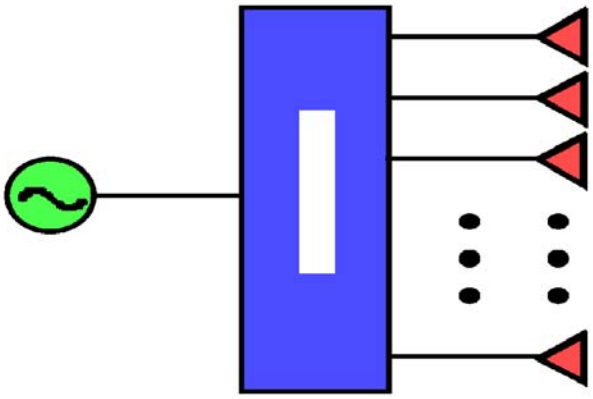
Collaboration: We are collaborating with Harvard University.

Future Direction: We hope to apply this technology to packing problems in MELCO, such as those that arise in laser cutting. Additionally, we can use our interactive system as an interface to other packing algorithms.

Contact: Neal Lesh
http://www.merl.com/projects/packing/

Lab: MERL Research Lab
Project Type: Research

Human-Guided Antenna Design



Optimization-based approaches to antenna design have enjoyed limited success. The task is often computationally intractable. Moreover, it is also often difficult to capture all relevant design issues and trade-offs in a single mathematical objective function. Therefore, human experts typically specify and refine antenna designs by hand, using computers only to evaluate their candidate designs by simulation. In this project we propose a middle ground between this traditional approach and fully automatic optimization - a human-guided interactive system.

Background and Objectives: The idea of using computer-based optimization for design tasks has been applied to many problems, including antenna design. However, this idea does not always work well: the optimization problems are often intractable; and it often proves impossible to consider all relevant design criteria in the optimization process. In this project we propose that the computer be used differently. Instead of having the computer search for a single optimal design, we program it to intelligently sample the large space of possible antenna designs, subject to user-supplied constraints. The task of choosing a final design from the computer-generated sampling is left to the human user, who can apply experience and judgment to recognize and then refine the most useful antenna design.

Technical Discussion: At the heart of our approach, the computer generates a sample set of candidates (called a population) of possible antenna designs and presents them to the user. The person's role is to help guide the computer in generating a desirable population, and ultimately to select an appropriate antenna. The two key components in our system are dispersion and visualization. Dispersion is the process by which we generate a representative sample of designs; it is an intelligent sampling process. A key requirement for dispersion is a mathematical function that quantifies the difference between two antenna designs. This difference metric is usually based on the performance characteristics of an antenna. Visualization is the process by which a person can examine the current population to examine trade-offs in the design space (e.g., cost and gain) and to constrain future searches. We are currently using MPV (multi-parametric visualization) as our visualization component. The design process is an iterative one, alternating between dispersion and visualization.

Collaboration: This project is a joint effort between MERL Technology Lab and MERL Research Lab in collaboration with Johosoken and with sponsorship from Denshihon.

Future Direction: In the first phase of this project we completed a proof-of-concept implementation of an interactive system for antenna design. In the upcoming year we will be focusing our efforts on visualizing and browsing large datasets of phased-antenna arrays.

Contact: Kathy Ryall, Darren Leigh
http://www.merl.com/projects/antenna/

Lab: MERL Technology Lab
Project Type: Research

Wireless Communications and Networking

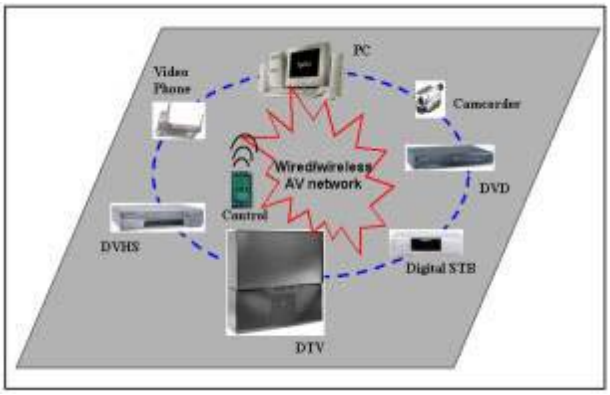
Communications and networking are pervasive in today's society. The explosion in the Internet and in communication technology has revolutionized the way people access information, ranging from e-commerce to "virtual government". Another aspect of this revolution is the untethering of voice and data communications, giving people the ability to talk and access information anytime, anywhere. These changes in paradigms will continue with the development of more and more high-speed, information-technology-enabled devices. From advanced wireless multimedia systems to simple integrated home networking, communications and networking technologies will be at the center of this continuing revolution.

At MERL, we aim to create new concepts and develop enabling technologies in the area of digital communications and networking. As an R&D team, we collaborate with other research organizations, publish our accomplishments, participate in and contribute to related industrial standardization activities, and measure our performance by the impact we have on both Mitsubishi Electric Corporation and the world. Our current focus is on broadband wireless communications, pervasive ad hoc networks, and integrated home networking.

For next generation broadband mobile communications, different technologies will be converged and integrated to provide high speed and seamless roaming access. MERL is an active participant in major industry standards, including 3GPP, and continues developing advanced technologies, such as MIMO (Multi-Input-Multi-Output) for HSDPA (High Speed Downlink Packet Access) and future systems. In the area of short-range wireless communications, MERL is an active member and contributor in the emerging technologies, especially ultra-wideband (UWB), ZigBee, sensor networks, and network security. We are also aggressively working on high throughput WLAN for home networking, wireless digital TV, simple control protocol (SCP) for power line communication, and integrated home networking.

Project Descriptions

Broadband A/V Home Networking



With the advance of communications and networking technologies, the networked home market has taken off for years. A networked home allows the people to have the simultaneous Internet access from different parts of the home, and to share the devices and contents within the home. This project focuses on audio/video home networking technologies that will allow the redistribution and sharing of A/V contents through a unified wired/wireless home network. The solution will lead to a new networked A/V consumer electronics business.

Background and Objectives: The new upcoming IEEE802.11e standard that will support the applications with QoS requirements is paving a way for customers to have a multimedia wireless home network. With the existing wired infrastructure, the future home network will be a mixed wired/wireless network. The main objective of the project is to develop an interworking protocol that allows the smooth and reliable connections between wired (i.e. 1394, Ethernet, and digital cable) clusters and wireless (i.e. 802.11) clusters.

Technical Discussion: Unlike the original 802.11 standard, 802.11e supports the real-time applications such as voice, video with guaranteed QoS, i.e. guaranteed bandwidth, bounded delay and so on. This makes wireless multimedia home networking a reality. The major technical hurdle here is to develop an interworking protocol between 802.11e and other existing wired standards such as 1394, Ethernet. In the project, we design the PAL (protocol adaptation layer) between 802.11e and 1394, such that 1394 data can be smoothly transmitted over 802.11e without loss of quality and vice versa. We also research and implement the mapping algorithms between IP DiffServ and 802.11e to optimize the network's performance.

Collaboration: This project is performed in collaboration with Sentansoken.

Future Direction: The future work will primarily focus on the PAL design between 1394 and 802.11e, and protocol mappings between 802.11e and IP DiffServ.

Contact: Daqing Gu
Lab: MERL Technology Lab
http://www.merl.com/projects/BroadbandAV/
Project Type: Advanced Development

Wireless Digital TV

IEEE802.11 WLAN is considered to play an important role in future multimedia wireless home networks. So, the support of IEEE802.11 in digital TV and other home electronics becomes a necessity. In this project, we focus on A/V transmission over IEEE802.11 with QoS. For DTV applications, this enables DTV to be directly connected with Internet and other A/V equipments without wires. Furthermore, the future DTV can be built with the integrated 802.11 interface for universal wireless access in the multimedia home networking environments.

Background and Objectives: The new upcoming IEEE802.11e standard that will support the applications with QoS requirements is paving a way for customers to have a multimedia wireless home network. Meanwhile, IEEE802.11 working group is working on another standard called 802.11n, with a target of minimum throughput of at least 100 Mbps. This standard will allow a more reliable wireless link for high data rate A/V transmission. The main objective of the project is to develop a reliable 802.11 interface for DTV and other A/V equipments.

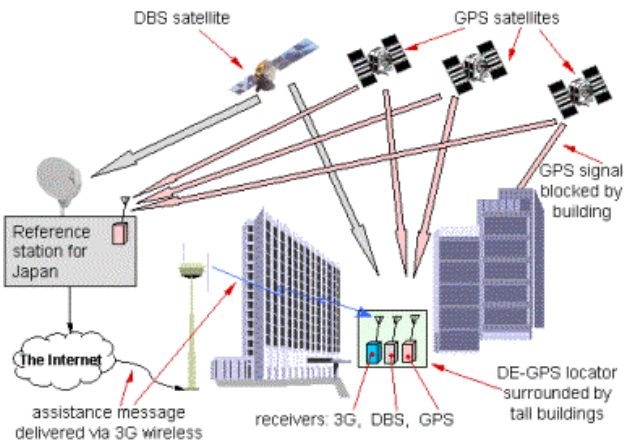
Technical Discussion: We currently use IEEE802.11a WLAN as a wireless carrier for A/V transmission. IEEE802.11a is a high-speed WLAN with data rate up to 54 Mbps, and is good enough for single MPEG stream transmission. The major technical hurdle here is to develop a bridge between 1394 and 802.11a devices such that 1394 data can be transmitted smoothly over 802.11a to a DTV for display. In our design, MPEG data is first packetized into 802.11a format, and then transmitted over an 802.11a. At the receiving end, the received 802.11a data is converted to1394 format for display. The PCs shown in the above figure are for demo purpose only. The final goal of the project will be DTV with an integrated 802.11 interface.

Collaboration: This project is performed in collaboration with MDEA.

Future Direction: For the future work, we will focus 802.11e/n standards for high throughput transmission with QoS.

Contact: Daqing Gu **Lab:** MERL Technology Lab
http://www.merl.com/projects/WirelessDigitalTV/ **Project Type:** Advanced Development

DE-GPS: DBS-Enhanced GPS



Using a conventional GPS receiver in dense urban areas, a large portion of the sky is frequently obstructed by buildings, making it impossible to "see" a sufficient number of GPS satellites. The DE-GPS technique allows a GPS receiver to use the signal from a DBS satellite when it is in a bad spot. To do so, the DE-GPS receiver must obtain assistance information from a remote server. This assistance information enables the DE-GPS receiver to calculate the range information from the non-GPS satellites for the location fix.

Background and Objectives: GPS is essential for car navigation systems. However, when the car is in an area where most of the sky is blocked by buildings, it is difficult for the GPS receiver to acquire four satellites for a complete location fix. If additional satellites such as DBS can be used for geo-location, the probability of the receiver acquiring four satellites will be greatly increased.

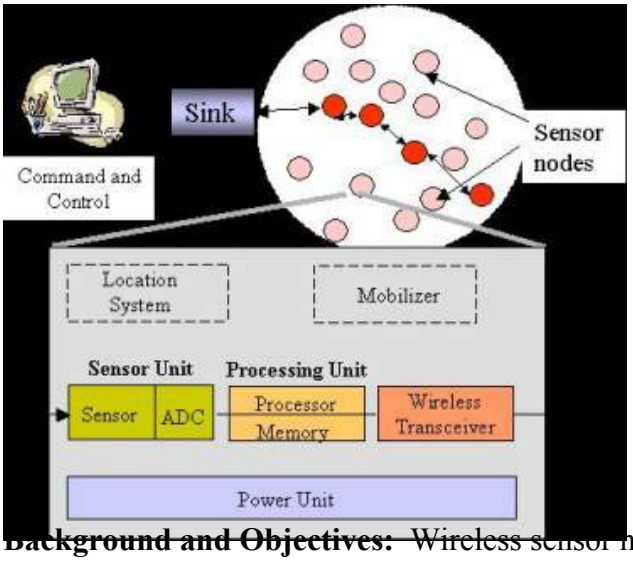
Technical Discussion: In dense cities, GPS receivers don't work well because tall buildings block GPS satellite signals. However, most urban areas are well covered by DBS satellites. DBS satellite signal can be used instead of GPS signal -- but assistance must be provided to receiver. Assistance is prepared by one reference station for all of Japan. Third Generation (3G) wireless service (e.g., FOMA) is used to send assistance from the reference station to the receiver. Additionally, the location fix can be more accurate than using only GPS because a DBS signal has a much wider signal bandwidth.

Future Direction: MERL is actively seeking a partner to develop this business opportunity.

Contact: George Fang
http://www.merl.com/projects/DE-GPS/

Lab: MERL Technology Lab
Project Type: Initial Investigation

Wireless Sensor Networks



Wireless sensor networks consists of a large number of densely deployed sensor nodes. These nodes incorporate wireless transceivers so that communication and networking are enabled. Additionally, the network possesses self-organizing capability so that little or no network setup is required. Ideally, individual nodes should be battery powered with a long lifetime and should cost very little. The applications for such networks are numerous and include: Inventory management, product quality monitoring, factory process monitoring, disaster area monitoring, biometrics monitoring, and surveillance.

Background and Objectives: Wireless sensor networks are similar to Ad-hoc networks in the sense that sensor networks borrow heavily on the self-organizing and routing technologies developed by the ad-hoc research community. However, a major design objective for sensor networks is reducing the cost of each node. For many applications the desired cost for a wirelessly enable device is less than one dollar.

Technical Discussion: MERL is currently investigating several technologies that are suitable for wireless sensor networks. MERL has been investigating network protocols that help conserve node battery power. Our technique incorporates a node's residual energy into the cost metric that is computed when determining what route to send packets on. By avoiding nodes that have little reserve battery power the overall lifetime of the network can be extended. This technique has been used to modify the Dynamic Source Routing (DSR) protocol and has shown promising results.

In addition to our work on network algorithms, MERL has extensive experience in Ultra-Wide Band (UWB) radio technology. This has considerable promise in that it can reduce the cost of radio communications. In UWB a series of short baseband pulses are used to communicate information, because communication occurs with baseband pulses no IF components are needed in the receiver, which reduces transceiver cost. Additionally, because of the signal's large bandwidth individual many multipath components are resolvable, this aids in UWB's ability to locate devices. MERL has developed IP in the area of low-cost low rate transceiver design.

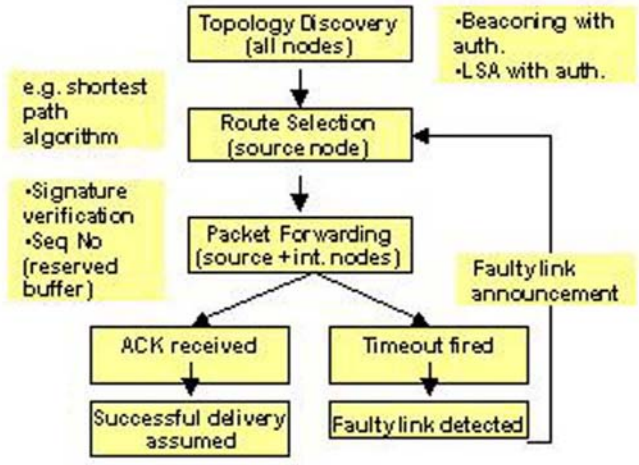
Collaboration: We have worked closely with Sentansoken/Sanden.

Future Direction: We will continue our research efforts in energy aware routing and UWB. As well as work towards incorporating our IP in relevant standards such as ZigBee and IEEE 802.15.

Contact: Philip Orlik
http://www.merl.com/projects/sensornet/

Lab: MERL Technology Lab
Project Type: Research

Denial of Service Attacks in Wireless Ad Hoc Networks



Denial of Service (DoS) is the degradation or prevention of legitimate use of network resources. The wireless ad hoc network is particularly vulnerable to DoS attacks due to its features of open medium, dynamic changing topology, cooperative algorithms, decentralization of the protocols, and lack of a clear line of defense is a growing problem in networks today. Many of the defense techniques developed on a fixed wired network are not applicable to this new mobile environment. How to thwart the DoS attacks differently and effectively and keep the vital security-sensitive ad hoc networks available for its intended use is essential.

Background and Objectives: This project concentrates on developing defense mechanism against certain types of DoS attacks in the Ad Hoc network environment.

Technical Discussion: The thrust of this project is to detect and defend against the black hole and misdirection DoS attacks by adding security features to Ad Hoc On Demand Distance Vector (AODV) routing algorithm via a protocol extension. Our extended protocol utilizes a packet forwarding mechanism with fault link detection, which is based on authentication of data and control packets. With authentication, the black hole attack can be thwarted since the malicious node does not hold the intended destination's secret key and thus cannot generate a valid signature on the destination acknowledgement (ACK). The combined use of ACKs, authentication, monotonically increasing non-wrapping sequence numbers and time-outs at the source and every intermediate router, in combination with fault announcement (FA), provide sufficient information to detect faulty links. Mathematical models are also established to analyze different types of DoS flooding attacks in transport layer. Traffic information obtained from the model can be used to develop methodologies and algorithms that can detect such attacks and defend a network against such assaults.

Collaboration: Certain parts of this project are done in collaboration with Princeton University and the Information Security Technology Department at Johosoken.

Future Direction: Reducing the computation and communication overhead of authentication is one primary goal in the near future. Message Authentication Codes (MACs) can be employed for this purpose, which provide authentication for pairs of routers. Therefore, the source of a data or control packet has to authenticate it to every downstream router separately. We also plan to enhance the client puzzle defense mechanism by using our flooding attack analysis models and developing the parameter specifications.

Contact: Johnas Cukier **Lab:** MERL Technology Lab
http://www.merl.com/projects/DenialServiceAttacks/ **Project Type:** Initial Investigation

MIMO Technology for 3GPP

HSDPA (high speed downlink packet access) is a data service in the 3rd generation W-CDMA (wideband CDMA) systems, which provides high-speed data transmission (up to 8-10 Mbps) in CDMA downlink to support multimedia services. In order to further increase the system capacity and enhance the quality of services, MIMO technology will be adopted in the next phase of the 3GPP standards, in which 20 Mbps of data transmission will be achieved. In this project, we developed the key technologies on fast link adaptation for HSDPA, with the focus on MIMO and transmit diversity, to enhance MELCO's IPRs and prepare the essential technologies for 4th generation wireless communication.

Background and Objectives: The objective of this project is to develop the key technologies on fast link adaptation for HSDPA (with focus on MIMO/STC) and make contributions to 3GPP standards to enhance MELCO's IPR portfolio on 3G systems. In 3GPP standards, Release 4 specifications provide efficient IP support enabling provision of services through an all-IP core network and Release 5 specifications focus on HSDPA to provide data rates up to approximately 10 Mbps to support packet-based multimedia services. MIMO systems are the work item in Release 6 specifications, which will support even higher data transmission rates up to 20 Mbps.

Technical Discussion: HSDPA is a packet-based data service in W-CDMA downlink with data transmission up to approximately 10 Mbps over a 5MHz bandwidth. This system is evolved from and backward compatible with Release 99 WCDMA systems, which is basically a voice service. Adaptive modulation and coding is used for data transmission to support multiple rates and multiple types of multimedia services. In order to reach higher peak rate (20 Mbps), MIMO technology is used, in which multiple antennas are implemented at both basestation and mobile UE terminals. At the transmitter, the information bits are divided into several bit streams and transmitted through different antennas. The transmitted information are recovered from the received signals at multiple receive antennas by using an advanced receiver. Due to the high data rate transmission, the trade off between complexity and system performance becomes an important issue, especially for the UE designs. The research in MERL focused on the MIMO transceiver design, system performance evaluation, advanced receiver design and extension of transmit diversity schemes with more than 2 transmit antennas.

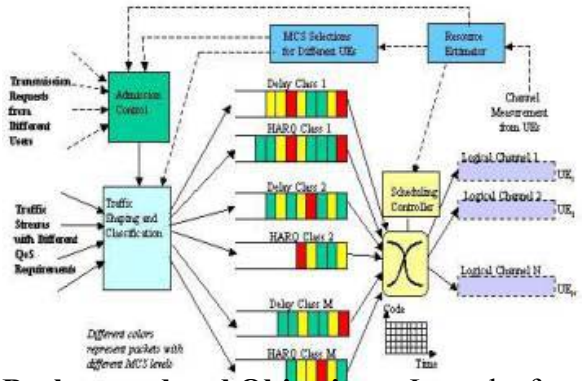
Collaboration: MERL MTL collaborated with Johosoken for the development of HSDPA MIMO systems.

Future Direction: MIMO system enhancements for HSDPA.

Contact: Jyhchau (Henry) Horng
http://www.merl.com/projects/HSDPA/

Lab: MERL Technology Lab
Project Type: Advanced Development

Dynamic Resource Control for Shared Downlink Wireless Channel



This project explores dynamic resource control and system optimization for shared downlink wireless channel such as 3GPP HSDPA (High Speed Downlink Packet Access). We propose a new resource control framework and particularly a dynamic multi-class traffic scheduling mechanism to provide QoS (Quality of Service) for multimedia services. Inter-layer resource interaction and the entire system optimization are also studied in this project.

Background and Objectives: It can be foreseen that multimedia services and applications such as video cellular phone and web browsing will proliferate in the next generation wireless network. Quality of Service (QoS) provision for multimedia services is regarded as a challenging task due to special characteristics of wireless links such as fading and mobility. This project aims to provide effective resource control techniques for multimedia over wireless networks.

Technical Discussion: We propose a new dynamic resource control framework integrated with adaptive modulation and coding (AMC) and hybrid automatic repeat request (H-ARQ) to support class-based applications. Within the framework, transmission requests are sent to the Admission Control module. The Admission Control module makes the decision whether or not to admit new transmission streams by computing the available resource from physical-layer resource measurement and existing traffics. Once admitted, traffic streams enter the Traffic Shaping and Classification module, and then are passed to different queues according to traffic. Traffic streams are classified according to delay and packet loss requirements. These QoS parameters decide the queue length, the weight of a queue and RED (Random Early Detection) configuration. Because H-ARQ retransmission may introduce large amount of extra traffic load when wireless channels are in bad condition, we assign two queues to each traffic class: original transmission queue and re-transmission queue. For each class, sub-classes (different colors) are specified according to MCS (Modulation and coding scheme). Both spreading codes and time frames are scheduled to UEs. We propose a new wireless scheduling algorithm, delay-sensitive Dynamic Fair Queueing (DSDFQ), to meet delay and loss requirements of multimedia applications as well as maintain high network efficiency. Our approach can easily adapt to load fluctuations from different traffic classes and dynamic wireless channel status affected by user mobility, fading and shadowing.

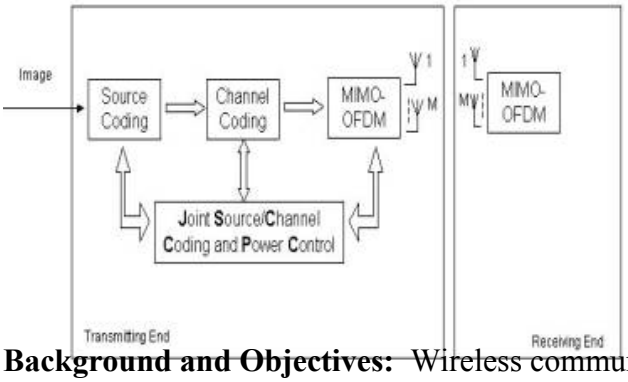
Collaboration: J. Zhang, D. Gu and P. Orlik from MERL-TL.

Future Direction: The future work includes how to reduce signaling load between UEs and node B by decreasing spreading code switch times for each UE at the scheduler.

Contact: Huairong Shao, Chia Shen
http://www.merl.com/projects/dynamic-resource-control/

Lab: MERL Research Lab
Project Type: Research

Energy-Efficient Wireless Image/Video Transmission



In general, the purpose of the energy-efficient transmission scheme being developed in MERL is to jointly adapt the source coding, channel coding and transmit power level of an image/video to channel conditions such that energy-consumption is minimized under a distortion constraint. We specifically study quality-progressive coded JPEG2000 images, which consists of multiple quality layers.

Background and Objectives: Wireless communication has become very important in today's life. Currently, it is mainly used for speech and data transmission. Together with the deployment of 3G and beyond systems, visual communications consisting of high-resolution image and video will be widely demanded. However, transmission of video in mobile wireless networks carries many challenges such as energy constraint, bandwidth limitation, severe channel conditions with varying bit error rate (BER), error propagation in compressed bit stream.

Energy consumption occurs due to computational complexities of source and channel coding and transmission of bit streams. The transmission energy cost depends on joules/bit at the transmitter and the size of the bit stream. Therefore, the schemes used for source and channel coding and the transmit power level determines overall energy consumption. The objective is to find the optimum set of coding schemes and transmit power level for each quality layer of a JPEG2000 image such that the overall energy consumption will be minimized under a distortion constraint.

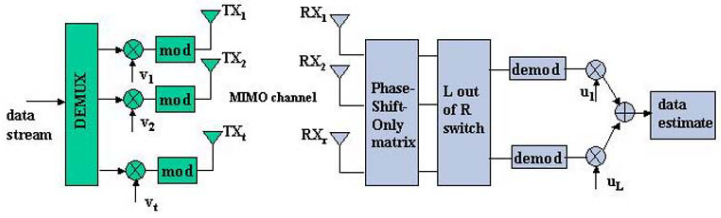
Technical Discussion: We address channel adaptive wireless transmission of a quality progressive JPEG2000 images with minimum energy consumption under a quality of service (QoS) constraint. In particular, based on the motion JPEG2000 variable rate space frequency codes, we develop an optimal and efficient joint source-channel coding scheme with transmit power control to minimize the consumed energy for point-to-point transmission. Protection-redundancy and transmit power level for each generated quality layer is determined according to a receiver feedback. Received SNR is used for estimating channel condition.

Collaboration: We are expecting that the output of the project will be beneficial to the Imaging Products Storage Department in Kyoto Works to improve the image transfer rate and quality for the wireless image delivery platform they have been developing.

Future Direction: We target to modify the scheme to provision energy-efficient high quality delivery in highly mobile multi-hop transmissions. The new applications that this technology can be deployed are handsets such as PDAs, cell-phones and ad-hoc surveillance systems.

Contact: Zafer Sahinoglu **Lab:** MERL Technology Lab
http://www.merl.com/projects/EnergyEfficient/ **Project Type:** Initial Investigation

Antenna Selection



Increasing the data rate is one of the most important goals for wireless data links, both for cellular (3G) applications, and for the enormously popular wireless LANs. As spectrum is expensive, the most promising way for achieving this is the use of multi-antenna systems. Such systems can be used to improve the transmission quality of a single data stream, or to allow the transmission of several, parallel, data streams. At the same time, transceiver complexity should be kept to a minimum in order to allow any new, multiple-antenna devices to be priced competitively with existing single-antenna solutions.

Background and Objectives: The "Broadband Mobile Communication Systems" project of Mitsubishi Electric, is investigating next-generation solutions for high-data rate communications, using both theoretical investigations and building a testbed for trying out the different methods for performance enhancement.

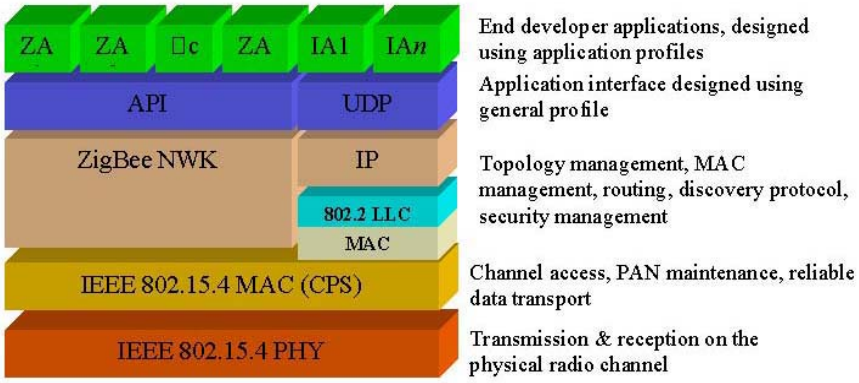
Technical Discussion: The main drawback of any multiple-antenna system (with N antennas) is the increased complexity, and thus cost, since it requires N complete transceiver chains. There are numerous situations where this high degree of hardware complexity is undesirable - this is especially important for the mobile station (MS). Additional antenna elements (patch or dipole antennas) are usually inexpensive, and the additional digital signal processing is becoming less of a burden as digital processing becomes ever more powerful. However, the RF components like low-noise amplifiers, mixers, etc., are very expensive. We investigate reduced-complexity schemes where the best L out of the available N signals at the antenna elements are received, downconverted, and processed. Traditionally, these schemes suffer from poor performance in highly correlated channels. We recently devised a phase-shift-based transformation that assures good performance in any type of channel.

Collaboration: Princeton University, Johosoken.

Future Direction: We have investigated our new scheme for the receive case and we are working on the application to the transmit case. In the near future, we will submit it to 3GPP and IEEE 802.11 for inclusion in the standards.

Contact: Andreas F. Molisch **Lab:** MERL Technology Lab
http://www.merl.com/projects/AntennaSelection/ **Project Type:** Advanced Development

ZigBee



The ZigBee Alliance is defining an international standard to enable reliable, cost-effective, low-power, wirelessly networked monitoring and control products. Some of the targeted applications include: home, building and industrial automation and control, PC peripherals, medical monitoring and toys. MERL is actively participating in this development and the figure on the left shows the ZigBee protocol stack that is currently under development.

Background and Objectives: Until very recently most international standards in the wireless communications space focused on providing either voice quality transmission or high rate data communications. 3GPP and IEEE 802.11 are well known examples. ZigBee, however, seeks to develop a standard in a completely new application space, that of extremely low-cost, low-rate data communication. The idea is to wirelessly enable all sorts of devices from gaming joy-sticks to simple light switches.

With the publication of the IEEE 802.15.4 standard, which defines, low-rate (<250Kbps) physical and MAC (Medium Access Control) layers some building blocks of ZigBee are complete. On top the 802.15.4 standard the alliance is defining a new network layer that incorporates ad-hoc features such as multi-hop mesh routing and self-organization. Additionally, network security algorithms and application profiles are to be defined.

Technical Discussion: MERL has made contributions to the development of the ZigBee network layer and some security algorithms. Specifically, a multi-hop mesh routing algorithm based on a simplified AODV (Ad-hoc On-demand Distance Vector) routing protocol was proposed jointly with Intel and Eaton Corp. This proposal has been adopted as part of the routing specification. Further modifications have been proposed to enable a heterogeneous network consisting of routing nodes that will not require any storage routing tables as well as more capable routing nodes with memory for tables. The solution is a combination of tree-based routing and AODV that allows trade-offs between optimal route selection and the cost of devices.

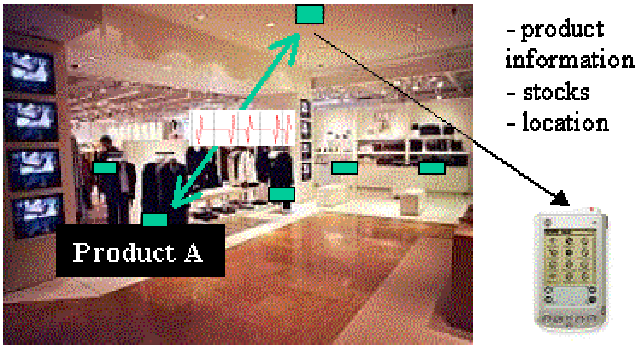
Collaboration: We have worked closely with MELCO representatives to the ZigBee Alliance and are collaboration with Sentansoken/Sanden and others for future development.

Future Direction: We will continue our participation and work in the ZigBee alliance. As well as work towards building ZigBee prototype device and ZigBee applications.

Contact: Philip Orlik
http://www.merl.com/projects/ZigBee/

Lab: MERL Technology Lab
Project Type: Research

Ultra Wideband Systems



Ultra Wideband (UWB) techniques are viewed by many companies as having the potential to combine high resolution for geolocation with communication for low cost and low power consumption. The task group of the High Data Rate section of the Wireless Personal Access Network standard (802.15.3a) starts working on the standard for alternative physical layer for piconets of 10 meter range and for a minimum data rate of 110Mbps. Furthermore a group of interest is gathering companies to create a study group for low data rate applications (802.15.4IGa).

Background and Objectives: In February 2002, FCC adopted a First Report and Order that permits the use of UWB devices. Radar and imaging companies were waiting for this decision for several years in order to commercialize their products and to look for new applications, including for short range communications. The low power consumption of UWB devices opens the door to applications requiring batteries such as storage data devices. Furthermore, the data rates which are considered by the task group of 802.15.3a allow several new applications like wireless Digital TV, high definition MPEG2 motion picture transferring, DVD playback and DV Camcorder. The main objectives are to influence IEEE standard activities in the area of Wireless Personal Area Network, to design a transceiver matching the MELCO needs for communication with high data rates and to prototype UWB radios for target applications.

Technical Discussion: There are several different techniques under the name Ultra Wideband. A train of very narrow pulses using a very wide bandwidth is one of them. It gives an efficient way to combat multipaths in indoor environments and provide a high time resolution for tracking, geolocation and imaging. But multiband systems with frequency hopping sequences using up converted pulses or OFDM symbols are also considered. The main difficulty is to generate a signal conforming to the power limitations from FCC and at the same time to be able to achieve high data rate for a reasonable price.

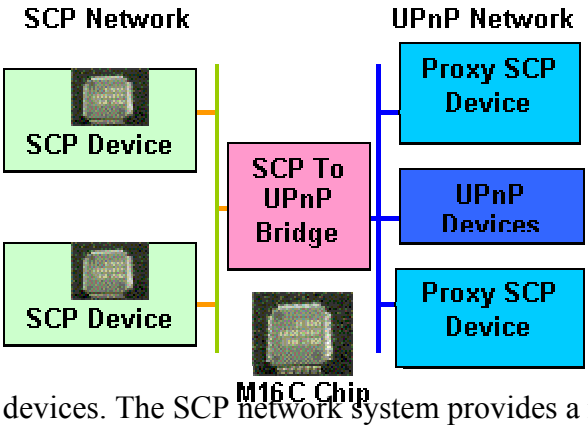
Collaboration: This project is performed in collaboration with Johosoken, Sentansoken, ITE, the Georgia Institute of technology and Princeton University.

Future Direction: A multiband approach using OFDM symbols will be investigated. We will continue to actively participate to the IEEE WPAN standard to define the alternative 802.15.3 physical layer and enhanced the MAC layer. We will also investigate how to apply our ideas to the low data rate applications targeted by the future 802.15.4a Study Group.

Contact: Andreas F. Molisch, Jinyun Zhang
http://www.merl.com/projects/uwb/

Lab: MERL Technology Lab
Project Type: Research

Simple Control Protocol



The Simple Control Protocol (SCP) is a non-IP based home networking protocol, proposed by Microsoft Corporation. It is a peer-to-peer, media independent protocol that allows manufacturers to produce small and intelligent devices that communicate with each securely and robustly. Currently, it uses the Power Line Carrier (PLC) as its physical layer.

An SCP network consists of a physical network and multiple logical networks. The physical network provides a common network medium that is used for communications among the devices. The SCP network system provides a mechanism for defining the logical networks and adequate security. The SCP is a complementary technology to Universal Plug-n-Play (UPnP), a IP based protocol. The devices in SCP network can be bridged into UPnP network. MERL has ported the SCP software onto Mitsubishi M16C chip. The diagram above shows a M16C chip and a sample SCP/UPnP network system.

Background and Objectives: The home networking is an important emerging area with vast opportunities for growth and revenue. Given the wide variety of home devices, no single networking technology will satisfy the customer requirements. As opposed to IP-based UPnP, the SCP targets a class of devices with limited processing resource, low bandwidth requirement, and a relatively low price. This class of devices mostly consists of home appliances and sensors. The main goal for networking of these devices is to provide automated operational control, diagnostics, and remote updating of firmware, etc.

Technical Discussion: The current implementation of the SCP/PLC runs on Arm7 and M16C based platforms. The software consists of several inter-connected modules such as lower level hardware dependent code, a multi-threaded pre-emptive RTOS, software for physical layer, and the protocol stack itself. The significant architectural differences of the two processors mean that the bulk of the porting effort consists of implementing the lower level processor specific code. As opposed to M16C, the Arm7 processor has several processing modes with separate register banks.

Collaboration: MERL is working closely with Renesas and Johosoken on the current SCP technology development and the future SCP application deployment.

Future Direction: The SCP is one of several emerging home networking technologies. It provides a good basis for the development of new low-cost hybrid systems such as sensor networks, energy monitoring system, automated meter reading. The SCP devices need to be interoperable with other devices such as UPnP device, HAVi device, etc. through the bridging technologies.

Contact: Jianlin Guo, Jinyun Zhang
http://www.merl.com/projects/SCP/

Lab: MERL Technology Lab
Project Type: Advanced Development

Concordia for Power Systems Business Units

The Concordia mobile agent technology has been successfully transferred to Mitsubishi Electric's power systems business units as a core technology in their solutions for the electric power industry.

Background and Objectives: The Concordia mobile agent technology has been well received by Mitsubishi Electric's power systems business units and has become a core technology in their systems solutions for the electric power industry. In the past few years, the Concordia program has focused much of its development efforts specifically to meet their requirements. Furthermore, we have conducted thorough technology transfer to these business units in order that formal product customization, maintenance, and support for Concordia can be established in Japan.

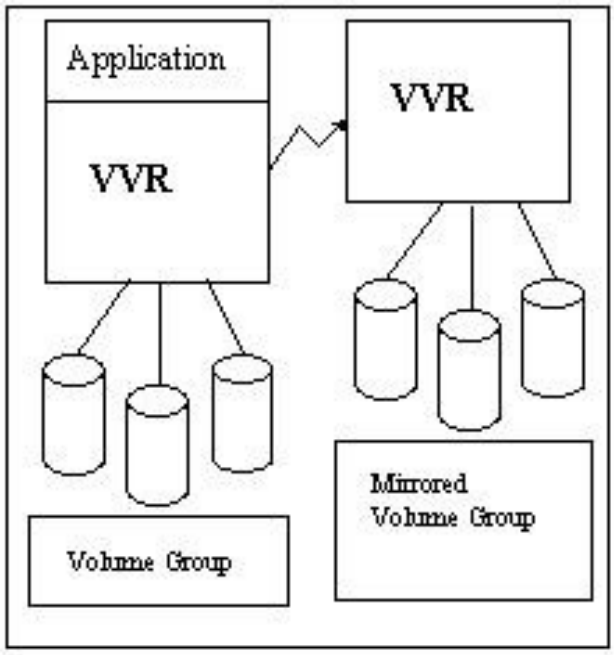
Technical Discussion: Some of the important Concordia product features that have been developed specifically for the power systems business units include: 1) streamlining of the Concordia core, 2) enhancing Concordia for wide area network support, and 3) enhancing Concordia security features. The effort on streamlining the Concordia core and reducing its memory footprint resulted in the subsequent development of Featherweight Concordia for embedded devices.

Collaboration: The Concordia program has been working very closely with Mitsubishi Electric's power systems business units for the past three years. MERL has also worked closely with the Texas A&M University and PSERC (Power Systems Engineering Research Center) to explore new applications of agent technology in the power systems domain.

Future Direction: We will continue to serve in the role of being the primary mobile agent technology provider to Mitsubishi Electric's business units.

Contact: David Wong **Lab:** MERL Technology Lab
http://www.merl.com/projects/concordiaSanden/ **Project Type:** Advanced Development

Network Replication



The Network Replication project is a cooperative effort with Veritas Software to provide network replication capability to their popular Volume Manager product, VxVM. Our technology has been licensed by Veritas for their Veritas Volume Replicator product (VVR).

The replication engine of VVR offers an ideal solution for the remote archiving of storage such as databases or file systems through its ability to replicate a virtually unlimited number of related data volumes while maintaining consistency. Replication uses standard networks without proprietary hardware and is highly resistant to system and network failure.

Veritas Software is a market leader in storage management software.

Background and Objectives: The Network Replication project evolved from an R&D exploration of network distributed storage. Establishment of our relationship with Veritas Software refined this objective to network replication of Veritas' logical volumes.

Technical Discussion: Veritas' customers use VxVM to protect their data from media failure by creating local mirrors, but remote mirrors were needed to protect against system and infrastructure failure. VVR replicates groups of logical volumes, maintaining consistency among them during replication. Changes to any member of the volume group are transparently captured and replicated to one or more remote locations.

VVR has flexible replication and configuration characteristics. It can replicate synchronously or asynchronously. Feedback allows input flow rates to be throttled if necessary to match available network bandwidth. It can simultaneously replicate volume groups to multiple remote sites while acting as a receptacle for volume groups replicated from other sites, and it can support an unlimited number of volume groups. Recent work has focused on integrating VVR with Veritas' Cluster Volume Manager (CVM), to allow replication from clusters.

Collaboration: Since 1997, CSL has worked exclusively with Veritas Software in the development of VVR, deriving market direction and market access from this relationship.

Future Direction: Veritas has purchased the rights to this product from Mitsubishi, so MERL will no longer be contributing to the product.

Contact: David Rudolph
http://www.merl.com/projects/netrep/

Lab: MERL Technology Lab
Project Type: Advanced Development