Mitsubishi Electric Research Laboratories (MERL)

Annual Report

July 2003 through June 2004

TR2004-00

(Published August 2004)

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 and our projects.

201 Broadway, Cambridge, Massachusetts 02139
617.621.7500

Production:

Adam Bogue, Karen Dickie, Janet O'Halloran, Richard C. Waters

Table of Contents

Mitsubishi Electric Research Laboratories

Mitsubishi Electric Research Laboratories (MERL) is the North American arm of the corporate 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 five principal technology sectors:

Computer Vision – featuring the observation of people in images.
Digital Communications - featuring wired networks and wireless transmission.
Digital Video – featuring encoding, decoding and analysis of video.
Off-the-Desktop Interaction and Display – featuring novel devices and interface concepts.
Sensor and Data Systems – featuring novel sensors, communication and system architectures.

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's output ranges from papers and patents, through proof-of-concept hardware and software prototypes, to modules for 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

MERL Organization

MERL consists of two laboratories, which share the same space in Cambridge, Massachusetts and collaborate closely to achieve groundbreaking results. The six members of the top management team work closely together, guiding all aspects of MERL's operation.

Mitsubishi Electric Research Laboratories
Dr. Richard C. (Dick) Waters (President, CEO & Research Fellow)

- Liaison to Japan
  Mr. Takashi Kan (EVP & CFO)
- Research Lab - *Basic & Applied Research, staff 21*
  Dr. Joe Marks (VP & Director)
- Technology Lab - *Applied Research & Advanced Development, staff 43*
  Dr. Kent Wittenburg (VP & Director)
  Dr. Huifang Sun (VP, Research Fellow & Deputy Director)
- Marketing & Business Development
  Mr. Adam Bogue (VP)

Richard C. (Dick) Waters *Ph.D., MIT, 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 Computing 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, he worked on computer-related projects involving computer architecture, parallel computing, digital broadcasting and image processing. Before coming to MERL in 2002, Takashi was general manager of the Multimedia Laboratory within MELCO's Information Technology 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 parsing. 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.

MERL Research Lab

The MERL Research Lab (MRL) pursues basic research in applied computing. Its efforts are directed towards applications of practical significance, but its time horizon is long (five or more years) and its appetite for technical challenge and risk is high. Located in Cambridge, Massachusetts, hometown of Harvard University and the Massachusetts Institute of Technology, the permanent staff is enriched by a large number of student interns, visiting scientists, consultants and academic collaborators, as well as collaboration with MTL and MELCO colleagues in Japan. All of MERL participates actively in the external research community, exposing its work to critical peer review, and publishing its results as quickly as possible.

MRL's primary areas of research are computer vision, audio & speech processing, sensor networks, optimization & control, 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, MRL has been successful.

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 MRL.

MERL Technology Lab

The research at MERL's Technology Laboratory (MTL) is intertwined with MRL. However, while 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 ideas into full-bodied forms 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 MELCO's CR&D labs worldwide. Second, we participate in international standards bodies, benefiting from the competitive challenges of the standards process.

A particular strength of MTL is its emphasis on collaboration across conventional academic and engineering disciplines. Expertise in the lab ranges from chip-level and board-level hardware through low-level communication protocol layers to software and applications. The friendly and cooperative culture of MTL, which is small by industry standards, allows a unique opportunity for cross-fertilization of research ideas that truly make a difference.

Mitsubishi Electric

Number 151 on Fortune magazine's 2003 list of the world's 500 largest corporations, Mitsubishi Electric Corporation (MELCO) has $29 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 recent worldwide slump lead to difficulties. However, energetic restructuring and cost cutting has lead to renewed profits and created a firm foundation for growth now that the Japanese economy is recovering.

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 2004 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.

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-Tokyo Financial Group, Mitsubishi Motors, Mitsubishi Heavy Industries, and Mitsubishi Trading Company (all four of which are also on the Fortune Global 500 list—Numbers 217, 224, 238, & 389 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 US Operations

Approximately 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 2004 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	

Mitsubishi Electric Corporate R&D

Number 55 on the IEEE Spectrum list of the top R&D spenders in 2002, MELCO has a global R&D network comprising five laboratories. 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), 18 people (Tokyo) Managing MELCO's R&D	Hatsuhon
Advanced Technology R&D Center (ATC) Research & Advanced Development: Dr. K.Kyuma (GM), 988 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. I. Arai (GM), 95 people (Ofuna) Industrial Design, Usability Studies	ID-ken
Mitsubishi Electric Research Laboratories (MERL) Research & Advanced Development: Dr. R.Waters (CEO), 81 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), 57 people (France & England) Wireless Communications, Digital Audio & Video	ITE

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 and Senior Members of technical societies. The second and third are best paper awards and technology awards received from outside organizations. The fourth is awards received from MELCO for MERL's contribution to MELCO products. Listed below are achievements and awards for the period of this Annual Report, July 1, 2003 through June 30, 2004.

Current Technical Society Fellows and Senior Members

Three (5%) of MERL's technical staff are fellows of professional societies:

Dr. Huifang Sun, Fellow Institute of Electrical and Electronic Engineers;
Dr. Charles Rich, Fellow American Association for Artificial Intelligence;
Dr. Candace L. Sidner, Fellow American Association for Artificial Intelligence.

A further Nineteen (30%) of MERL's technical staff are Senior Members of the Institute of Electrical and Electronic Engineers (IEEE). (Only 7% of the 361,000 members of the IEEE are Senior Members.)

Best Paper Awards

Viola, P.; Jones, M.J.; Snow, D., "Detecting Pedestrians Using Patterns of Motion and Appearance", *IEEE International Conference on Computer Vision (ICCV)*, Vol. 2, pp. 734-741, October 2003. [Best Paper Award]

Molisch, A.F.; Zhang, X.; Kung, S-Y; Zhang, J., "DFT-Based Hybrid Antenna Selection Schemes for Spatially Correlated MIMO Channels", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Vol. 2, pp. 1119-1123, September 2003. [Excellent Paper-Most Highly Evaluated]

Technology Awards From Outside Organizations

In November 2003, MERL's Senior Research Scientist, Ramesh Raskar was recognized as a winner of the Global Indus Technovator Award, instituted at MIT. This honor recognizes the top 20 Indian technology innovators worldwide from more than 150 nominations from 15 countries.

Awards From MELCO

In early 2004, MERL staff received a "MELCO Corporate R&D Director's" award for DiamondTouch hardware and software (pages 95-96).

In early 2004, MERL staff received a "MELCO Valuable Invention" award for a patent on projector mosaicing. (Subsequent related work to this is presented on page 100.)

Business Impacts

This section details the impact of MERL on MELCO's business in four areas: product features, system components, licensing, and standards contributions. In each of these areas, there is continuing revenue from MERL technology that had its initial impact in previous years. This section presents only those items whose first impact occurred in the 12 months covered by this annual report.

A dream of MERL is to create a new high volume product for MELCO. We have not yet achieved this, but we have contributed important new features to a number of products. For such features, we take the date on which the product with the feature was first produced for sale as the date of MERL's impact on MELCO.

A large part of MELCO's business is in the form of large custom systems for business or government. MERL has contributed components to a number of such systems. For these components, we take the date at which the contract for the first system using it was signed as the date of impact. It can often be a year or two before the system is completely delivered to the customer, but payments are usually made to MELCO in stages and the moment of contract signing is the moment when the payment becomes assured.

A different kind of way that MERL can impact MELCO is by making standard contributions. This may or may not lead to direct revenue via licensing. However, in any case, it allows MELCO to keep closely in touch with important standards and to shape these standards for maximum benefit to MELCO. For standard contributions, we take the date at which a contribution is included in a draft of the standard as the impact date. It is typically a year or more after that before a standard becomes final and yet more years before the existence of the standard directly benefits MELCO. However, the date of inclusion in a draft is the date at which it becomes reasonably assured that MERL will have a beneficial impact on MELCO.

A final way that MERL can impact MELCO is by licensing MERL IP to third parties and obtaining direct revenue as a result. For licensing, we take the date on which a license agreement is signed as the impact date. MERL has signed several licenses, but as it happens none were signed in the last 12 months.

The following subsections detail what MERL's impact on MELCO has been in the past year. In addition, they summarize how this impact was achieved. It is worthy of note that there are several distinct models of how impact can be achieved ranging from work specifically requested by MELCO to finding an application in MELCO for a technology developed independently by MERL.

Recommendation Engine for Diaprism

In August 2003, the Information Systems business unit (ISHon) began selling recommendation software as a companion product of its Diaprism commercial database hardware product. The heart of the software is a recommendation algorithm developed by MERL that can make predictions (e.g., of customer interests) based on information in a database. The recommendation software is

only a small part of the Diaprism product line, but such new features are essential to keep the product line fresh.

Data modeling, data mining, optimization and control are key aspects of MERL's research. We expect to contribute to MELCO business in a number of ways in this area. Current MERL work includes: dimensionality reduction for data modeling (page 118), various kinds of data mining (page 120), optimization for elevator control (page 119), and man-machine collaborative optimization (Page 121).

Details: In 2001, an MRL researcher working on fundamental questions in computer vision developed a new way to do Singular Value Decomposition (SVD). SVD is used to analyze the fundamental structure of high dimensional data. Suppose that a database contains 1,000 100x100-pixel pictures of faces and you want to know what features of these pictures (e.g., the shape of the face or the size of the nose) are the best ways to tell one picture from another. (This is a question that is fundamental to good face recognition systems.) You can view this data as a 10,000 by 1,000 matrix of intensity values (one row for each pixel and one column for each picture). Using SVD you can determine, for instance, what the six most important features in the data are.

SVD is extremely powerful and used in many applications. However, as traditionally calculated, it has two problems. First, it is quite costly to compute on the kinds of huge matrices it is usually applied to and must be entirely recomputed whenever any change is made in the matrix (e.g., when a new face is added to (or removed from) the data base above. Second, SVD requires all the data to be present in the matrix. This is often not the case due to incomplete data collection. When there is missing data, it has to be imputed (estimated) by some method before SVD can be applied. This process can also be costly and can introduce errors.

Faced with these problems in the application he was attacking, the MRL researcher invented a new approach to SVD called Incremental Imputative SVD (IISVD). This deals with incomplete data in an efficient and elegant way and can incorporate new data incrementally without having to reprocess the entire matrix. This latter feature leads to dramatic reduction in computation time in a variety of well-known SVD applications.

The researcher immediately began to look for applications where IISVD would have the greatest benefit and began to look at the problem of making recommendations. As an example, he began to experiment with some data about movie preferences that had been collected by the University of Wisconsin. This data specifies for thousands of people and thousands of movies, some information about which movies each person likes and dislikes. This data is extremely sparse, because information only exists for a few dozen movies for each person. The job of a recommendation engine is to fill in the missing data and estimate how people would rate movies they have not yet seen. The highly rated movies can then be recommended to them.

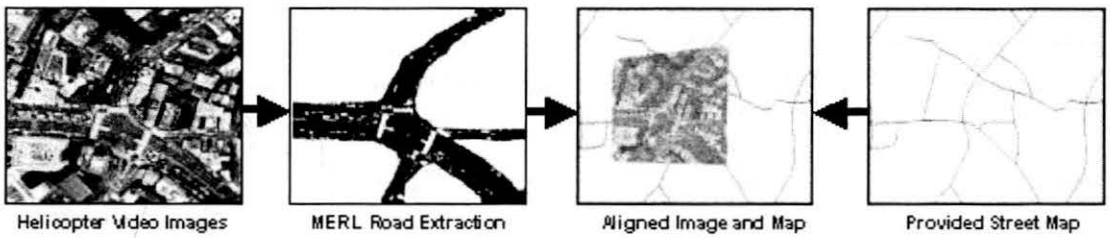
SVD does a reasonable job with this task, but slowly due to the enormous amount of data that must be processed whenever new information is added to the database. (A real application might have to deal with hundreds of thousands of people.) Using IISVD good results can be obtained so rapidly that new people or preferences can be added and recommendations can be updated almost instantly.

A prototype movie recommender was implemented and shown to a number of people in Japan in the fall of 2001. It was seen by a researcher at MELCO's Johosoken laboratory who made the

connection to ISHon's Diaprism database product line. At the request of Johosoken, MERL upgraded the prototype, adding a number of new features. Based on the revised prototype, Johosoken created code for inclusion in the Diaprism product line. This code was beta tested with a customer in the fall of 2002 and became a MELCO product in 2003.

Road Recognition in the "Heli-Tele" System

In the fall of 2003, the Japanese government signed a contract with MELCO's Kobe works for the delivery of a system called "Heli-Tele". The system uses a MERL component to automatically locate roads in aerial photographs taken by helicopters, this information is then used to align the photographs with pre-existing maps, so that they can be shown as overlays on these maps. MERL's technology is only a small part of the system as a whole; however, it provides a key part of the functionality.



Computer-vision-based image analysis is a major area of research at MERL. This has lead to a wide range of technologies including tracking (page 54), face detection (page 58), face recognition (page 56), face-based surveillance video analysis (page 55), and human activity detection (page 51).

Details: One of MELCO's many businesses is the design and construction of earth observation satellites. Knowing that MELCO is interested in producing the software systems used for analyzing satellite data, MERL has investigated how its technology might be used in such systems. One such investigation was done by an MTL researcher in 2002. Extending his earlier graduate work. He built a prototype system for identifying roads in satellite images.

This prototype was shown to various people in MELCO in late 2002 and early 2003. It did not generate much interest from people working on satellite data, but it was seen by a researcher at MELCO's Johosoken laboratory who was working with the Kobe works in an effort to obtain the contract for the Heli-Tele system.

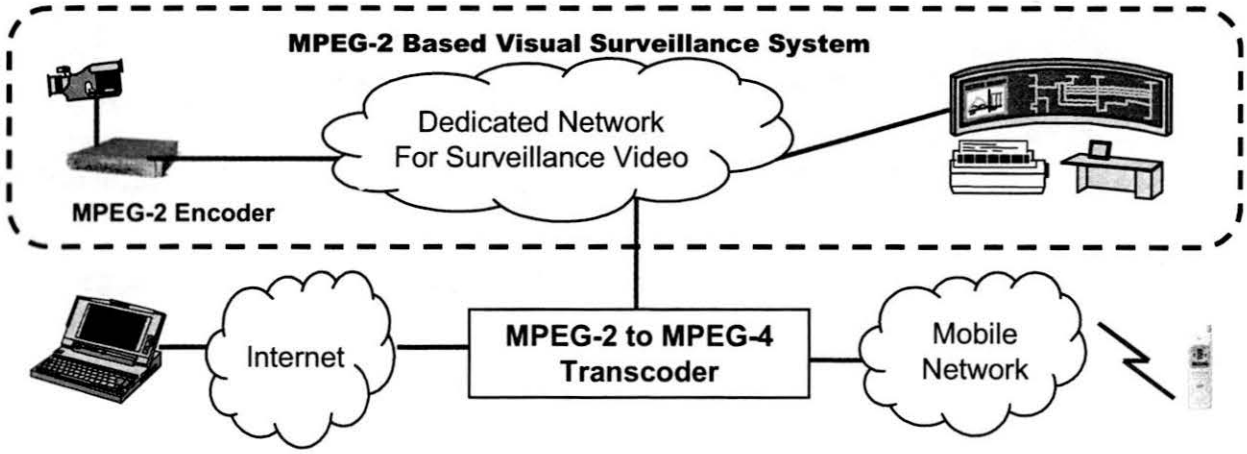
In the Heli-Tele system, helicopters with cameras are used to survey an area, e.g., in a disaster situation. The helicopters are equipped with GPS units and can report the approximate orientation and position (within 20 meters) from which each image is taken. This allows approximate alignment of the images with map data. However, to get exact alignment, the GPS data is not enough. Upon seeing the MERL demonstration, the Johosoken researcher had the idea that exact alignment could be achieved by matching roads recognized in an image, with the roads in the map data. To test this idea, he asked MERL to adapt its prototype to aerial images instead of satellite images.

It turns out that the technology needed to recognize roads many pixels wide in aerial images is entirely different from the technology needed to recognize roads a fraction of a pixel wide in

satellite images; however, the MTL researcher new how to do this task equally well and produced a prototype system in the summer of 2003. This prototype was used in demonstrations that were critical to MELCO's efforts to win the Heli-Tele contract and was delivered as part of an initial version of the system in January 2004. MERL is currently working on an improved road recognition module, which will be delivered as part of the final system in August of 2004.

MPEG-2/4 Transcoding

In February 2004, MELCO's Koriyama works began shipping a PC-based product (the BC-5600) for converting MPEG-2 encoded video into MPEG-4 encoded video. The prime purpose of the system is to take multiple MPEG-2 streams from surveillance cameras and convert (transcode) them into compact, low-resolution MPEG-4 streams so they can be communicated effectively over the Internet. The transcoding is done entirely in software using a module developed jointly by MERL and Johosoken that incorporates software designed and written by MERL (page 88).



The BC-5600 is just one of what we expect will be several applications of MERL's transcoding module in Melco. Another business unit is planning to use MERL's transcoding in a surveillance related product in the coming year. In addition, we are discussing a number of potential applications in the area of video storage and transmission for entertainment.

Details: The standard approach to transcoding is to use decoding hardware to decode the source video and then use encoding hardware to create output in the target encoding. However, this is both costly and inflexible due to the need for special purpose hardware.

In 1999, a pair of MTL researchers began a project on software-only transcoding, based on pioneering work done by one of them several years earlier. In 2000, this work resulted in key methods for doing spatial resolution reduction (making a picture with a small number of pixels from one with a large number of pixels) for MPEG-2 and MPEG-4. These methods are fast in part because they operate directly on the compressed form of the video without having to decompress it. This work was considered groundbreaking by the scientific community and led to several awards.

The work also led to a practical PC-based software-only transcoder for converting high resolution MPEG-2 into lower resolution MPEG-4. (The methods of compression used by MPEG-2 and MPEG-4 are very similar. However, MPEG-4 is more flexible and allows for

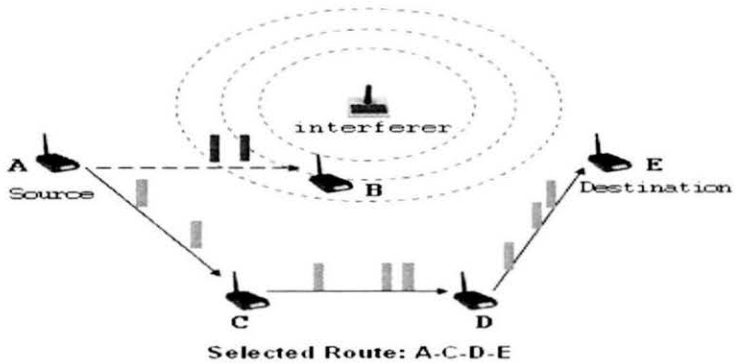
various low resolution and/or low frame rate video encodings.) This transcoder was presented to various people in Melco in late 2000 and early 2001 and generated considerable interest

In recent years, many large surveillance systems (from MELCO and others) have been built around cameras that produce MPEG-2 encoded output. This is highly effective as a means for communication video over a dedicated digital network. However, the large bit-rate needed for MPEG-2 means that it is not practical to send the video out over the Internet for remote monitoring. For remote viewing, one would like to have lower resolution video in MPEG-4; however, Melco was hesitant to go in this direction because hardware transcoding is expensive and cannot conveniently deal with the multiple, varying frame rate video streams one typically encounters in surveillance situations. This particular need for a superior method of transcoding was not known to MERL, but it was known to key groups in Melco, which became interested in MERL's transcoder as soon as they heard about it.

At the request of Melco, MERL focused its transcoder work on surveillance related applications starting in 2002 and produced a product ready software module in 2003. At MELCO's further request, we are currently maintaining this module and investigating ways to extend it to other compression standards, other kinds of computational hardware, and other applications.

Link Quality Adaptation in ZigBee

In March 2004, MERL's contribution "Link-Quality-Indicator-Based Routing Protocol" was included in draft v0.8 of the ZigBee standard (page 69). This is a fundamental contribution to the standard specifying how to implicitly assemble a group of ZigBee nodes into a network and how to route messages over this network.



MERL continues to work actively on the ZigBee standard and ZigBee applications. This includes further proposals to the ZigBee standard (pages 70-72), work on an SCP/ZigBee bridge (page 78), and applications of ZigBee to sensor networks.

Details: MELCO was a founding member of the ZigBee alliance in 2002. The goal of this alliance is to create a standard for low cost, medium data rate (20-200 Kbits/sec), short-range (30 meters) wireless communication over ad hoc networks. The prime target is monitoring and control applications in home automation; however, there are industrial and sensor network applications as well. A basic motivation is to produce something that is cheaper and better than BlueTooth.

ZigBee is relevant to many product lines within MELCO including home appliances, home entertainment, and industrial systems. However, the primary initial force behind MELCO's

support of the alliance was the Semiconductor business unit (Hanpon), whose goal is to make ZigBee interface chips. This goal remains strong in Renesas Technologies Corporation, which was formed by the merger of most of MELCO and Hitachi's semiconductor businesses in 2003.

MTL has a long history of involvement in home networking including work with standards like IEEE 1394, Havi, HomeAPI and SCP. MTL also has a long history of working with Hanpon/Renesas. This included working with Hanpon on an SCP/PLC interface chips in 2001-03 (page 77).

As a result, it was only natural that Hanpon turned to MERL for help when Hanpon joined with other companies to initiate the ZigBee standard. MERL jumped into ZigBee standard activities at the end of 2002 and was able to make proposals shortly thereafter. Because MERL was there at the right time with the right knowledge, we were able to get our contribution accepted as part of the fundamental basis of ZigBee.

Point-Based Rendering in MPEG-4

In March 2004, MERL's contribution "Point based Rendering for MPEG-4 AFX" (page 87) was included in the draft of the Animation Framework eXtension (AFX) which is part of the Synthetic Natural Hybrid Coding (SNHC) section of the MPEG-4 standard. MPEG-4 has long included the ability to support polygon graphics. MERL's contribution adds support for point-based graphics.

In addition to working on point-based graphics in the context of the MPEG-4 standard, MERL is participating in the MPEG 3D (Multi-View Video Coding) discussion, which may mature into a standard. MRL is also working on technology for live 3D TV (page 86). In addition, key technology for capturing point-based graphics data (page 57) is being used as a basis for 3D face recognition (page 56).

Details: In the fall of 1999, an MRL researcher began work on point-based graphics in collaboration with a researcher at the Swiss Federal Institute of Technology (ETH) in Zurich. This lead to major papers on rendering images based on point-based graphics at the SIGGRAPH conferences in 2000 and 2001, which are some of the most cited papers in this new sub field of graphics. MERL's focus then switched to scanning technologies for obtaining point-based graphics data from real objects. This led to further groundbreaking work in 2001-2003.

To date, most graphics is polygon-based. The 3D shape of an object is modeled as surfaces created by combining large numbers of (usually very small) triangles edge to edge. The appearance of the object is modeled by applying (usually very small) "textures" (image segments) to the faces of the triangles. An image of the object from a particular viewpoint is created by determining which location(s) on which textured triangle(s) correspond to each pixel in the desired image.

In contrast, point-based graphics models an object as a cloud of points (analogous to pixels in a 2D image) each of which has a 3D position and a color. An image from a particular viewpoint is created by determining which point(s) correspond to each pixel in the desired image. The key potential advantage of point-based graphics is that it can be directly captured from a real object

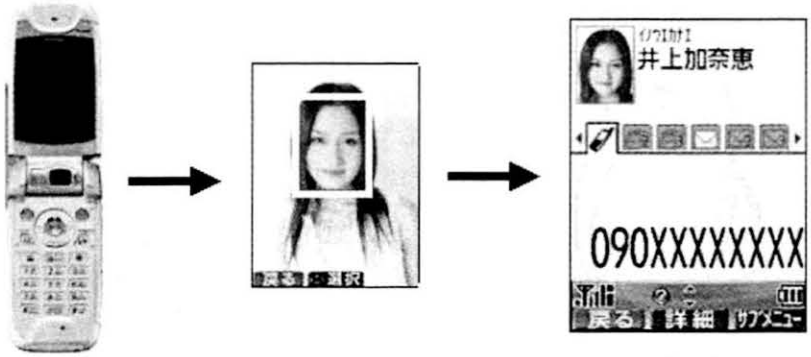
in analogy with taking a photograph and the hope that processing of point-based graphics could be faster because the data structures involved are simpler and more uniform.

The work on point-based graphics at MERL was initially spurred by the desire of the semiconductor business unit (Hanpon) to get into the business of 3D graphics rendering chips. However, Hanpon eventually decided not to move in that direction. As a result, the MRL researcher looked for other ways that MERL's point-based graphics technology could benefit MELCO.

In collaboration with MTL researchers, he introduced these ideas into the MPEG 3D discussion in early 2003. Later that year, they introduced the ideas into MPEG-4, leading to the accepted contribution noted above. In support of the contribution, reference code was implemented and delivered to MPEG-4 in the summer of 2004.

Face Detection in the D-506i Cell Phone

In May 2004, MELCO started to ship a new camera-equipped NTT DoCoMo cell phone (model D-506i). This phone uses a fast face detection algorithm from MERL. When the user takes a picture with the phone, the face detector can be used to automatically generate thumbnail images of the faces in the picture. These can be stored as part of address book entries. Using caller ID, the picture of a person pops up when he or she calls. This is only one small feature of the phone, but it is an advertised point of differentiation between the D-506i and phones from other manufacturers.



Face found in phone photo

Phonebook entry with face

The D-506i is just one of what we hope will be many applications of fast face detection to MELCO business. In particular, the face detector and the basic approach underlying it have become a key foundation for a significant part of MERL's computer vision efforts. This includes extending the approach to the temporal domain to recognize actions in video (page 51), a fast and accurate method of face recognition (page 58) and analysis of surveillance data based on the faces shown in it (page 55).

Details: In late 2000, MRL embarked on an effort to apply its computer vision expertise to the observation of people in video, with the goal of contributing to MELCO business in the areas of surveillance and access control. This led in 2001 to the staffing of a computer vision application group in MTL as a complement to the existing research group in MRL. In 2002 it led to a major 3-year project on "Computer Human Observation".

The development of MERL's face detection algorithm began in 2001 as a collaboration between a newly hired researcher at MRL and a newly hired researcher at MTL based on work they had done in the previous year. Before their work, there were a number of accurate methods for finding faces in images, but they were computationally expensive and therefore slow. The key advance of their work was combining high accuracy on full frontal faces with by far the world's fastest speed.

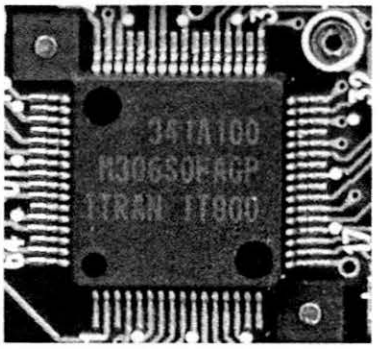
The algorithm operates in two parts. Computation-intensive computer learning techniques operating on large amounts of data are used to `train' a classifier that can determine whether or not a particular part of an image is a face. The image features used by the classifier are extremely simple, making it possible to evaluate the classifier very rapidly using a lightweight program. This allows real-time location of faces in video.

The initial version of the face detector was demonstrated to MELCO in the summer of 2001. In the ensuing year, the algorithm was redesigned, fundamentally improving the operation of both the classifier and the training algorithm. The system was also extended to operate on profile faces.

This improved face detector was demonstrated to MELCO in the summer of 2002. One of the people that saw this demonstration was a researcher at MELCO's Johosoken laboratory. He became interested in the detector and began to experiment with it.

The Johosoken researcher made the initial connection to the cell phone business unit. After discussion with MERL, he reimplemented MERL's classifier evaluation code for the cell phone platform and created a prototype application. His code, combined with a classifier trained by MERL, supports the D-506i.

The Renesas SCP/PLC Chip



In May 2004, Renesas Technology Corporation began production of the M603S, a single-chip Power Line Communication (PLC) IC for cost-effective smart home networking. The chip combines ITRAN Communications Ltd's IT800 power line modem and Renesas's M16C microprocessor. The M16C supports Microsoft's UPnP-compatible Simple Control Protocol (SCP) by means of microcode provided by MERL (page77).

MERL continues involvement in a range of projects related to home networking. This includes work on an SCP/ZigBee bridge (page 78), contributions to the ZigBee standard (pages 70-72) and work on broadband power line communications for HomePlug Audio/Video.

Details: One promising approach to low data rate networking of control information inside homes is to use PLC over existing electrical wiring. Microsoft's SCP is a lightweight alternative to the Internet Protocol (IP) for power-line communicating in support of Microsoft's Universal Plug and Play (UPnP) home networking standard.

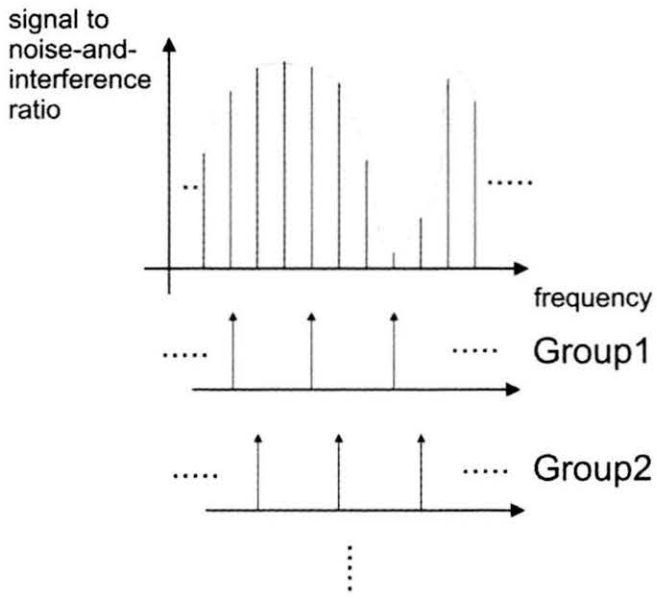
Motivated by the prospect of high volume sales of interface chips, interest in various kinds of home networking by MELCO's semiconductor business unit (Hanpon) goes back many years. This interest continues unabated following the merger of most of MELCO's semiconductor operations with most of Hitachi's semiconductor operations to form Renesas in 2003.

MTL has a long history of working with Hanpon/Renesas. One of the labs that were eventually merged into MTL was founded in 1993 to support a digital TV chip-set project. MTL also has a long history of involvement in home networking including work with standards like IEEE 1394, Havi and HomeAPI. This included working with Hanpon on an interface chip for the Japanese PLC standard Echonet in 2000.

As a result, it was only natural that Hanpon turned to MERL for help when it put together a project with Microsoft and ITRAN to create an SCP/PLC chip. In 2001, MERL experimented with SCP at Hanpon's request. In 2002-03, MERL created the microcode for the M603S at Hanpon/Renesas's request. This was done in close collaboration with a semiconductor lab in Japan, which did all the hardware design of the chip and Microsoft, which provided the highest-level code supporting SCP. MERL wrote the main body of microcode, which provides the support environment expected by Microsoft's code.

Symbol Spreading in MBOA Standard

In May 2004, MERL's contribution "Symbol Spreading Technique" was included in draft v0.8 of the Multi-Band OFDM Alliance (MBOA) standard for Ultra Wide Band (UWB) wireless communication. By using multiple OFDM subcarriers to encode each symbol, MERL's technology modifies the MBOA system to exploit the frequency diversity inherent in a UWB channel thereby reducing transmission errors caused by interference and fading (page 67).



MERL's Symbol Spreading work is just one of a range of efforts at MERL on wireless communications in general (pages 66-78) and Ultra-Wide Band (UWB) communications in particular (pages 66-67). A particular focus of our work is on contributions to standards.

Details: UWB was developed about 15 years ago, mainly for military purposes. The prime attraction of UWB to the military is that it is a spread-spectrum system with a large spreading factor. This results in a low power spectral density (which makes it difficult to detect) and a high interference rejection capability (which makes it almost impossible to jam). Furthermore, the

different frequency components inherent in a UWB signal provide a high frequency diversity and thus high reliability.

Recently, UWB has been approved for civilian use in the USA at extremely low power levels. At these levels, UWB can either support very high data rates (> 100Mbit/s) for short range communications (< 10m), or low data rates for longer ranges. High data rate communications are attractive for so called Personal Area Networks (PANs), e.g., for communicating high data rate information such as video within a home or office.

In late 2002, MERL began an initiative within MELCO on UWB. In collaboration with MELCO's Johosoken laboratory, MERL pursued a number of efforts including participating in the IEEE 802.15.3a PAN standard effort. The goal of 802.15.3a is short range (3-10 meter) high data rate (100-400 megabits per second) communication using UWB.

Two major groups have emerged within 802.15.3a. At MERL's suggestion, MELCO joined with six other companies in June 2003 to found one of these groups, the Multi-Band OFDM Alliance (MBOA), which has since grown to more than 150 members. (It is difficult at best to influence an important standard by going it alone. One needs allies.) MERL then focused its work within 802.15.3a on the MBOA proposal.

As sometimes happens, the 802.15.3a process became deadlocked due to a 60/40 split between the MBOA and the other major power group. (A 3/4 majority is needed for either approach to succeed.) As a result, the MBOA decided to form a special interest group by itself and put forward a standard before the end of 2004 along with compliant products. As noted above, MERL has been successful in making an impact on the MBOA standard.

The essence of the the MBOA proposal is to divide the available spectrum into several subbands of 500MHz width each and use high-speed pseudo-random switching between subbands to allow multiple simultaneous transmissions with minimal interference. Within the current subband, Orthogonal Frequency Division Multiplexing (OFDM) is used to communicate 100 streams of data simultaneously using 100 difference subfrequencies. The frequencies are close together and interfere with each other, but the modulation used for the individual data streams are chosen to be mathematically `orthogonal' so the data streams can nevertheless be reliably decoded by a receiver.

A problem with the above approach is that each individual OFDM subfrequency is subject to fading and interference from other sources. This can lead to errors in individual data streams even when other streams are communicated clearly. The essence of MERL's proposal is to mix the data streams together so that multiple subfrequencies are used for each data stream (and each subfrequency combines information from several data streams). This allows realiable communication of every data stream even if communication on a few subfrequencies is blocked.

Technical Staff

By far the most important assets of MERL are its people. The following pages present the capabilities and interests of MERL's technical staff members. More detailed information about the current activities of the technical staff can be seen by looking at the publications list and project descriptions elsewhere in this report. Further information can be obtained by looking at people's individual web pages on "http://www.merl.com/people".

Shmuel Avidan *Ph.D., Hebrew University, Jerusalem, Israel, 1999*
Research Scientist MRL

Avidan's research focus is computer vision with occasional detours into computer graphics and machine learning. He joined MERL in 2004 after three years at MobilEye, where he developed detection and tracking algorithms for vision-based driver assistance systems. In addition, he held a faculty position at the Interdisciplinary Center, Herzlya, Israel. Previously, he worked at Microsoft Research on modeling environments from collections of images.

Ali Azarbayejani *Ph.D., Massachusetts Institute of Technology, 1997*
Principal Technical Staff MTL

Azarbayejani's thesis was on computer-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 research. There, he led the development of new markets in the film and video post-production industry for vision-based software. In 2003, he joined MERL with interests in technology, software, and business development.

Paul A. Beardsley *Ph.D., Oxford University, 1992*
Senior Research Scientist MRL

Beardsley's thesis work was on applications of projective geometry to 3D recovery from images. His current focus is on 3D scanning, stereo vision for surveillance particularly looking at depth and 3D connectivity cues to aid segmentation of individuals in a crowd, and hand-held projectors together with novel modes of use. In support of a range of vision research at MERL, he is working on the Diamond3D vision library.

Ghulam M. Bhatti *Ph.D., Boston University, 1998*
Member Technical Staff MTL

For his thesis, Bhatti specialized in distributed and parallel discrete event simulation. Before joining MERL in 2000, he worked as a Sr. Software Engineer at Evare LLC, Inc, developing software for a network switch and implementing an RSA cryptographic scheme. He also worked at Excel Tech. Ltd. (XLTEK) developing embedded software for a portable EEG device. Currently, he is working on Home Networking and Digital TV.

Matthew E. Brand *Ph.D., Northwestern University, 1994*
Senior Research Scientist MRL

Brand studies unsupervised learning from sensory data. His results include spectral solutions for reconstructing manifolds from samples, decision-theoretic elevator group control, a linear-time online SVD, video-realistic synthesis of humans' recovery of non-rigid 3D shape from ordinary video, and an entropy optimization framework for learning. He has received best paper awards in computer vision (CVPR2001) and scheduling (ICAPS2003).

Dirk Brinkman *J.D., Suffolk University Law School, 1990*
Patent Counsel

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 I. Cukier *M.Sc., Polytechnic Institute of New York, 1985*
Senior Principal Technical Staff MTL

Cukier joined MERL in 1996. His initial focus was on digital systems for CATV, RF microwave transmitters & receivers, and front-ends for 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

Prior to becoming a lawyer, Curtin received his B.S. in Marine Engineering from the Massachusetts Maritime Academy and spent six years as an engineering officer aboard U.S. flag merchant ships engaged in worldwide trade. Before turning his attention to patent law and joining MERL in 2001, he spent two years as an attorney in private practice.

Paul H. Dietz *Ph.D., Carnegie Mellon University, 1995*
Principal Technical Staff MTL

Before joining MERL in 2000, Dietz 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. At MERL, Paul has been leading efforts developing new user interface technologies.

Ajay Divakaran *Ph.D., Rensselaer Polytechnic Institute, 1993*
Team Leader / Senior Principal Technical Staff MTL

Divakaran was an Assistant Professor with the Department of Electronics and Communications Engineering, University of Jodhpur, India, in 1985-86. He was a Scientist with Iterated Systems Inc., Atlanta, GA from 1995 to 1998. He joined MERL in 1998 and was an active contributor to the MPEG-7 video standard. His current research interests include video analysis, summarization, indexing, compression, and related applications

Alan W. Esenther *M.Sc., Boston University, 1993*
Principal Technical Staff MTL

Esenther enjoys human-computer interaction (HCI) design, distributed software development, graphical user interfaces and Internet technologies. His recent work has focused on 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).

George Fang *B.Sc., California Institute of Technology, 1990*
Member Technical Staff MTL

After graduating from college, Fang took a job at MELCO's Kyoto Works. During 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. After joining MERL in 2001, his research has focused on wireless and network technologies.

James Fang *B.Sc., Columbia University, 1992*
Member Technical Staff MTL

Fang did some graduate work at Columbia before joining MELCO's US TV operation in 1995. He worked on consumer televisions for three years before transferring to MERL in 1998. He is currently working on digital wireless communications.

Clifton L. Forlines *Master of HCI, Carnegie Mellon University, 2001*
Research Associate MRL

Forlines' research interests include the design and evaluation of novel user interfaces. His 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 user evaluation of three projects, MediaFinder, TimeTunnel, and DiamondSpin.

James L. Frankel *Ph.D., Harvard University, 1983*
Consulting Scientist MRL

Frankel's interests lie primarily in the areas of computer systems (architecture, operating systems, and distributed computing) and their application to ubiquitous computing devices. He was the principal designer of the C* language at Thinking Machines Corporation and was central to many other projects there.

Sarah F. Frisken *Ph.D., Carnegie Mellon University, 1991*
Senior Research Scientist MRL

Frisken (formerly Gibson) has research interests in computer graphics, volume visualization and physically based modeling. Her current work is with Adaptively Sampled Distance Fields (ADFs), a general representation of shape for computer graphics, which provides intuitive manipulation, and editing of smooth surfaces with fine detail. Applications of ADFs include digital sculpting, 3D design, color gamut representation and font rendering.

Daqing Gu *Ph.D., State University of New York at Stony Brook, 1996*
Principal Technical Staff MTL

After joining MERL in 1999, Gu 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 MTL

Guo worked at Waterloo Maple for a year and a half as a software developer before joining MERL in 1998. He primary research interests include home networks, digital broadcasting, and wireless computing.

Bret A. Harsham *Massachusetts Institute of Technology*
Principal Technical Staff MTL

Harsham joined MERL in 2001 to pursue interests in speech 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. Harsham's other technical interests include distributed architectures, knowledge representation and language theory.

Frederick J. Igo, Jr. *B.A., LeMoyne College, 1982*
Senior Principal Technical Staff MTL

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

Michael J. Jones *Ph.D., Massachusetts Institute of Technology, 1997*
Principal Technical Staff MTL

Jones joined MERL in the fall of 2001 after 4 years at the Digital/Compaq Cambridge Research Laboratory. His main area of interest is computer vision. 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 morph able models.

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

Kadel-Garcia is a member of MERL's 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 MTL

Kato went to work for MELCO in 1991. He worked on the design of compute servers based on Intel architecture chips and network security products as a hardware engineer at MELCO's Johosoken laboratory before accepting a multi-year assignment to MERL in 2002. His current research interest is in sensor networks, RFID tags and data mining.

Hao-Song Kong *Ph.D., Sydney University, 1998*
Member Technical Staff MTL

Kong has research interests in neural networks, digital signal processing, image processing, video transcoding, video transmission and networking. He worked at Motorola's Australian Research Center for 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 of wired and wireless IP networks.

Christopher H. Lee *Ph.D., Carnegie Mellon University, 2000*
Visiting Scientist MRL

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.

Jin Ho Lee *M.S., Yonsei University, 2004*
Research Associate MRL

Lee is in the midst of pursuing a PhD degree in the Department of Computer Science and Engineering at Ohio State University. At MERL, he is participating in a 3D face recognition project and a face reflectance measurement project. His research interests include computer graphics, scientific visualization and computer vision.

Darren L. Leigh *Ph.D., Harvard University, 1998*
Senior Research Scientist MRL

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). His current research includes DiamondTouch multi-user touch technology and sensor networks.

Neal B. Lesh *Ph.D., University of Washington, 1998*
Senior Research Scientist MRL

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 the collaborative navigation of digital data (the Personal Digital Historian project). Before coming to MERL, he worked as a postdoc with James Allen at the University of Rochester.

Sergei Makar-Limanov *Ph.D., Stanford University, 1994*
Principal Technical Staff MTL

After spending a few years in academia, Makar-Limanov decided to join the "real world" where he worked on software projects such as computer aided manufacturing for PTC corporation, supply management for Kewill PLC, and automated software testing tools at Empirix Corporation. Sergei joined MERL in 2001 to work on the Concordia project. He is now working with the vision group on creating a database framework for video surveillance applications.

Wojciech Matusik *Ph.D., MIT, 2003*
Visiting Scientist MRL

Matusik is a Consulting Scientist at MERL, as well as a Visiting Scientist in the Computer Science and Artificial Intelligence Laboratory at MIT. His primary research lies in computer graphics with an emphasis on modeling based on measured data.

Neelesh B. Mehta *Ph.D., California Institute of Technology, 2001*
Member Technical Staff MTL

Mehta worked at AT&T Research Labs (Wireless Systems Group) and Broadcom before joining MERL's digital communications group at MERL. His areas of interest include physical layer communication technologies such as MIMO, MIMO-OFDM, link adaptation techniques, multiple access techniques, and system performance evaluation studies of 3G systems.

Koji Miyahara *M.Sc., Kyushu University, 1988*
Senior Principal Technical Staff MTL

Miyahara joined MELCO in 1988. He worked at MELCO's Johosoken laboratory before accepting a multi-year assignment to MERL in 2002. He was a visiting researcher at the University of California, Irvine, from 1999 to 2000. His research interests include user interfaces, intelligent agents and information filtering.

Baback Moghaddam *Ph.D., Massachusetts Institute of Technology, 1997*
Senior Research Scientist MRL

Moghaddam's research interests are in computational vision with a focus on probabilistic visual learning, statistical modeling and pattern recognition with applications in biometrics and computer-human interface. While at MIT he developed a fully automatic vision system, which won DARPA's 1996 "FERET" Face Recognition Competition.

Andreas F. Molisch *Ph.D., Technical University Vienna, 1994*
Senior Principal Technical Staff MTL

Molisch's current research interests are multiple-antenna systems, wireless channel measurement and modeling, ultra wideband 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.

Yves-Paul N. Nakache *M.Sc., E.S.I.E.E., 2000*
Member Technical Staff MTL

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

Daniel N. Nikovski *Ph.D., Carnegie Mellon University, 2002*
Research Scientist MRL

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 and traffic prediction. He also has varied interests in the field of data mining.

Philip V. Orlik *Ph.D., State University of New York at Stony Brook, 1999*
Principal Technical Staff MTL

Orlik joined MERL's digital communications and networking group in 2000. His research interests include wireless and optical communications, networking, queuing theory, and analytical modeling.

Kadir A. Peker *Ph.D., New Jersey Institute of Technology, 2001*
Member Technical Staff MTL

Peker finished his thesis on content-based video indexing and summarization using motion activity while working at MERL. At MERL, he has also worked on home networks and multimedia networking, attending UPnP working groups. His current research interests include video indexing, browsing and summarization, video presentation techniques, and video mining.

Georgiy Pekhteryev *M.Sc., Kharkiv Aviation Institute, Ukraine, 1994*
Member Technical Staff MTL

Pekhteryev joined MERL in 2002, where he has applied his software expertise to a range of projects. His current interests are focused on wireless and wired home networking, network technologies.

Ronald N. Perry *B.Sc., Bucknell University, 1981*
Senior Research Scientist MRL

Prior to joining MERL in 1998, Perry 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 are centered on key algorithms in computer graphics.

Hanspeter Pfister *Ph.D., State University of NY at Stony Brook, 1996*
Associate Director / Senior Research Scientist MRL

Pfister was 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. Pfister was the general chair of the IEEE Visualization 2002 conference in Boston.

Erik A. Piip
Computer Network Services Manager

PiiP 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 DEC in multiple roles; service delivery, support, fault management and analysis, and server management design.

Fatih M. Porikli *Ph.D., Polytechnic University, 2002*
Member Technical Staff MTL

Porikli's 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. Before I joined to MERL in 2000, he worked for Hughes Research Labs, Malibu, CA (1999) and AT&T Research Labs, Holmdel, NJ (1997).

Stanley W. Pozerski *BA Computer Systems, Daniel Webster College, 1987*
Systems Network Administrator

Pozerski'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 personal computers for production testing & manufacturing of chemicals and controlling multi-axis rotary assembly machines. Currently, Stan supports Windows and Linux clients, servers, networking, and the wide variety of PC applications used at MERL.

Bhiksha Raj *Ph.D., Carnegie Mellon University, 2000*
Research Scientist MRL

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. Prior to joining MERL, Raj worked at Compaq's Cambridge MA research lab.

Ramesh Raskar *Ph.D., University of North Carolina at Chapel Hill, 2002*
Research Scientist MRL

Raskar joined MERL in 2000. Prior to that, he was in 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. His current work includes topics from non-photorealistic rendering, computer vision and intelligent user interfaces.

Charles Rich *Ph.D., Massachusetts Institute of Technology, 1980*
Distinguished Research Scientist MRL

The long-term focus of Rich's research is making interacting with a computer more like interacting with a person. As co-founder and co-director of the Programmer's project at the MIT Artificial Intelligence Lab. in the 1980s, he pioneered research on intelligent assistants for software engineering. For the past several years, he has been working on a technology, called Collagen, for building collaborative interface agents based on human discourse theory.

David C. Rudolph *M.S., University of Illinois, 1989*
Principal Technical Staff MTL

Since joining MERL in 1990, Rudolph has contributed to several systems software projects, including 6 years spent in the "Network Replication" project, which is now being successfully marketed by Veritas Software Corporation as VVR (Veritas Volume Manager). Before joining MERL, David spent 3 years at Data General. His M.S. research focused on system software and performance analysis for massively parallel architectures.

Kathleen Ryall *Ph.D., Harvard University, 1997*
Principal Technical Staff MTL

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. For 3 years before joining MERL, Kathy was an Assistant Professor of Computer Science at the University of Virginia.

Zafer Sahinoglu *Ph.D., New Jersey Institute of Technology, 2001*
Member Technical Staff MTL

Sahinoglu worked at AT&T Shannon Labs in 1999, and joined MERL in March 2001. His research interests include home networking, QoS in video streaming & multicasting, wireless image sensor networks, traffic self-similarity and biomedical signal processing. He has made significant contributions to the emerging MPEG-21 and ZigBee standards.

Bent K. Schmidt-Nielsen *B.S. University of California at San Diego, 1971*
Senior Principal Technical Staff MTL

Schmidt-Nielsen spent 7 years at Dragon Systems applying speech recognition to useful products. At MERL he is paying a lot of attention to making speech interfaces robust and usable. He 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 L. Schwenke *M.S., Worcester Polytechnic Institute, 1988*
Principal Technical Staff MTL

Before joining MERL in 1988, Schwenke worked at Raytheon on image processing and satellite communications systems. At MERL he worked on the design and simulation of CPU hardware and a wide range of software development projects including multi-user virtual reality, mobile agents on the Internet, and multi-modal interfaces. He is an active member of the W3C VoiceXML and Multimodal working groups.

Huai-Rong Shao *Ph.D., Tsinghua University, 1999*
Research Scientist MRL

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 Huai-Rong Shao worked at the Microsoft Research laboratories in Beijing and Redmond.

Chia Shen *Ph.D., University of Massachusetts, 1992*
Associate Director / Senior Research Scientist MRL

Shen's current research focuses on shared interactive surfaces. Previously, she led the MidART project, which MELCO has 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 E. Shipman *M.Sc., Carnegie Mellon University, 1985*
Principal Technical Staff MTL

Shipman's technical interests and background are in real-time and distributed operating systems research and development. At MERL, he has worked on network filer Replication and multi-user virtual reality projects, as well as on smaller efforts related to MPEG-7 and fingerprint recognition.

Candace L. Sidner *Ph.D., Massachusetts Institute of Technology, 1979*
Senior Research Scientist MRL

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 was a researcher at Bolt Beranek Newman, Digital Equipment Corp., and Lotus Development Corp. In addition, she was a visiting scientist at Harvard University and past President of the Association for Computational Linguistics.

Paris Smaragdis *Ph.D., Massachusetts Institute of Technology, 2001*
Research Scientist MRL

Paris Smaragdis joined MERL in 2002. His main interests are auditory scene analysis and self-organizing computational perception. Before coming to MERL he was a postdoctoral associate at MIT. His most recent work has been on sound source separation, multimodal statistics and audio classification.

Jay E. Thornton *Ph.D., University of Michigan, 1982*
Group Manager MTL

Thornton worked at Polaroid Corporation for many years, first in the Vision Research Laboratory and then as manager of the Image Science Laboratory. There, he worked on problems in color reproduction, image quality and image processing. He joined merl in 2002 as manager of 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*
Member Technical Staff MTL

Van Baar's interests are in the fields of Computer Graphics, Scientific Visualization, Computer Vision and HCI. He first came to MERL as an intern in 1997. He joined MERL full-time in 1999. 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*
Team Leader / Senior Principal Technical Staff MTL

Anthony Vetro joined MERL in 1996. 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 K. Woelfel *M.S., Rutgers University, 1992*
Principal Technical Staff MTL

Before joining MERL in February 2001, Joe worked at Dragon Systems, where he led small teams developing extensible voice architecture. In the years before that, Joe worked on the development of a statistical process control software package at GE-Fanuc. At MERL, his efforts have been focused on the application of computer vision technology to surveillance.

Peter P. Wolf *B.S., Yale University, 1983*
Senior Principal Technical Staff MTL

Wolf is an expert in Speech Technologies and a broad range of Software Engineering tools and practices. While his 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. Wolf is currently exploring the use of speech recognition to retrieve information with applications for cell phones, PDAs, automobiles and home entertainment.

David W. H. Wong *Ph.D., University of Connecticut, 1991*
Group Manager MTL

David Wong is the manager of MTL's Human-Centric and Analytic Systems, a group. His technical interests include various flavors of agent technology, data mining and visualization techniques, as well as 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.

Christopher R. Wren *Ph.D., Massachusetts Institute of Technology, 2000*
Research Scientist MRL

Wren's research area is Perception for Human-Computer Interaction. While his 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. As part of his thesis work at MIT, he developed a system for combining physical models with visual evidence in real time to recover subtle models of human motion.

Jun Xin *Ph.D., University of Washington, 2002*
Member Technical Staff

Xin joined MERL in 2003. His research interests include digital video processing and multimedia communication. His current focus is on video transcoding, video codec optimization and next generation video compression schemes.

Jonathan S. Yedidia *Ph.D., Princeton University, 1990*
Research Scientist MRL

Yedidia's graduate work focused on theoretical condensed-matter physics, particularly the statistical mechanics of systems with quenched disorder. In 1997, he changed his focus to computer software and worked for a company called Viaweb on a shopping search engine, which has since become Yahoo's shopping service. At MERL since 1998, Yedidia's particularly interest is in the development of new methods belief propagation in constraint networks.

William S. Yerazunis *Ph.D., Rensselaer Polytechnic Institute, 1987*
Senior Research Scientist MRL

Yerazunis has worked in a number of fields including: optics, vision processing, and signal processing, computer graphics, artificial intelligence parallel symbolic computation, radio astronomy and SETI, transplant immunology, virtual and augmented reality (Diamond Park and SPLINE), real-time sensing and ubiquitous computing, and real-time statistical categorization of text (for spam filtering).

Fangfang Zhang *M.S., Brandeis University, 2000*
Member Technical Staff MTL

Prior to joining MERL, Zhang worked at CMGI, as a Software Engineer for the web-dialup service development team. At MERL, she has worked on various projects including the Concordia Project and the Speech Project. She is currently a member of the Computer Human Observation Project.

Jinyun Zhang *Ph.D., University of Ottawa, 1991*
Group Manager MTL

Zhang manages MTL's digital communication and networking group. Before joining MERL in 2001, Zhang worked for Nortel Networks for 10 years where she held engineering and 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 and real-time embedded S/W for wireless communications.

Recent Major Publications

The following lists the 148 major publications by members of the MERL staff over the past year. (This is an average of more than 2.2 papers per technical staff member). 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

An asterisk (*) appears before the 62 (42%) 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.

* Palanki, R.; Yedidia, J.S., "Rateless Codes on Noisy Channels", *IEEE International Symposium on Information Theory (ISIT)*, June 2004 (TR2004-038)

Radhakrishnan, R.; Xiong, Z.; Divakaran, A.; Memon, N., "Time Series Analysis and Segmentation Using Eigenvectors for Mining Semantic Audio Label Sequences", *IEEE International Conference on Multimedia and Expo (ICME)*, June 2004 (TR2004-063)

Xiong, Z.; Radhakrishnan, R.; Divakaran, A.; Huang, T.S., "Effective and Efficient Sports Highlights Extraction Using the Minimum Description Length Criterion in Selecting GMM Structures", *IEEE International Conference on Multimedia and Expo (ICME)*, June 2004 (TR2004-061)

Divakaran, A.; Vetro, A.; Kan, T., "Towards Maximizing the End-User Experience", *IEEE International Conference on Multimedia and Expo (ICME)*, June 2004 (TR2004-055)

Kong, H-S; Nie, Y.; Vetro, A.; Sun, H.; Barner, K.E., "Coding Artifact Reduction Using Edge Map Guided Adaptive and Fuzzy Filter", *IEEE International Conference on Multimedia and Expo (ICME)*, June 2004 (TR2004-056)

Peker, K.A.; Divakaran, A., "Adaptive Fast Playback-Based Video Skimming Using a Compressed-Domain Visual Complexity Measure", *IEEE International Conference on Multimedia and Expo (ICME)*, June 2004 (TR2004-060)

* Porikli, F.M.; Haga, T., "Event Detection by Eigenvector Decomposition Using Object and Frame Features", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2004 (TR2004-029)

  Feris, R.; Turk, M.; Raskar, R.; Tan, K-H; Ohashi, G., "Exploiting Depth Discontinuities for Vision-Based Fingerspelling Recognition", *IEEE Workshop on Real-Time Vision for Human-Computer Interaction (RTV4HCI)*, June 2004 (TR2004-99)

  Yedidia, J.S., "Sparse Factor Graph Representations of Reed-Solomon and Related Codes", *IEEE International Symposium on Information Theory (ISIT)*, June 2004 (TR2004-097)

  Brand, M.; Kang, K.; Cooper, D.B., "Algebraic Solution for the Visual Hull", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), June 2004 (TR2004-101)

* Matusik, W.; Loper, M.; Pfister, H., "Progressively-Refined Reflectance Functions from Natural Illumination", *Eurographics Symposium on Rendering (EGSR)*, June 2004 (TR2004-065)

* Ma, Z.; Shao, H-R; Shen, C., "A New Multi-path Selection Scheme for Video Streaming on Overlay Networks", *IEEE International Conference on Communications (ICC)*, June 2004 (TR2004-064)

* Sahinoglu, Z.; Catovic, A., "A Hybrid Location Estimation Scheme (H-LES) for Partially Synchronized Wireless Sensor Networks", *IEEE International Conference on Communications (ICC)*, June 2004 (TR2003-142)

* Gezici, S.; Molisch, A.F.; Poor, H.V.; Kobayashi, H., "The Trade-off Between Processing Gains of Impulse Radio Systems in the Presence of Timing Jitter", *IEEE International Conference on Communications (ICC)*, June 2004 (TR2004-015)

* Gezici, S.; Kobayashi, H.; Poor, H.V.; Molisch, A.F., "Effect of Timing Jitter on the Trade-off Between Processing Gains", *IEEE International Conference on Communications (ICC)*, June 2004 (TR2003-109)

  Porikli, F.M.; Li. X, "Traffic Congestion Analysis in Compressed Video Without Tracking", *IEEE Intelligent Vehicles Symposium (IV)*, June 2004 (TR2004-019)

* Raskar, R.; Ilie, A.; Yu, J., "Image Fusion for Context Enhancement and Video Surrealism", *International Symposium on Non-Photorealistic Animation and Rendering (NPAR)*, ISBN: 1-58113-887-3, pp. 85-152, June 2004 (TR2004-039)

* Brown, J.C.; Smaragdis, P., "Independent Component Analysis for Automatic Note Extraction from Musical Trills", *Journal of the Acoustical Society of America*, Vol. 115, Issue 5, pp. 1851-2634, May 2004 (TR2004-078)

  Li, Y.; Molisch, A.F.; Zhang, J., "Practical Approaches to Channel Estimation and Interference Suppression for OFDM Based UWB Communications", *IEEE CAS Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communications (MWC)*, May 2004 (TR2004-095)

* Dietz, P.H.; Raskar, R.; Mihelic-Booth, S.; van Baar, J.; Wittenburg, K.B.; Knep B., "Multi-Projectors and Implicit Interaction in Persuasive Public Displays", *Advanced Visual Interfaces (AVI)*, ISBN: 1-58113-867-9, pp. 209-217, May 2004 (TR2004-021)

* Lesh, N.B.; Mitzenmacher, M., "Interactive Data Summarization: An Example Application", *Advanced Visual Interfaces (AVI)*, ISBN: 1-58113-867-9, pp. 183-187, May 2004 (TR2004-094)

Xia, M.; Vetro, A.; Liu, B.; Sun, H., "Rate-Distortion Optimized Bit Allocation for Error-Resilient Transcoding", *IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2004 (TR2004-047)

Raskar, R.; van Baar, J.; Beardsley, P., "Display Grid: Ad-hoc Clusters of Autonomous Projectors", *Society for Information Displays International Symposium (SID)*, May 2004 (TR2004-033)

Kong, H-S; Vetro, A.; Sun, H., "Edge Map Guided Adaptive Post-Filter for Blocking and Ringing Artifacts Removal", *IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2004 (TR2004-003)

Kung, W-Y; Kong, H-S; Vetro, A.; Sun, H., "Error Resilient Methods for Real-Time MPEG-4 Video Streaming", *IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2004 (TR2004-049)

Nakache, Y-P N.; Orlik, P.V.; Gifford, W.M.; Molisch, A.F.; Ramachandran, I.; Fang, G.; Zhang, J., "Low-Complexity Ultra Wideband Transceiver with Compatibility to Multiband - OFDM", *Conference on Ultra Wideband Systems and Technologies (UWBST)*, May 2004 (TR2004-051)

Gezici, S.; Kobayashi, H.; Poor, H.V.; Molisch, A.F., "Optimal and Suboptimal Linear Receivers for Time-Hopping Impulse Radio Systems", *Conference on Ultra Wideband Systems and Technologies (UWBST)*, May 2004 (TR2004-052)

Zhou, S.; Chellappa, R.; Moghaddam, B., "Intra-Personal Kernel Space for Face Recognition", *IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, pp. 235-240, May 2004 (TR2004-025)

Tufvesson, F.; Molisch, A.F., "Ultra-Wideband Communication Using Hybrid Matched Filter Correlation Receivers", *IEEE Semiannual Vehicular Technology Conference (VTC)*, May 2004 (TR2004-016)

Minnen, D.C.; Wren, C.R., "Finding Temporal Patterns by Data Decomposition", *IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, May 2004 (TR2004-054)

Raj, B.; Singh, R.; Stern R.M., "On Tracking Noise with Linear Dynamical System Models", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2004 (TR2004-042)

Lee, J.; Moghaddam, B.; Pfister, H.; Machiraju, R., "Finding Optimal Views for 3D Face Shape Modeling", *IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, pp. 31-36, May 2004 (TR2004-024)

DeVault, D.; Rich, C.; Sidner, C.L., "Natural Language Generation and Discourse Context: Computing Distractor Sets from the Focus Stack", *International Florida Artificial Intelligence Research Symposium (FLAIRS)*, May 2004 (TR2004-004)

* Porikli, F.M., "Trajectory Distance Metric Using Hidden Markov Model Based Representation", *European Conference on Computer Vision (ECCV)*, May 2004 (TR2004-030)

* Brand, M.; Antone, M.; Teller, S., "Spectral Solution of Large-Scale Extrinsic Camera Calibration as a Graph Embedding Problem", *European Conference on Computer Vision (ECCV),* May 2004 (TR2004-100)

* Porikli, F.M., "Trajectory Pattern Detection by HMM Parameter Space features and Eigenvector Clustering", *European Conference on Computer Vision (ECCV)*, May 2004 (TR2004-032)

Lau, K.T.; Baldwin, S.; Shepherd, R.L.; Dietz, P.H.; Yerazunis, W.S.; Diamond, D., "Novel Fused-LEDs Devices as Optical Sensors for Colorimetric Analysis", *Talanta,* Vol. 63, Issue 1, pp.167-173, May 2004 (TR2003-150)

Divi, V.; Forlines, C.; van Gemert, J.V.; Raj, B.; Schmidt-Nielsen, B.; Wittenburg, K.; Woelfel, J.; Wolf, P.; Zhang, F.,, "A Speech-In List-Out Approach to Spoken User Interfaces", *Human Language Technology Conference*, May 2004 (TR2004-023)

* Brand, M.E.; Nikovski, D.N., "Risk-Averse Group Elevator Scheduling", *Elevcon World Congress on Vertical Transportation Technologies (ELEVCON)*, April 2004 (TR2004-045)

* Brand, M.E.; Nikovski, D.N., "Optimal Parking in Group Elevator Control", *IEEE International Conference on Robotics & Automation (ICRA)*, April 2004 (TR2004-044)

* Shen, C.; Vernier, F.D.; Forlines, C.; Ringel, M., "DiamondSpin: An Extensible Toolkit for Around-the-Table Interaction", *ACM Conference on Human Factors in Computing Systems (CHI)*, ISBN: 1-58113-702-8, pp. 167-174, April 2004 (TR2004-005)

* Ringel, M.; Ryall, K.; Shen, C.; Forlines, C.; Vernier, F., "Release, Relocate, Reorient, Resize: Fluid Techniques for Document Sharing on Multi-User Interactive Tables", *ACM Conference on Human Factors in Computing Systems (CHI)*, ISBN: 1 58113 703 6, pp. 1441-1444, April 2004 (TR2004-022)

* Lee, C.H.; Lesh, N.B.; Sidner, C.L.; Morency, L-P; Kapoor, A.; Darrrell, T., "Nodding in Conversations with a Robot", *ACM Conference on Human Factors in Computing Systems (CHI)*, ISBN: 1-58113-703-6, pp. 785-786, April 2004 (TR2004-018)

Vetro, A.; Sun, H., "Minimum Distortion Sampling of Multiple Image Sequences", *International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*, April 2004 (TR2004-001)

Lesh, N.B.; Marks, J.; McMahon, A.; Mitzenmacher, M., "Exhaustive Approaches to 2D Rectangular Perfect Packings", *Information Processing Letters*, Vol. 90, Issue 1, pp. 7-14, April 2004 (TR2004-093)

* Molisch, A.F.; Win, M.Z., "MIMO Systems with Antenna Selection", *IEEE Microwave Magazine*, ISSN: 1527-3342, Vol. 5, Issue 1, pp. 46-56, March 2004 (TR2004-014)

* Dai, H.; Molisch, A.F.; Poor, H.V., "Downlink Capacity of Interference-Limited MIMO Systems with Joint Detection", *IEEE Transactions on Wireless Communications*, ISSN: 1536-1276, Vol. 3, Issue 2, pp. 442-453, March 2004 (TR2004-069)

* Porikli, F.M.; Wang, Y., "Automatic Video Object Segmentation Using Volume Growing and Hierarchical Clustering", *EURASIP Journal on Applied Signal Processing*, ISSN: 1536-1276, Vol. 3, Issue 2, pp. 442-453, March 2004 (TR2004-012)

Gezici, S.; Kobayashi, H.; Poor, H.V.; Molisch, A.F., "Performance Evaluation of Impulse Radio UWB Systems with Pulse-Based Polarity Randomization in Asynchronous Multi-user Environments", *IEEE Wireless Communications and Networking Conference (WCNC)*, March 2004 (TR2004-017)

* Palanki, R.; Yedidia, J.S., "Rateless Codes on Noisy Channels", *Conference on Information Sciences and Systems (CISS)*, March 2004 (TR2004-037)

Vetro, A., "MPEG-21 Digital Item Adaptation: Enabling Universal Multimedia Access", *IEEE Multimedia*, ISSN: 1070-986X, Vol. 11, Issue 1, pp. 84-87, January - March 2004 (TR2004-002)

* Molisch, A.F., "A Generic Model for MIMO Wireless Propagation Channels in Macro- and Micro Cells", *IEEE Transactions on Signal Processing*, ISSN: 1053-587X, Vol. 52, Issue 1, pp. 61-71, January 2004 (TR2004-013)

* Molisch, A.F.; Zhang, X., "FFT-Based Hybrid Antenna Selection Schemes for Spatially Correlated MIMO Channels", *IEEE Communication Letters*, ISSN: 1089-7798, Vol. 8, Issue 1, pp. 36-38, January 2004 (TR2004-070)

* Moghaddam, B.; Tian, Q.; Lesh, N.B.; Shen, C.; Huang, T. S., "Visualization and User-Modeling for Browsing Personal Photo Libraries", *International Journal of Computer Vision*, Vol. 56, Issue 1-2, pp.109-130, January 2004 (TR2004-026)

* Zhou, S.; Chellappa, R.; Moghaddam, B., "Appearance Tracking Using Adaptive Models in a Particle Filter", *Asian Conference on Computer Vision (ACCV)*, January 2004 (TR2004-027)

Ilie, A.; Raskar, R.; Yu, J., "Fixed Camera Surveillance Using Image Fusion", *Asian Conference in Computer Vision (ACCV)*, January 2004 (TR2004-020)

Yerazunis, Y.S., "The Spam-Filtering Accuracy Plateau at 99.9% Accuracy and How to Get Past It", *MIT Spam Conference*, January 2004 (TR2004-091)

Asai, K.; Nishikawa, H.; Kudo, D.; Divakaran, A., "MPEG-7 Meta-data Enhanced Encoder System for Embedded Systems", *SPIE Conference on Visual Communications and Image Processing (VCIP)*, Vol. 5308, pp. 119-123, January 2004 (TR2004-009)

Goh, K-S; Miyahara, K.; Radhakrishan, R.; Xiong, Z.; Divakaran, A., "Audio-Visual Event Detection Based on Mining of Semantic Audio-Visual Labels", *SPIE Conference on Storage and Retrieval for Multimedia Databases*, Vol. 5307, pp. 292-299, January 2004 (TR2004-008)

Porikli, F.M., "Real-time Video Object Segmentation for MPEG-encoded Video Sequences", *SPIE Conference on Real-Time Imagining VIII*, Vol. 5297, pp. 195-203, January 2004 (TR2004-011)

Divakaran, A.; Miyaraha, K.; Peker, K.A.; Radhakrishnan, R.; Xion, Z., "Video Mining using Combinations of Unsupervised and Supervised Learning Techniques", *SPIE Conference on Storage and Retrieval for Multimedia Databases*, Vol. 5307, pp. 235-243, January 2004 (TR2004-007)

Porikli, F.M., "Automatic Image Segmentation by Solving Eikonal Equation Based on Gaussian Mixture Models", *IS&T / SPIE Symposium on Electronic Imaging: Science & Technology*, January 2004 (TR2004-010)

* Sidner, C.L.; Kidd, C.D.; Lee, C.H.; Lesh, N.B., "Where to Look: A Study of Human-Robot Engagement", *ACM International Conference on Intelligent User Interfaces (IUI)*, ISBN: 1-58113-815-6, pp. 78-84, January 2004 (TR2003-123)

Sidner, C.L.; Dzikovska, M., "A First Experiment In Engagement For Human-Robot Interaction In Hosting Activities", *Natural, Intelligent and Effective Interaction in Multimodal Dialogue Systems*, N.O. Bernsen, L. Dybkjaer, J. van Kuppevelt (eds), January, 2004 (TR2003-134)

Chen, D.; Gu, D.; Zhang, J., "Supporting Real-Time Traffic with QoS in IEEE 802.11e Based Home Networks", *Consumer Communications and Networking Conference (CCNC)*, pp. 205-209, January 2004 (TR2004-006)

* 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, December 2003 (TR2003-064)

* Molisch, A.F.; Foerster, J.R.; Pendergrass, M., "Channel Models for Ultra Wideband Personal Area Networks", *IEEE Wireless Communications*, ISSN: 1536-1284, Vol. 10, Issue 6, pp. 14-21, December 2003 (TR2004-071)

* Radhakrishan, R.; Xiong, Z.; Divakaran, A.; Ishikawa, Y., "Generation of Sports Highlights Using a Combination of Supervised & Unsupervised Learning in Audio Domain", *International Conference on Pacific Rim Conference on Multimedia*, Vol. 2, pp. 935-939, December 2003 (TR2003-144)

* 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, December 2003 (TR2004-081)

Yedidia, J.S., "Sparse Factor Graph Representations of Reed-Solomon and Related Codes", *DIMACS Workshop on Algebraic Coding*, December 2003 (TR2003-135)

Brand, M., "Minimax Embeddings ", *Neural Information Processing Systems (NIPS)*, December 2003 (TR2004-102)

* 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 Telecommunications Conference (GLOBECOM)*, Vol. 2, pp. 839-843, December 2003 (TR2003-078)

* Almers, P.; Tufvesson, F.; Molisch, A.F., "Keyhole Effects in MIMO Wireless Channels-Measurements and Theory", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 4, pp. 1781-1785, December 2003 (TR2003-136)

* Sahinoglu, Z.; Orlik, P.V., "Optimum Power Compensation for Error Propagation in Relay Assisted Wireless Networks", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 1, pp. 382-386, December 2003 (TR2003-141)

* Zhang, X.; Molisch, A.F.; Kung, S-Y, "Phase-Shift-Based Antenna Selection for MIMO Channels", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 2, pp. 1089-1093, December 2003 (TR2004-072)

* Molisch, A.F.; Win, M.Z.; Winters, J.H., "Reduced Complexity Transmit/Receive Diversity Systems", *IEEE Transactions on Signal Processing*, ISSN: 1053-587X, Vol. 51, Issue 11, pp. 2729-2738, November 2003 (TR2003-044)

  Sheng, H.; Haimovich, A.M.; Molisch, A.F.; Zhang, J., "Optimum Combining for Time Hopping Impulse Radio UWB Rake Receivers", *IEEE Conference on Ultra Wideband Systems and Technologies (UWBST)*, pp. 224-228, November 2003 (TR2004-073)

* Wittenburg, K.B.; Forlines, C.; Lanning, T.; Esenther, A.W.; Harada, S.; Miyachi, T., "Rapid Serial Visual Presentation Techniques for Consumer Digital Video Devices", *ACM Symposium on User Interface Software and Technology (UIST)*, ISBN: 1-58113-636-6, pp. 115-124, November 2003 (TR2003-019)

  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., October 2003 (TR2003-034)

  Raj, B.; Seltzer, M.L.; Reyes-Gomez, M.J., "Speech Recognizer Based Maximum Likelihood Beamforming", *NSF Workshop on Perspectives on Speech Separation*, October 2003 (TR2003-087)

* Messerges, T.S.; Cukier, J.I; Kevenaar, T.A.M.; Puhl, L.; Struik, R.; Callaway, E., "A Security Design for a General Purpose, Self-Organizing, Multihop Ad Hoc Wireless Network", *ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN)*, ISBN: 1-58113-783-4, pp. 1-11, October 2003 (TR2003-114)

  Xie, L.; Chang, S-F; Divakaran, A.; Sun, H., "Unsupervised Mining of Statistical Temporal Structures in Video ", *Video Mining*, Rosenfeld, A.; Doermann, D.; Dementhon, D., ISBN: 1402 075 499, October 2003 (TR2003-132)

  Gaganis, V.; Pasakakis, N.; Smaragdis, P.; Varotsis, N., "Deconvolution of Overlapping HPLC Peaks of Aromatic Hydrocarbons using Independent Component Analysis", *Conferentia Chemometrica (CC)*, October 2003 (TR2004-090)

* Raskar, R.; van Baar, J.; Beardsley, P.A.; Forlines, C.; Dietz, P.H.; Esenther, A.W.; Leigh, D.L; Ryall, K.; Shen, C.; Shipman, S.E.; Yerazunis, W.S., "Intelligent Clusters and Collaborative Projector-based Displays", *NSF Workshop on Collaborative Virtual Reality and Visualization (CVRV)*, October 2003 (TR2003-086)

  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)*, pp. 177-180, October 2003 (TR2003-139)

* Reyes-Gomez, M.J.; Raj, B.; Ellis, D.P.W., "Multi-Channel Source Separation by Beamforming Trained with Factorial HMMS", *IEEE Workshop on Applications of Signal Processing to Audio and Acoustic (WASPAA)*, pp. 13-16, October 2003 (TR2003-088)

  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 (TR2004-075)

  Li, Y.G.; Molisch, A.F.; Zhang, J., "Channel Estimation and Signal Detection for UWB", *International Symposium on Wireless Personal Multimedia Communications (WPMC)*, October 2003 (TR2003-074)

  Foerster, J.R.; Pendergrass, M.; Molisch, A.F., "A UWB Channel Model for Ultra Wideband Indoor Communications", *International Symposium on Wireless Personal Multimedia Communications (WPMC)*, October 2003 (TR2004-074)

  Molisch, A.F.; Nakache, Y-P N.; Orlik, P.V; 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)*, October 2003 (TR2003-072)

  Molisch, A.F.; Li, Y.G.; Nakache, Y-P; Orlik, P., Miyak, M.; Wu, Y.; Gezici, S.; Sheng, H.; Kung, S.Y.; Kobayaashi, H.; Poor, H.V.; Haimovich, A.; Zhang, J., "A Low-Cost Time Hopping Impulse radio system for High Date Rate", *International Symposium on Wireless Personal Multimedia Communications (WPMC)*, October 2003 (TR2003-129)

* Jones, M.J.; Viola, P., "Face Recognition Using Boosted Local Features", *IEEE International Conference on Computer Vision (ICCV)*, October 2003 (TR2003-025)

* Moghaddam, B.; Lee, J.H.; 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)*, pp. 20-27, October 2003 (TR2003-084)

* Levin, A.; Viola, P.; Freund, Y., "Unsupervised Improvement of Visual Detectors Using Co-Training", *IEEE International Conference on Computer Vision (ICCV)*, pp. 626-633, October 2003 (TR2003-127)

* Viola, P.; Jones, M.J.; Snow, D., "Detecting Pedestrians Using Patterns of Motion and Appearance", *IEEE International Conference on Computer Vision (ICCV)*, Vol. 2, pp. 734-741, October 2003 (TR2003-090)

* Shakhnarovich, G.; Viola, P.; Darrell, T., "Fast Pose Estimation With Parameter Sensitive Hashing", *IEEE International Conference on Computer Vision (ICCV)*, Vol. 2, pp. 750-757, October 2003 (TR2003-128)

* Shen, C.; Everitt, K.M.; Ryall, K., "UbiTable: Impromptu Face-to-Face Collaboration on Horizontal Interactive Surfaces ", *ACM International Conference on Ubiquitous Computing (UbiComp)*, October 2003 (TR2003-049)

* Wren, C.R.; Srinivasa, R., "Self-configuring, Lightweight Sensor Networks for Ubiquitous Computing", *International Conference on Ubiquitous Computing (UbiComp)*, pp. 205-206, September 2003 (TR2003-024)

* Dietz, P.H.; Yerazunis, W.S.; Leigh, D.L., "Very Low Cost Sensing and Communication Using Bidirectional LEDS", *International Conference on Ubiquitous Computing (UbiComp)*, October 2003 (TR2003-035)

  Sidner, C.L.; Lee, C.H.; Lesh, N.B., "Engagement Rules for Human-Robot Collaborative Interactions", *IEEE International Conference on Systems, Man & Cybernetics (CSMC)*, ISSN: 1062-922X , Vol. 4, pp. 3957-3962, October 2003 (TR2003-050)

* Martinian, E.; Yedidia, J.S., "Iterative Quantization Using Codes on Graphs", *Allerton Conference on Communications, Control, and Computing*, October, 2003 (TR2003-120)

* Sahinoglu, Z.; Vetro, A., "Mobility Characteristics for Multimedia Service Adaptation", *Signal Processing: Image Communication*, Vol. 18, Issue 8, pp. 699-719, September 2003 (TR2003-053)

* Molisch, A.F.; Petrovic, R., "Reduction of the Error Floor of Binary FSK by Nonlinear Frequency Discriminators", *IEEE Transactions on Wireless Communications*, ISSN: 1536-1276, Vol. 2, Issue 5, pp. 1101-1107, September 2003 (TR2003-041)

  Porikli, F.M., "Road Extraction by Point-wise Gaussian Models", *SPIE Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery IX*, Vol. 5093, pp. 758-764, September 2003 (TR2003-059)

* Sahinoglu, Z.; Orlik, P.V., "Regenerator versus Simple Relay with Optimum Transmit Power Control for Error Propagation", *IEEE Communication Letters*, ISSN: 1089-7798, Vol. 7, Issue 9, pp. 416-418, September 2003 (TR2003-054)

  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)*, September 2003 (TR2003-140)

  Huang, Q.; Cukier, J.I.; Kobayashi, H.; Liu, B.; Zhang, J., "Fast Authenticated Key Establishment Protocols for Self-Organizing Sensor Networks", *International Conference on Wireless Sensor Networks and Applications (WSNA)*, ISBN: 1-58113-746-8, pp. 141-150, September 2003 (TR2003-102)

  Vetro, A., "Transcoding Scalable Coding & Standardized Metadata", *International Workshop Very Low Bitrate Video (VLBV)*, IBSN: 3-540-2008109, Vol. 2849, pp.15-16, September 2003 (TR2003-113)

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)*, Vol. 1, pp. 5-8, September 2003 (TR2003-118)

Wada, H.; Ogata, M.; van Baar, J.; Raskar, R., "An Experimental Prototype of Multiple Projectors Display System for Visual Systems", *Forum on Information Technology*, September 2003 *(Text in Japanese Only)*

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)*, Vol. 1, pp. 29-32, September 2003 (TR2003-116)

Porikli, F.M., "Inter-Camera Color Calibration by Cross-Correlation Model Function", *IEEE International Conference on Image Processing (ICIP)*, Vol. 2, pp. 133-136, September 2003 (TR2003-103)

Guillamet, D.; Moghaddam, B.; Vitria, J., "Higher-Order Dependencies in Local Appearance Models", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, Vol. 1, pp. I 213-216, September 2003 (TR2003-093)

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)*, ISSN: 1522-4880, Vol. 1, pp. 161-164, September 2003 (TR2003-117)

Peker, K.A.; Divakaran, A., "An Extended Framework for Adaptive Playback-Based Video Summarization", *SPIE Internet Multimedia Management Systems IV*, Vol. 5242, pp. 26-33, September 2003 (TR2003-115)

Radhakrishan, R.; Xiong, Z.; Divakaran, A.; Raj, B., "Investigation on Effectiveness of Mid-Level Feature Representation for Semantic Boundary Detection in News Video", *SPIE Internet Multimedia Management Systems IV*, Vol. 5242, pp. 74-80, September 2003 (TR2003-119)

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)*, Vol. 3, pp. 2276-2280, September 2003 (TR2003-048)

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)*, Vol. 3, pp. 2009-2013, September 2003 (TR2003-122)

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)*, Vol. 3, pp. 2064-2067, September 2003 (TR2003-047)

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)*, Vol. 1, pp. 877-881, September 2003 (TR2003-069)

Molisch, A.F.; Zhang, X.; Kung, S-Y; Zhang, J., "DFT-Based Hybrid Antenna Selection Schemes for Spatially Correlated MIMO Channels", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Vol. 2, pp. 1119-1123, September 2003 (TR2003-070)

Huang, Q.; Kobayashi, H.; Liu, B.; Gu, D.; Zhang, J., "Energy/Security Scalable Mobile Cryptosystem", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Vol. 3, pp. 2755-2759, September 2003 (TR2003-079)

Sidner, C.L.; Lee, C.H.; Lesh, N.B., "Engagement by Looking: Behaviors for Robots When Collaborating with People", *Diabruck: Workshop on the Semantics and Pragmatics of Dialogue*, Kruiff-Korbayova and Kosny (eds), Univ of Saarland, pp. 123-130, September, 2003 (TR2003-130)

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

Lamere, P.; Kwok, P.; Walker, W.; Gouvea, E.; Singh, R.; Raj, B.; Wolf, P.P., "Design of the CMU Sphinx-4 Decoder", *Eurospeech*, September 2003 (TR2003-110)

Singh, R.; Warmuth, M.; Raj, B.; Lamere, P., "Classification with Free Energy at Raised Temperatures", *Eurospeech*, September 2003 (TR2003-022)

* Almers, P.; Tufvesson, F.; Molisch, A.F., "Measurement of Keyhole Effect in a Wireless Multiple-Input Multiple-Output (MIMO) Channel", *IEEE Communication Letters*, ISSN: 1089-7798, Vol. 7, Issue 8, pp. 373-375, August 2003 (TR2003-043)

Suzuki, N.; Iwata, M.; Sasakawa, K.; Komaya, K.; Nikovski, D.N.; Brand, M.E., "A Calculation Method of Car Arrival Time with Dynamic Programming in Elevator Group Control", *Conference of the Institute of Electrical Engineers of Japan*, August 2003 *(Text Appears in Japanese Only)*

Gezici, S.; Fishler, E.; Kobayashi, H.; Poor, H.V., Molisch, A.F., "A Rapid Acquisition Technique for Impulse Radio", *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, Vol. 2, pp. 627-630, August 2003 (TR2003-046)

Molisch, A.F.; Zhang, J.; Miyake, M., "Time Hopping and Frequency Hopping in Ultra-wideband Systems", *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, Vol. 2, pp. 541-544, August 2003 (TR2003-092)

Zhang, X. M.; Shi, Y-Q; Xu, W-Q; Vetro, A.; Sun, H., "Optimal 2-d Interleaving for Robust Multimedia Transmission", *International Conference on Information Technology: Research and Education (ITRE)*, pp. 74-78, August 2003 (TR2003-148)

Molisch, A.F., "MIMO Systems with Antenna Selection: An Overview", *IEEE Radio & Wireless Conference (RAWCON)*, pp. 167-170, August 2003 (TR2004-014)

* Brand, M.E., "Continuous Nonlinear Dimensionality Reduction by Kernel Eigenmaps", *International Join Conference on Artificial Intelligence (IJCAI)*, August 2003 (TR2003-021)

* Lesh, N.B.; Marks, J.W.; McMahon, A.; Mitzenmacher, M., "New Heuristic and Interactive Approaches to 2D Rectangular Strip Packing", *International Joint Conference on Artificial Intelligence (IJCAI)*, August 2003 (TR2003-018)

* Garland, A.; Lesh, N.B.; Rich, C., "Responding to and Recovering from Mistakes During Collaboration", *International Joint Conference on Artificial Intelligence (IJCAI)*, August 2003 (TR2003-017)

* Nikovski, D.N.; Brand, M.E., "Marginalizing Out Future Passengers in Group Elevator Control", *Conference on Uncertainty in Artificial Intelligence (UAI)*, ISBN: 0-127-05664-5, pp. 443-450, August 2003 (TR2003-062)

* Matusik, W.; Pfister, H.; Brand, M.E.; McMillan, L., "A Data-Driven Reflectance Model", *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 22, Issue 3, pp. 759-769, July 2003 (TR2003-083)

* Raskar, R.; van Baar, J.; Beardsley, P.A.; Willwacher, T.; Rao, S.; Forlines, C., "iLamps: Geometrically Aware and Self-Configuring Projectors", *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 22, Issue 3, pp. 809-818, July 2003 (TR2003-023)

  Sidner, C.L.; Lee, C.H.; Lesh, N., "The Role of Dialog in Human Robot Interaction", *International Workshop on Language Understanding and Agents for Real World Interaction*, July, 2003 (TR2003-063)

* Zhou, S.; Chellappa, R.; Moghaddam, B., "Adaptive Visual Tracking and Recognition using Particle Filters", *IEEE International Conference on Multimedia and Expo (ICME)*, Vol. 2, pp. 349-352, July 2003 (TR2003-095)

  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, July 2003 (TR2003-099)

  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)*, Vol. 2, pp. 41-44, July 2003 (TR2003-097)

  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)*, Vol. 2, pp. 361-364, July 2003 (TR2003-016)

  Porikli, F.M.; Divakaran, A., "Multi-Camera Calibration, Object Tracking and Query Generation", *IEEE International Conference on Multimedia and Expo (ICME)*, Vol. 1, pp. 653-656, July 2003 (TR2003-100)

  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)*, Vol. 2, pp. 417-420, July 2003 (TR2003-098)

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 Multimedia and Expo (ICME)*, Vol. 3, pp. 401-404, July 2003 (TR2003-149)

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)*, Vol. 3, pp. 29-32, July 2003 (TR2003-101)

Xiong, Z.; Radhakrishnan, R.; Divakaran, A.; Huang, T.S., "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 Multimedia and Expo (ICME)*, Vol. 3, pp. 397-400, July 2003 (TR2004-082)

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 Communications
- Digital Video
- Off the Desktop Interaction and Display
- Sensor and Data Systems

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

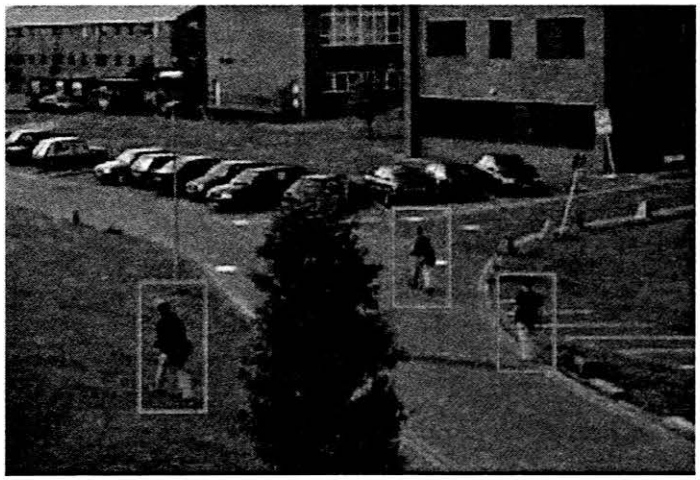
Computer Vision is the branch of computer science concerned with the analysis of images to extract information about the world. This is the same function that the human visual system provides (although perhaps accomplished through different mechanisms). As sensor and computer hardware drops in cost, 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 pedestrian detection, facial feature finding, face recognition, and gender and race classification. Last fall this work was honored with the Marr award at a primary computer vision conference. 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. Once having detected and tracked objects, we now strive to analyze the activity or event-taking place.

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, audio-visual event detection, intelligent video browsing systems and fusion of imaging sensors. 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

Pedestrian Detection



See Color Figure 1

MERL is developing new technology for automatically detecting pedestrians in video sequences. Detecting and tracking pedestrians is important for surveillance purposes. It can be used to sound an alarm if an intruder is in a restricted area or to aid in browsing hours of surveillance video by skipping to the next part of the video where a person was seen.

Background and Objectives: The problem of detecting people in low resolution surveillance video is a fundamental problem for surveillance. The problem is difficult because the pedestrian may be very small in the image, making the amount of information contained in the pixels small. Furthermore, there may be background motion in the scene (such as trees waving in the breeze or a cloud shadow passing by) which makes motion detectors faulty. To overcome these problems, we learn a model that encompasses both the appearance (pixels) and the motion of pedestrians. The pedestrian model is fast to evaluate which makes it feasible to search for pedestrians across the frames of a video.

Technical Discussion: To detect pedestrians in a video sequence, we learn a foreground model of the motion and appearance of pedestrians from example video sequences. The pedestrian detector builds on earlier face detection work at MERL. The earlier work is extended by using motion information as well as appearance information. This is in contrast to most prior work which attempts to build a model of the background. In our case, no background model is used. The detector learned uses a set of simple motion and appearance features. The appearance features are simple rectangle features acting on a single frame of the video. The motion features are simple rectangle features acting on the difference image between successive frames of the video. The optimal set of motion and appearance features are learned from a large library of possible features using the AdaBoost learning algorithm.

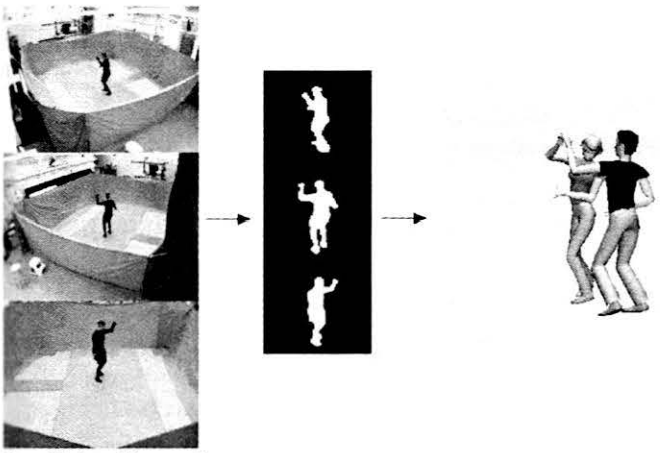
Collaboration: Discussions with the Physical Security System Project at Sentansoken, the Visual Communication Systems Department at Gunden and the Imaging Storage Products Department at Kyoden.

Future Direction: There are many ideas for further improvements such as using more than two frames for modeling motion and experimenting with different motion features. There are also ideas for improving the speed of the detector which will be important for productization.

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

Lab: MERL Technology Lab
Project Type: Research

Human Activity Determination



See Color Figure 2

The goal of this project is to classify complex motions (activities) of human bodies. Instead of relying on image data alone, we are using a 3D motion database that contains representative motions. The motions are captured in a MoCap studio and depend on the application of the system. The users then perform similar motions in front of inexpensive video cameras. The system will automatically classify the user's motions into categories and execute appropriate commands (e.g., raise an alarm). Applications include physical security, assistive technologies for handicapped people, remote supervision of physiotherapy, or remote motion training for athletes.

Background and Objectives: The insight behind the system is that while simple vision processing may provide incomplete and inaccurate information about the user's movements, with the addition of domain knowledge from a previously captured motion database, plausible classification is possible. Perhaps more surprisingly, this process can be performed interactively, with less than a second of delay between the capture of the video and the classification of the motion.

Technical Discussion: The system combines information about the user's motion contained in silhouettes from several viewpoints with domain knowledge contained in a motion capture database. In our system, the user performs in front of one to three video cameras and the resulting silhouettes are used to estimate his or her orientation and body configuration based on a set of discriminative local features. Those features are selected by a machine learning algorithm during a pre-processing step. Sequences of motions that approximate the user's actions are extracted from the motion database and scaled in time to match the speed of the user's motion.

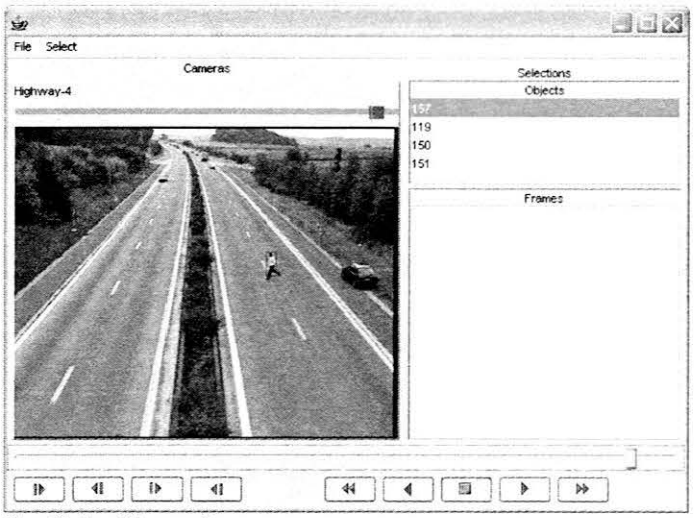
Collaboration: Jessica Hodgins and Liu Ren, CMU.

Future Direction: Define assistive technologies for handicapped people, implement and improve current system.

Contact: Hanspeter Pfister
http://www.merl.com/projects/actdet/

Lab: MERL Research Lab
Project Type: Research

Unusual Event Detection



See Color Figure 3

The ultimate goal of most surveillance systems is automatic detection of events and suspicious activities thereby triggering alarms (detection) as well as reducing the volume of data presented to human operator (retrieval). Event detection requires interpretation of the "semantically meaningful object actions." To achieve this task, the gap between the numerical features of video objects and the symbolic description of their meaningful activities needs to be bridged. Highway monitoring, airport surveillance, building access control are just a few of the several important applications.

Background and Objectives: Past work on event detection has mostly consisted of extraction of object trajectories followed by a supervised learning using parameterized models for actions. Models are usually predefined dynamic patterns of movements. Such methods assume that 'all' unusual event events can be modeled, which requires off-line training. However, it is not viable to foresee every possible event. Besides, the nature of event varies depending on the application, thus event modeling becomes even more challenging.

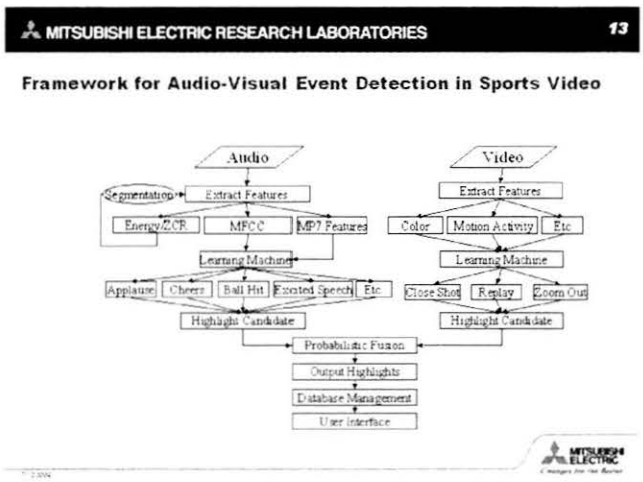
Unlike the past work cited above, we employ an unsupervised learning method. Our method does not require definition of what is usual and what is not. We define usual as the high recurrence of events that are similar. As a result, unusual is the group of events that are not similar to the rest. This enables as to detect multiple unusual events. The developed technique enables fusing different type of features, e.g. histograms, HMM's scalars, etc. as well.

Technical Discussion: We develop an event detection framework that has two significant advantages over past work. First, we introduce an extended set of time-wise and object-wise statistical features including not only the trajectory coordinates but also the histograms and HMM based representations of object's speed, orientation, location, size, and aspect ratio. These features enable detection of events that cannot be detected with the existing trajectory features reported so far. Second, we introduce a spectral clustering algorithm that can automatically estimate the optimal number of clusters. First, we construct feature-wise affinity matrices from the pair-wise similarity scores of objects using the extended set of features. To determine the usual events, we apply eigenvector decomposition and obtain object clusters. We show that the number of eigenvectors used in the decomposition is proportional to the optimal number of clusters.

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

Lab: MERL Technology Lab
Project Type: Research

Audio-Visual Event Detection for Consumer and Surveillance Video



Audio-visual analysis enables us to detect diverse events ranging from sports highlights to criminal activities in elevators. The focus of this project is to combine the audio-visual cues to get higher accuracy as well as deeper understanding of audio-visual events. Our initial results show that our audio-visual fusion-based event detection improves upon on techniques that rely on audio or video alone.

Background and Objectives: In the past we used exclusively video or exclusively audio to detect events for both consumer and surveillance applications. Such event detection enables us to detect live events as well as to summarize the content, which in turn enables rapid browsing of stored video, which is applicable to both personal video recorders and surveillance recorders. We have used motion activity to detect patterns corresponding to sports highlights as well as changes in highway traffic density. We have also used audio to detect sports highlights by finding long stretches of audience reaction in the form of cheering and applause. While both modes provide powerful cues, we find that we hit a ceiling with the detection performance. That motivates us to examine various combinations of audio and video cues for detecting a wide range of events.

Technical Discussion: The challenge lies in finding the right audio-visual fusion after systematic domain knowledge acquisition. We have tackled the systematic domain knowledge acquisition by developing multimedia-mining techniques that help us discover patterns in the content. Such unsupervised analysis of the content reveals the intrinsic nature of the content and enables systematic choice of key event classes and training of the corresponding detectors. We can then further investigate audio-visual fusion techniques that best exploit the underlying characteristics of the content. In other words, our aim is to develop audio-visual fusion techniques that adapt to the data. Such adaptive techniques are essential to cope with the wide range of variety of consumer and surveillance video that we have to tackle. In addition to the variety across the different genres of video, there is a significant diversity within genres as well. We aim to develop a common core of audio-visual fusion techniques that work across the aforementioned wide variety of events, in combination with application specific techniques.

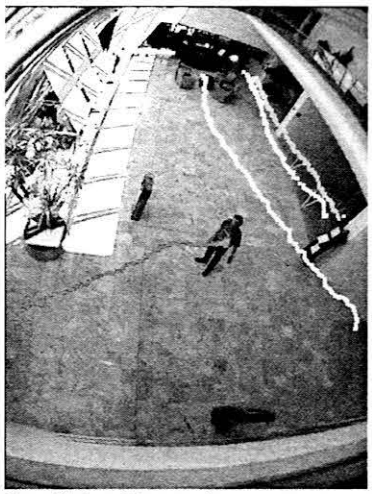
Collaboration: Joho-soken, Sentan-Soken

Future Direction: We plan to extend our techniques on both the consumer and surveillance video fronts to a broader variety of content.

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

Lab: MERL Technology Lab
Project Type: Research

Video Object Tracking



See Color Figure 4

As a part of the Physical Security SK, we developed robust, computationally efficient, unsupervised multi-object tracking techniques that requires minimum initialization effort and fine-tuning for stationary camera setups. Object tracking is a key technology of most surveillance products developed at Mitsubishi Electric.

Background and Objectives: Accurate object segmentation and tracking under the constraint of low computational complexity presents a challenge. Generally speaking, tracking of objects can be done either by back-tracking or by forward-tracking. The back-tracking based approach segments foreground regions in the current image and then establishes the correspondence of regions between the previous image. The forward-tracking approach estimates the positions of the regions in the current frame using the segmentation result obtained for the previous image. For establishing correspondence, several object templates are utilized. The limitation of this approach is a single template is not sufficient for wide variety of applications, e.g. human tracking and traffic monitoring require different templates. A possible forward-tracking technique is mean-shift analysis. Mean-shift is a nonparametric density gradient estimator. It is employed to derive the object candidate that is the most similar to a given model while predicting the next object location. This method provides accurate localization, and it is computationally fast. However, the mean-shift tracker is not automatic since it requires initial model properties, i.e. object boundary, etc.

Technical Discussion: We developed an automatic and real-time object-tracking algorithm by integrating a mixture model-based background subtraction into a mean-shift based forward tracking mechanism. We combined these methods to accomplish an automatic and robust tracker that can handle high resolution color video in real-time. We also address other main difficulties, i.e. managing sudden illumination changes in the scene, overcasts, shadows, and correspondence problems. We find human face and arms by applying a skin color mask which is formulated in the RGB space by offline training. Our method generates a reference image using pixel-wise mixture models, finds changed part of image by background subtraction, removes shadows by analyzing color and spatial properties of pixels, determines objects, and tracks them in the consecutive frames. Currently, we are improving our state-of-art tracker by integrating particle filter based posterior probability estimation methods. In addition to color histograms, we now use edge, motion, and appearance models to guide the mean-shift procedure. We use a novel inside/outside scale adaptation.

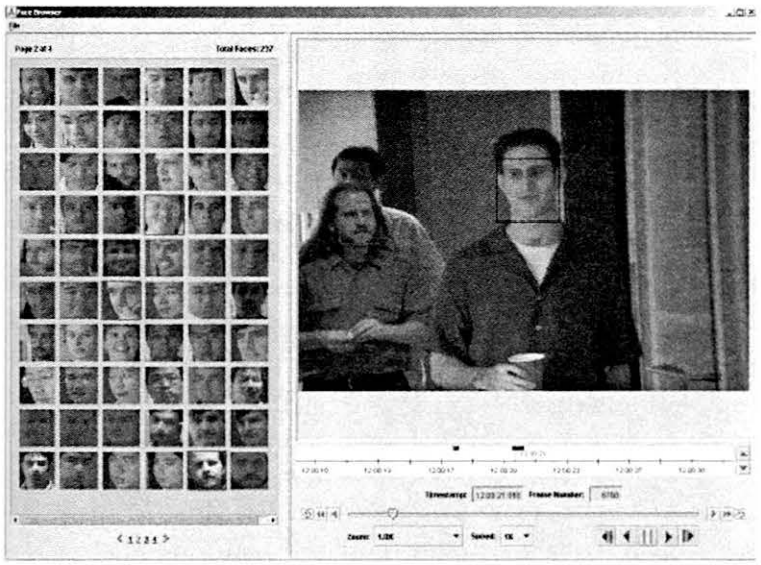
Collaboration: SentanSoken, Gunden, MRL.

Future Direction: This summer we are extending our solutions for PTZ camera tracking.

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

Lab: MERL Technology Lab
Project Type: Research

Face Based Browsing for Surveillance Applications



See Color Figure 5

Surveillance applications often collect a large amount of video data. Currently the surveillance applications do not allow the user to quickly search the collected data for an occurrence of a particular individual. This project gives the user this capability by displaying a pallet filled with the faces of all the individuals observed by a particular camera during a particular time. The user can then navigate to the video segments that contain a particular person. It is further possible to arrange the people in order of similarity with respect to a particular individual. This is an application of the Video Warehouse project.

Background and Objectives: Surveillance applications are meant to help users understand "the Four W's" (Who, What, Where and When) of a situation. Current surveillance applications only allow search on the last two of these, the where and when. This makes browsing surveillance data a very tedious and time consuming process. By adding the "Who" component to the searching capability of a surveillance application, we hope to relive the tedium of the process and make it faster.

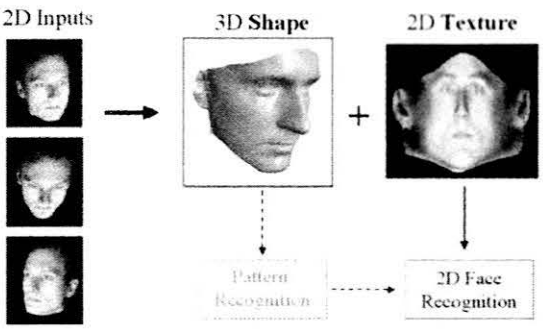
Technical Discussion: The browsing application leverages face detection, object tracking and face recognition technologies in order to provide the user with a most concise list of faces possible. Efficient database techniques have been utilized to make searching through image metadata as fast as possible.

Future Direction: Expanding search types to enable more sophisticated search capabilities. Improving the display to allow concurrent browsing through faces gathered by multiple cameras, and further simplification of the UI to reduce the cognitive load on the user.

Contact: Sergei Makar-Limanov, Jay Thornton
http://www.merl.com/projects/face-browsing/

Lab: MERL Technology Lab
Project Type: Advanced Development

3D Face Recognition



By applying 3D face models to robust automatic face recognition we are addressing the most critical factors limiting performance: illumination and pose variation. We have developed a novel system for capturing the 3D shape of a human face from a sequence of sparse 2D silhouettes from multiple cameras (or video) at affordable cost and with no manual user interaction. Using silhouettes decouples the geometric subtleties of the human face from the nuances of shading and texture. Our framework presents several computational and algorithmic advantages over the existing techniques for 3D face modeling. We are now applying our modeling framework to illumination- and pose-invariant 2D/3D face recognition with promising results.

Background and Objectives: The "state-of-the-art" is the "3D Morphable Models" of Blanz & Vetter [SIGGRAPH'99] which uses an analysis-by-synthesis framework for capturing 3D models from 2D photograph(s). Models are fit by finding the shape/texture parameters which will render a synthetic 2D image which best matches the observed image. The key advantage of our silhouette-based approach is that it does not rely on dense image/texture correspondence in order to estimate our model's shape parameters. Instead, the face shape is estimated directly by way of its own intrinsic cues: the occluding contours (as represented by the object's silhouettes). The texture information, on the other hand, is simply lifted and post-processed after the shape estimation stage is completed.

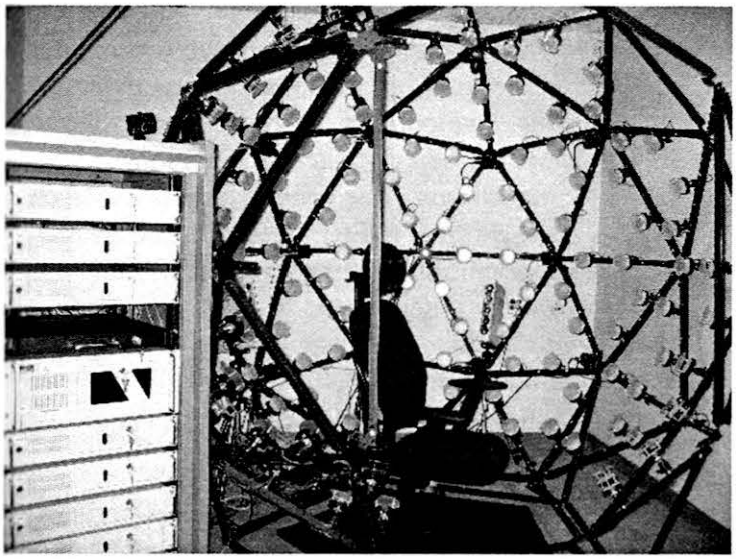
Technical Discussion: We use a linear combination of "eigenheads" obtained by Principal Component Analysis (PCA) of a training set of laser-scanned 3D human faces. The PCA coefficients are used as model parameters. We establish correspondence between faces with an efficient error metric (boundary weighted XOR). Our parameter estimation uses a "downhill simplex method" (which requires no gradients) and can be readily adapted to existing graphics hardware for computational speedup. Moreover, the resulting parameter recovery is surprisingly robust with respect to partial and noisy input silhouettes (with both positive and negative clutter). Our overall model acquisition pipeline is considerably faster (x10) than existing state-of-the-art techniques which rely on dense correspondence (eg. optical flow in "Morphable Models") and is robust with respect to illumination and texture variation across the face and we have now achieved near automatic model-fitting using MERL's state-of-the-art face/feature detection technology.

Future Direction: Having established and demonstrated superior performance in silhouette-based 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: Baback Moghaddam, Hanspeter Pfister
http://www.merl.com/projects/3Dfacerec/

Lab: MERL Research Lab
Project Type: Research

Face Scanning



We have built a scanning system that is able to acquire images of human faces from different viewpoints under varying illumination. The system also captured the 3D geometry of the face using a commercial face scanning system. The scanner hardware consists of a geodesic dome with 16 high-resolution color cameras and 150 computer-controlled white LED lights. We plan to acquire face data from a large cross-section of the population.

Background and Objectives: Our research aims at understanding how the human face changes over a person's lifetime and as a function of expression (e.g. how does the face change when somebody frowns, smiles, scowls, etc.). We are interested in measuring both the geometric deformations of a human's face as they age, speak and change expression and the variation in the surface reflectance properties of their skin. There are several applications of understanding this process including: naturally aging images of missing children/fugitives from a single photograph, face recognition, digitally aging actors for entertainment purposes and re-animating faces of actors to match alternative sound tracks (e.g. foreign language dubs).

Technical Discussion: In this study we will measure several aspects of a subject's face including 3D geometry, small-scale mesostructure (e.g. wrinkles, bumps, etc.) and surface reflectance in a variety of expressions and for the subject pronouncing different words. Each subject will also be asked to complete a questionnaire about their lifestyle and age. Additionally, subjects will be asked to provide photographs of themselves over the course of their life that will be digitized with a scanner.

Collaboration: Tim Weyrich and Markus Gross, ETH Zurich; Jovan Popovich and Daniel Vlasic, MIT; Jason Lawrence, Princeton University; Henrik Wann Jensen, UC San Diego

Future Direction: Scanning of a large cross-section of the population; development of appropriate models for aging, reflectance, and face geometry; applications of the model to specific domains.

Contact: Hanspeter Pfister
http://www.merl.com/projects/facescanning/

Lab: MERL Research Lab
Project Type: Research

DiamondClassify

Output of Face Detector on Test Images

The DiamondClassify project has created a group of libraries intended to support the development of products using the MERL object detection and recognition algorithms. The code is largely based on the original MERL detection and recognition libraries, but substantial improvements have been made to code readability and organization to bring the code to product-level status.

Background and Objectives: The MERL object detection and face recognition library, developed as part of the CHO SK, has grown into a large and unwieldy library, containing both detection and recognition code intended for use in products, and significant amounts of obsolete and experimental research and training code. The primary goal of the DiamondClassify project has been to break this large monolithic library into several smaller, more manageable units by separating the code into libraries intended for use in products, and libraries solely used for research and training purposes. A secondary goal has been to make the code simpler and easier to understand by improving its readability and organization.

Technical Discussion: The monolithic library was first broken down into several smaller libraries consisting of utility, object detection, face recognition, and training code. These libraries were designed to form a hierarchy, such that each one depends solely on the libraries below it in the hierarchy. This means that a product that needs to perform object detection only needs the utility and detection libraries, not the recognition or training ones. In dividing the monolithic library into the hierarchical libraries, it was sometimes sufficient to move an entire file into the upper levels of the hierarchy, while in other cases it was necessary to separate a large C++ class into multiple classes, one containing object detection functionality, and another containing training functionality, for example. In these cases, inheritance was used to maintain a relationship between the new classes.

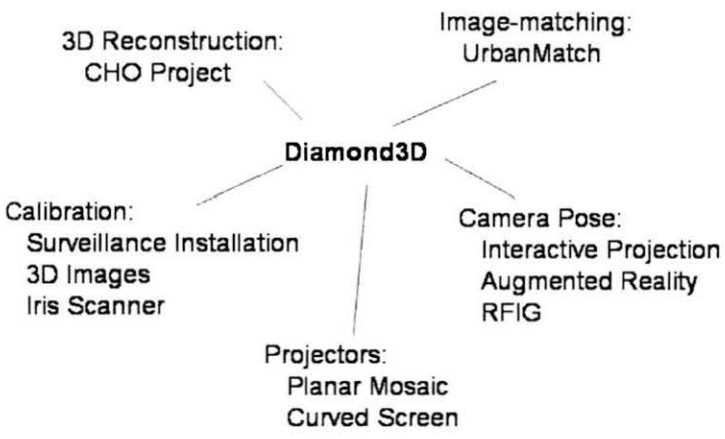
Collaboration: We are working with the Sensing Information Processing systems Department at Sentansoken. We are also collaborating with the Display systems Technology Department at Johosoken.

Future Direction: As new detection and recognition algorithms become available through ongoing research at MERL, we expect the DiamondClassify library to expand to encompass this new work. This will allow us to make advances in research available for use in products in a timely manner.

Contact: David Rudolph, Michael Jones
http://www.merl.com/projects/diamondclassify/

Lab: MERL Technology Lab
Project Type: Advanced Development

Diamond3D Computer Vision Library



Diamond3D is a software library for creating computer vision applications. It has three characteristics (a) it is easy to create new applications, (b) it is easy to reuse high-level vision components between applications, (c) it provides extensive functionality to do with feature and region detection, calibration, recovery of camera pose, image matching, and 3D reconstruction. The library has been used for multiple projects including stereo for surveillance, multi-projector systems, RFIG, and augmented reality.

Background and Objectives: Many computer vision applications utilize low-level components like feature and region detection. In addition, 3D-related applications utilize calibration, recovery of camera pose, image matching, and 3D reconstruction. This library offers reusable components that are easy to include in a new or existing application. The primary value is reduced time to create vision applications, with less time spent on maintenance and debugging because the components are well-tested from previous applications. An additional value is that the library offers easy creation of GUIs and visualization of results, providing a range of standard functionality like reading from a camera, read/write from disk, examining images by zoom or mouse pointing etc.

Technical Discussion: The library is in C++ and compiles on Windows or Linux. Each vision component is designed as a complete Model-View-Control (MVC) component. All vision components conform to a single standard interface, to aid the developer. Each component provides, in addition to its core computer vision algorithm, (a) extensive checking to prevent and report inappropriate use, (b) ready-to-use view functionality for displaying the results of the algorithm. The goal is allow easy reuse of components, automatic reporting of incorrect use, and easy visualization of the algorithm's behavior. The library currently supports tens of applications using about 50 different vision components, and building on several hundred lower-level C++ classes.

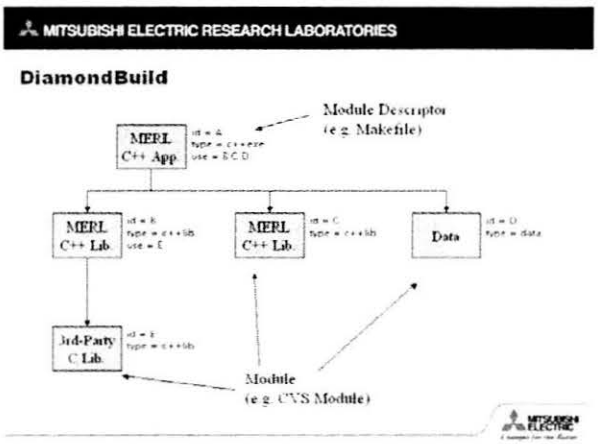
Collaboration: Dr. Miyata, SentanSoken.

Future Direction: The library is expanding to include new functions.

Contact: Paul Beardsley
http://www.merl.com/projects/d3d/

Lab: MERL Research Lab
Project Type: Advanced Development

DiamondBuild



DiamondBuild is a portable build system designed to enable professional and reliable software management in a highly heterogeneous and diverse software development environment such as the one present at MERL. DiamondBuild provides a lightweight framework that enables effective software sharing, integration, and delivery to our customers.

Although the system has been driven by MERL's own complex software management problems, these problems are common in software development, especially in the open-source community. The DiamondBuild project has synergy with existing open-source efforts to develop a portable and open build system. Thex ultimate objective of this project is to provide a mature prototype and specification to bolster this open-source effort, which MERL can then benefit from.

Background and Objectives: MERL's software development environment is characterized by usage of multiple operating systems, programming languages, compilers, and third-party open-source software packages. Furthermore many different versions of these tools are in use concurrently and individual projects follow diverse software development practices, making it difficult to share and integrate software internally and even more difficult to reliably transfer complex projects to our customers.

The objective of the DiamondBuild project is to establish a portable modular framework and build system that results in good and consistent software development and management practices througout MERL and results in reliable and effective transfer to our customers in Japan.

Technical Discussion: DiamondBuild has already evolved through a number of manifestations, including an initial version built entirely in "Make" and several versions augmented with "Perl". DiamondBuild has enabled software porting, sharing, management, and delivery of a number of software projects. Various usage problems associated with performance (re-build speed), reliability (due to OS-related differences), and absolute correctness (various unusual usage cases) remain to be solved before true widespread use is promoted.

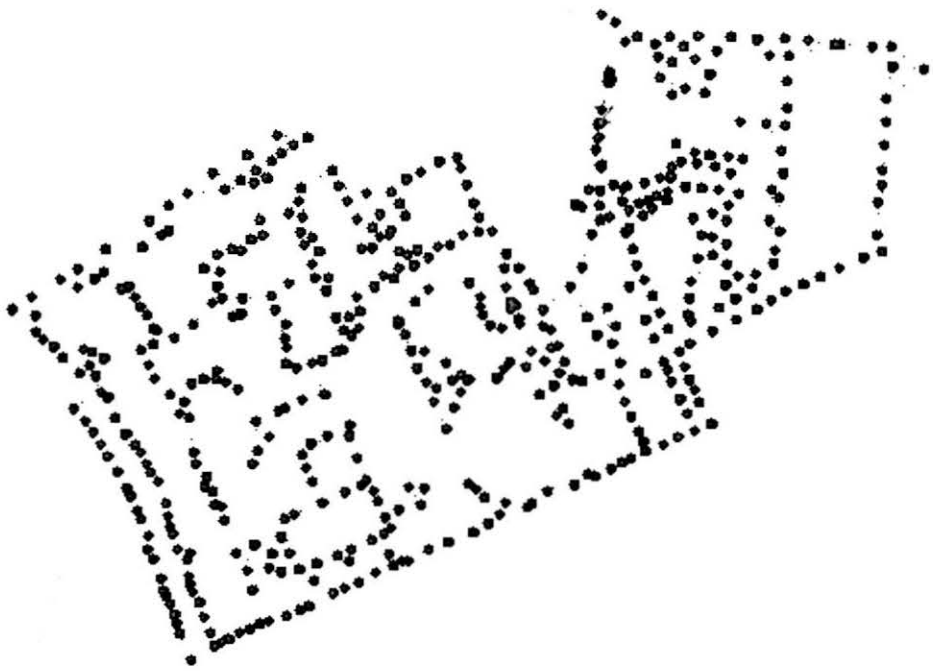
Collaboration: DiamondBuild has become the software management system for several major internal software projects, including Diamond3D and DiamondClassify. Successful deliveries of software to MELCO have been made to both Johosoken and Sentansoken using the DiamondBuild system.

Future Direction: A current effort is underway to re-implement the bulk of the system in "Perl", with the same "Make" front-end. This effort seeks to solve the major outstanding problems and make the tool more effective for internal and external users. This prototype will serve as the basis for transferring the project to the open-source community.

Contact: Ali Azarbayejani
http://www.merl.com/projects/diamondbuild/

Lab: MERL Technology Lab
Project Type: Advanced Development

Camera Network Calibration



Cameras are distributed throughout a city. From their images, we determine the 3D locations of the cameras and many of the environmental features they view. This makes it possible to quickly survey and build 3D models of cities.

Background and Objectives: Computer vision has intensely studied the problem of building 3D models of rooms and plazas from a dense set of images. We seek to build a 3D model of an entire city from a sparse set of images that might be collected by walking around with a camera.

Technical Discussion: By examining the vanishing points in images, it is possible to determine camera orientations. By examining environmental features (such as building corners) viewed by two cameras, it is possible to determine the direction from one camera to the other. The missing information is the relative locations of the cameras and environmental features in 3D. We developed a closed-form least-squares solution, and computed centimeter-accurate localizations from kilometer-scale camera networks.

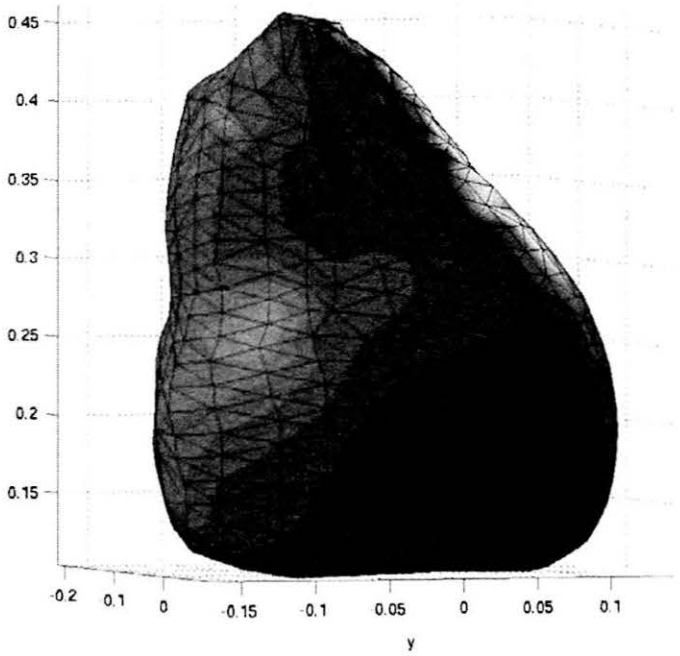
Collaboration: Seth Teller (MIT) and Matthew Antone (BBN).

Future Direction: This is part of a larger city-mapping project that DARPA would like to commercialize.

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

Lab: MERL Research Lab
Project Type: Initial Investigation

Surface Reconstruction



See Color Figure 6

We recover the 3D shape of an object from its silhouettes. The main innovation is that no correspondences are computed, significantly reducing computational complexity. The image at right shows the reconstructed surface of a pear.

Background and Objectives: 3D-from-silhouettes has long been studied in computer vision as a cheap way to 3D-scan objects. Current methods have problematic time complexity and generally produce very jagged 3D surfaces.

Technical Discussion: Technically, this is known as the convex hull problem. Previous methods requires expensive point matching to determine where silhouettes cross in 3-space. We project the problem into a dual space and show that by exploiting simple continuity principles in that space, a smooth manifold approximating the convex hull can be estimated without matching.

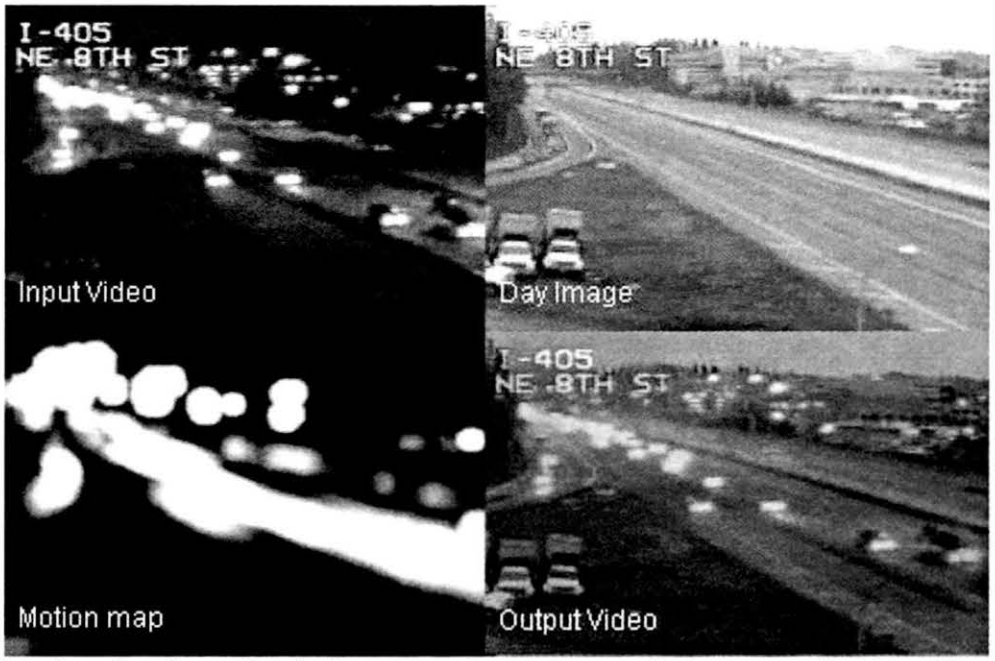
Collaboration: Kongbin Kang (Brown University)

Future Direction: The current method requires camera calibration. We would like to relax this requirement.

Contact: Matthew Brand
http://www.merl.com/projects/surface-reconstruction/

Lab: MERL Research Lab
Project Type: Research

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.

Collaboration: Adrian Ilie and Jingyi Yu at University of North Carolina at Chapel Hill.

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

Lab: MERL Research Lab
Project Type: Research

Digital Communications

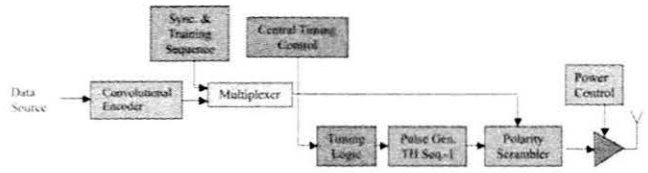
Digital communications and networking are pervasive in today's society. With advanced technologies at physical layer, medium access control layer and network layer, it will provide high speed communication capability for transmissions of voice, data as well multimedia information with quality of service over the air or wireline, and connect people at anywhere and anytime. From advanced wireless multimedia systems to simple integrated home networking, communications and networking technologies is at the center of a continuing revolution.

At MERL, our goal is to seek new business and technology trends in the area of digital communications and networking. We are not only conducting fundamental researches on communication theories, but also developing new core technologies for emerging applications. Our current focus is on broadband wireless communications, pervasive ad hoc networks, and digital home networking.

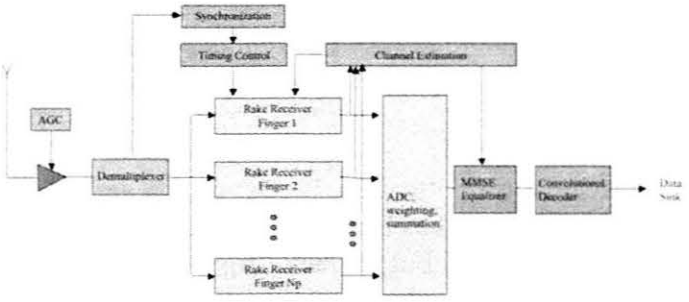
Over the past year, ultra-wideband (UWB) has attracted lots of attention. It can provide very high data rate for short-range wireless communications. MERL has become an important player in the field. We developed not only time-hopping impulse radio, but also investigated technologies to enhance the multi-band OFDM proposal, such as the symbol spreading technique. ZigBee was a very promising networking technology developed for many applications including sensor networks. As a key player, MERL contributed to the ZigBee Alliance in many ways including link quality based network protocol, security toolbox, and application profiles. For broadband mobile communications, MERL continued to develop MIMO (Multi-Input-Multi-Output) technologies, such as sub-group rate control and joint RF-baseband processing, and contribute to 3GPP standard activity. In the area of wireless LAN, MERL continued IEEE802.11 development, focused on 802.11e for QoS, high efficient MAC and high throughput PHY for next generation WLAN. MERL has completed SCP development for power line communication and started working on integrated home networks.

Project Descriptions

Impulse Radio



a)



b)

Impulse radio systems communicate by encoding each symbol as a sequence of short pulses. Due to the short pulse duration, the spectrum of the signal is very wide, and thus fulfills the requirements of the FCC for ultrawideband systems. Due to the absence of frequency upconversion, low-cost transceiver structures are possible. In our project, we investigated advanced impulse radio systems that have good performance, and are especially suitable for transmission with high data rates.

Background and Objectives: In 2002, the American frequency regulator (FCC) has allowed unlicensed ultrawideband (UWB) communications. Since that time, there is growing interest in UWB, both for high-data-rate communications (wireless UWB, wireless IEEE 1394), as well as for sensor networks with geolocation capabilities. Impulse radio systems show good potential for both applications.

Technical Discussion: While impulse-radio systems show good potential, they have several problems: the collection of energy from many multipath components, increased acquisition times, and difficulties in shaping the spectrum to fill the FCC mask as closely as possible. Our research has resulted in patented solutions for all of those problems, resulting in high-performance UWB transceiver designs.

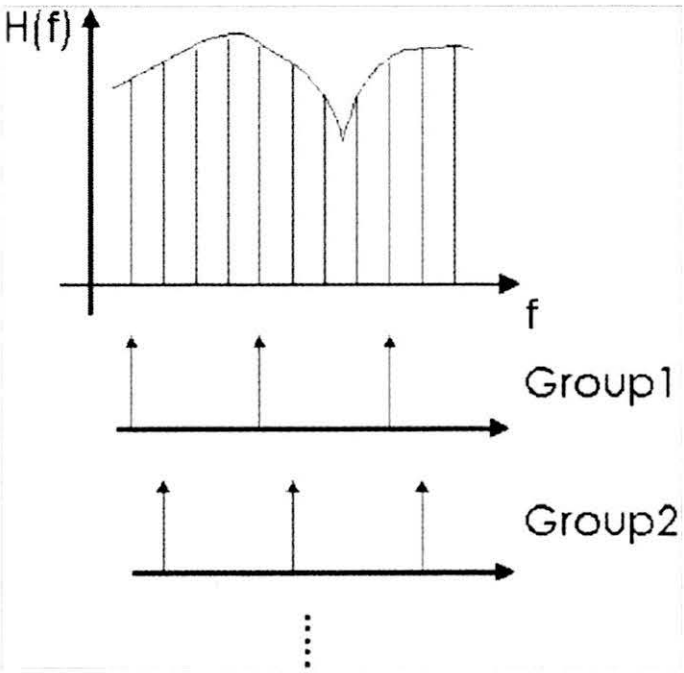
Collaboration: Joho-soken

Future Direction: The investigation of transmitted-reference schemes, and adaptation of our high-data-rate schemes to lower rates will be at the center of our future research.

Contact: Jinyun Zhang, Andreas F. Molisch
http://www.merl.com/projects/impulseradio/

Lab: MERL Technology Lab
Project Type: Research

Symbol Spreading Multiband OFDM



OFDM is a modulation format that divides the available spectrum into a number of narrowband channels (subcarriers); modulation is performed in each of those subcarriers. This modulation method shows considerable advantages for systems with high data rates, as it reduces errors due to intersymbol interference, including ultrawideband (UWB) communications. We are working on an enhancement that spreads the information over multiple carriers. This scheme (related to so-called multicarrier-CDMA) improves the robustness of the transmission with respect to interference and fading.

Background and Objectives: OFDM, in conjunction with time-frequency-interleaving, is used for high-data-rate (110 - 480 Mbit/s) wireless communications according to the MBOA standard. Such high data rates have applications for short-range wireless LANS, linking of home entertainment components (wireless IEEE 1394), and wireless UWB. Ultrawideband systems have attracted considerable attention in recent years, as the FCC has approved unlicensed UWB operation, subject to fulfillment of a mandated spectral mask.

Technical Discussion: While OFDM avoids problems with intersymbol interference, it does not exploit the frequency diversity inherent in a wideband channel. Some of this diversity is recovered when coding across tones is used; however, for weak codes, this is not sufficient. On the downside, spreading can lead to noise enhancement; this is avoided by maximum-likelihood detection of the received signal.

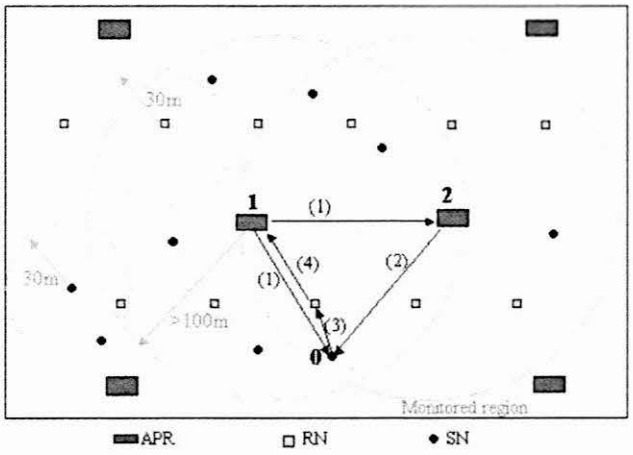
Collaboration: Joho-soken.

Future Direction: Efficient detection schemes will be investigated.

Contact: Jinyun Zhang, Andreas F. Molisch
http://www.merl.com/projects/symbolspreading/

Lab: MERL Technology Lab
Project Type: Research

Location Tracking for Ad-hoc Networks



The hybrid location tracking system we are developing consists of three device types: SNs, RNs and APRs. The SNs are simple sensor devices with a radio circuitry of very short communication range (e.g., 10-30 meters). They are only capable of TOA tracking of received signals, when required. The RNs are low-cost devices with simple routing capabilities and ability to measure the received signal strength (RSS). They also have a very short communication range (e.g., 10-30meters). The APRs are stationary known-location devices. The communication range of the APR is very long (e.g., > 100 meters). They are responsible for collecting messages from RNs and SNs, and forwarding them to the Central Monitoring Unit (CMU) through an aggregation point. The APRs also distribute commands received from the central monitoring unit to the corresponding RNs and SNs. The accuracy of the location estimation is determined jointly by TOA and RSS measurements. RSS measurements provide enhancement to only TOA based estimation at short distances.

Background and Objectives: The difference with respect to existing location estimation techniques exists in taking into account the heterogeneities of devices in a network such as communication ranges, time synchronization and routing capabilities etc. Locating techniques may be used for estimating sensor locations, tracking moving objects within a region and locating other physical entities such as sources of sounds. Applications requiring a certain degree of knowledge of sensor locations include environmental and structural monitoring (e.g., water quality monitoring and in-door air quality monitoring), indoor user tracking (e.g., locating of patients and medical personnel in hospitals and employee tracking in companies) and others. The scheme being developed supports both indoor and outdoor applications at low mobility.

Technical Discussion: TOA/RSS has better overall location accuracy than TOA. This is proven with expressions of Cramer-Rao Bounds of the estimation error. In particular, the use of RSS measurements mitigates the difficulties in locating devices within the close proximity of the reference devices, which are inherent to the time of arrival-based schemes. We show that the mutual placement of APRs and RNs is an additional tool that can be used to provide the desired location estimation accuracy in the network, even if the number of RNs and APRs is fixed.

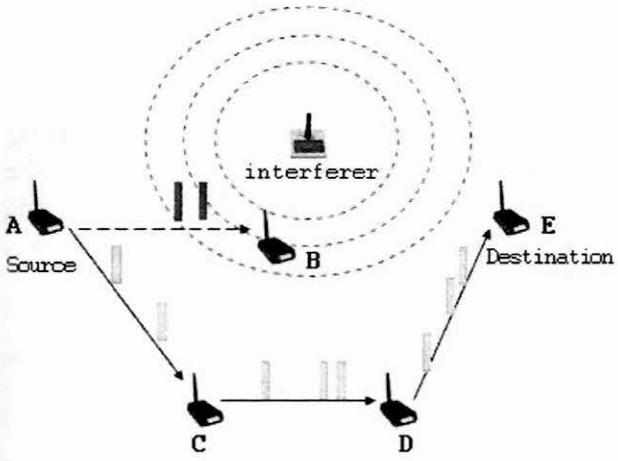
Collaboration: Johosoken

Future Direction: The prototype development of a wideband system solution will start at the end of 2004. An impulse radio UWB based ranging proposal is being prepared for IEEE 802.15.4a standards.

Contact: Zafer Sahinoglu, Jinyun Zhang
http://www.merl.com/projects/locationtracking/

Lab: MERL Technology Lab
Project Type: Research

Link Quality Indicator Based ZigBee Network Routing Protocol



Selected Route: A-C-D-E

The ZigBee Alliance is defining an international standard to enable reliable, cost-effective, low-power, wirelessly networked monitoring and control products. A ZigBee network is built on the principles of ad-hoc network and includes features such as automatic unsupervised network formation and multi-hop mesh networking. MERL is actively participating in this development. Specifically, we have been investigating methods to improve the route selection in mesh networks.

Background and Objectives: A multihop 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. However, the performance of any routing algorithm will be impacted by the selection of routes over which data is eventually transmitted. Simple routing algorithms that choose the minimum number of hops have been shown to lead to poor network performance in wireless environments. Recently our objective has been on developing enhancements that improve the overall reliability of the network by investigating methods to improve route selection. The approach taken has been to use existing features of the IEEE 802.15.4 physical and MAC layer to aide in route selection.

Technical Discussion: In a wireless multi-hop network it is preferable to transmit data over links that have a high probability success. In this way the number of the total number of retransmission required to successfully delivery data to the destination can be reduced. The ZigBee network layer is built on top of the IEEE 802.15.4 standard which enables the measurement of link quality between neighboring nodes in the network. This measurement is in the form of a Link Quality Indicator (LQI) value that is reported with each received packet. By averaging over several LQI values an estimate of the link quality can be obtained and therefore an estimate of the probability of successful transmission is available to the route selection algorithm.

Collaboration: We have worked closely with MELCO representatives to the ZigBee Alliance and are collaboration with Johosoken, Sentansoken/Shaden 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, Zafer Sahinoglu
http://www.merl.com/projects/zigbeenwk/

Lab: MERL Technology Lab
Project Type: Research

ZigBee Application Profile Contributions

ZigBee™ Alliance

We have been active in the ZigBee Applications Framework Group to develop an infrastructure and specific profiles that will properly support Mitsubishi Electric's products that would benefit from wireless personal-area mesh networking. We have been assisting MELCO in producing specific application profiles as the profiles are being created. So far, we have been working on the Industrial Plant Monitoring profile and the Split HVAC profile.

Background and Objectives: MERL has been active in the network and security layers of ZigBee since ZigBee's inception. This project is addressing MELCO's needs to produce profiles to support new devices.

Technical Discussion: Contributions to the ZigBee Applications Framework Group have taken many different forms. They include: submitting comments on proposed documents, assisting MELCO to produce new profiles for Industrial Plant Monitoring and Split HVAC, presenting MELCO profiles to ZigBee, providing feedback from ZigBee to MELCO, assuring that the lighting profile is well-designed and allows interoperability, lobbying for different policies and procedures for creating new profiles, participating in weekly telephone conference calls, and attending ZigBee meetings. We have been instrumental in modifying the procedure required for proposing new application profiles. In particular, rather than producing an MRD as part of a proposal, a PAR, or Profile Approval Request, to show market requirements will suffice. Also, rather than a design phase where the committee would request and review multiple proposals in the area of a new profile, it was agreed that issuing a Call for Participation in the new application area would be sufficient. In addition, rather than requiring three conforming implementations, it was agreed that just having the compliance group involved with a test suite would be sufficient to show conformance.

In working with the Split HVAC profile proposal, we have resolved issues regarding how to transmit large bulk data, how to deal with sharing data among nodes in the network, and several other questions.

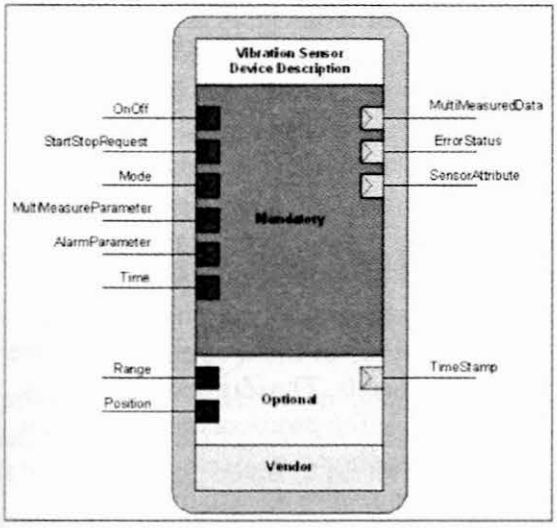
Collaboration: We are working with Kushiro-san, Nakata-san, and Yamamoto-san in the Living Environment Systems Laboratory. We have been in touch with Inoue-san in the Gunma works. We are collaborating with several members of MERL MTL who are active in the network and security layers of ZigBee.

Future Direction: We will continue to attend ZigBee Application Framework Group meetings and will continue to help develop new profiles.

Contact: Jamie Frankel
http://www.merl.com/projects/zigbeeprofile/

Lab: MERL Research Lab
Project Type: Research

ZigBee for Industrial Plant Monitoring



ZigBee is an emerging standard for low rate, low power, short range communications networks. The underlying MAC/PHY technology is IEEE 802.15.4. ZigBee is expected to drive diverse set of sensor network applications from home automation to environmental monitoring, and from asset management to automated meter reading etc. In this project, we target to develop an industrial plant monitoring (IPM) application profile and prototype system based on ZigBee technology. The application profile will consist of temperature sensor and vibration sensor device descriptions. These descriptions will also be contributed to the ZigBee Application Profile Group for standardization. The first draft was already submitted to and discussed in ZigBee in February 2004. The prototype end devices will have M16C microcontrollers on board.

Background and Objectives: The standardization of the IPM application profile to provide interoperability among different vendors is important. The profile defines every device object that are to be used in the profile, such as light sensor, temperature sensor and light switch etc. The early demonstration of the prototype IPM system will facilitate the actual deployment phase of the monitoring network requested by MELCO external customers. The monitoring system will be a useful tool for maintenance and process control, assisting in repair/rebuild decisions and optimizations.

Technical Discussion: The prototype devices consist of IEEE 802.15.4 MAC and PHY, ZigBee networking protocol and application support sublayer (APS) functionalities, which are compliant with the official ZigBee specifications and implemented on Mitsubishi M16C boards. The application devices are programmed on M16C chips. The ADXL210E by Analog Devices is used for vibration measurements within 0-1kHz range. The vibration data is dynamically preprocessed prior to transmission to the central monitoring unit (CMU) according to network conditions and measurement values.

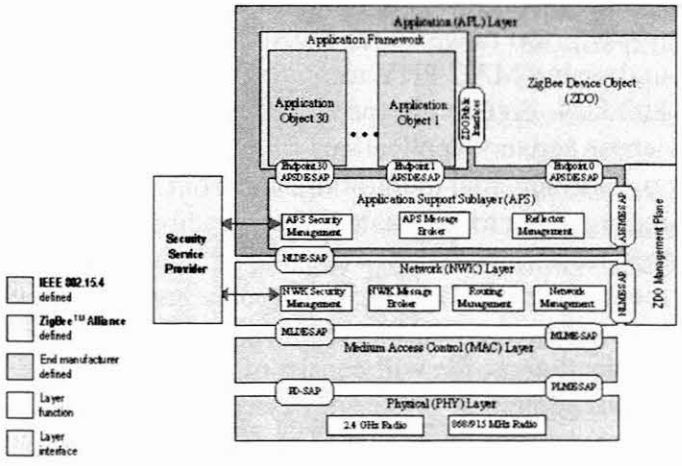
Collaboration: MEUS, Shaden, Sentansoken.

Future Direction: Development and testing of the IPM protype will be completed in September 2004. The profile authorization request will be issued in July 2004 in ZigBee member meeting; and active participation in the profile task group for standardization is targeted. We would like to extend this project to structural and environmental monitoring applications (e.g., cherry fields). Location awareness will be added to the system features. Software will be provided to control data acquisition hardware, to provide an interface to the controller and to interpret acquired signals to ease diagnosis.

Contact: Zafer Sahinoglu
http://www.merl.com/projects/zigbeeipm/

Lab: MERL Technology Lab
Project Type: Advanced Development

ZigBee Security Toolbox



The ZigBee Alliance is developing a very low-cost, very low power consumption, two-way, wireless communications standard. Solutions adopting the ZigBee standard will be embedded in consumer electronics, home and building automation, industrial controls, PC peripherals, medical sensor applications, toys and games. Many of these applications have security needs. The ZigBee Alliance is developing the protocols required for a wireless, adhoc network in order to make it as secure as possible for the average consumer. Security services provided for ZigBee include methods for key establishment, key transport, frame protection, and device management. These services form the building blocks for implementing security policies within a ZigBee device.

Background and Objectives: MERL has participated and contributed to the ZigBee Alliance in terms of key exchange protocols and secure routing techniques. Presently, we are working on secure installation of a ZigBee network in both the home and commercial environments. Another concern with this type of network is the topic of Denial of Service. If and when ZigBee networks become more widespread, the potential for an attack on those networks becomes more probable. One popular form of attack is making the network unusable by the user by flooding the network with packets or interfering with the radio traffic.

Technical Discussion: Network security for any type of network becomes necessary as that network becomes more widespread. Security for networks involves a family of protocols that prevent adversaries from eavesdropping by use of encryption, spoofing by use of authentication, and replay attacks by use of the nonce. Other areas of security involve key distribution and exchange protocols. Keys are required for encryption and decryption of messages. Key distribution centers (a.k.a. trust centers) must be maintained within a network. Recovery protocols for the nodes that drop out due to loss of power or some other malady must be able to regain access to the network quickly and seamlessly.

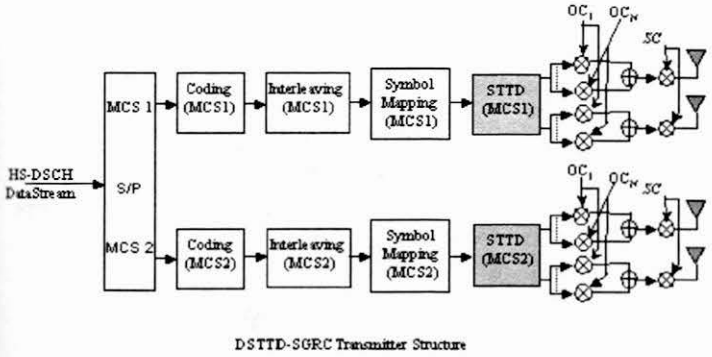
Collaboration: MERL is to pursue collaboration with two departments within Johosoken.

Future Direction: More advanced techniques for routing around malicious nodes, once discovered, is under discussion. Secure methods for authenticating a new node that is trying to associate with a network must be developed. Distribution of group keys, which is a shared key within a select set of nodes, are also being considered for future development.

Contact: Johnas Cukier, Ghulam Bhatti
http://www.merl.com/projects/zigbeesec/

Lab: MERL Technology Lab
Project Type: Research

Sub-Group Rate Control for MIMO Systems in 3GPP Standard



DSTTD-SGRC Transmitter Structure

The third generation partnership project (3GPP) is currently considering MIMO technology for next generation wireless cellular systems to increase the system capacity, achieve higher peak rates, and increase the coverage area. For transmitter with up to 4 transmit antennas, we have proposed a MIMO technique called double space-time transmit diversity with sub-group rate control (DSTTD-SGRC) to achieve these goals. DSTTD-SGRC requires lower feedback bandwidth, is backward compatible, and works even for mobiles with only two receive antennas. It has also been shown to perform well even at higher mobile speeds.

Background and Objectives: MIMO technology is currently a work item in 3GPP. The spatial channel model, which plays a crucial role in determining the capacity gains achievable by MIMO, has also been finalized in 3GPP. The current focus in MIMO systems is the finalization of a system-level simulation methodology and performance criteria to evaluate and compare candidate proposals. Several MIMO proposals have been submitted by participating companies, for example, Per Antenna Rate Control by Lucent, Multi-path Diversity by Nortel, etc. DSTTD-SGRC provides an acceptable trade-off between the conflicting demands of higher system capacity, robust performance, and minimal complexity.

Technical Discussion: In DSTTD-SGRC, the four transmit antennas are divided into two sub-groups. Adaptive modulation and coding along with STTD-based transmission are used by each group to transmit data. Two transmit antennas belong to each sub-group transmit and use the same modulation and coding scheme (MCS) and transmit at the same rate. The data rate of each sub-group is either independently or jointly adjusted based on feedback from the mobile about the MIMO channel state. By combining sub-group rate control with the STTD structure, which is mandatory in Rel'99 mobiles, higher throughput, enhanced link diversity, as well as backward compatibility are achieved. The enhanced link diversity also makes the system more robust to feedback delay-induced estimation errors. In addition, it requires only half the feedback of schemes that control the MCS for every transmit antenna required.

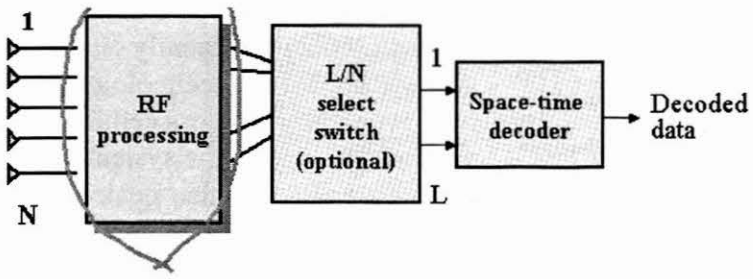
Collaboration: Johosoken

Future Direction: Future activities include building a system simulator to evaluate performance of proposed scheme. Several aspects such as pilot design, feedback mechanisms, etc., are to be finalized.

Contact: Neelesh Mehta, Jinyun Zhang
http://www.merl.com/projects/dsttd_sgrc_3gpp/

Lab: MERL Technology Lab
Project Type: Initial Investigation

Joint RF-Baseband Processing for Multiple Antenna Systems



Receiver with RF pre-processing

While MIMO systems can achieve significantly higher rates, they require complex hardware and considerable signal processing power. We propose and investigate novel joint RF-baseband designs that tackles this problem. A joint RF-baseband design combines RF pre-processing with selection, if necessary, and subsequent baseband signal processing. We have determined the optimal joint designs for the two cases when the spatial RF pre-processing filter is based on the time-varying instantaneous channel state and when it is based only on the time-invariant statistical information, such as the spatial covariance of the transmitter and receiver. In both cases, it performs close to or, even, as well as the full complexity solution.

Background and Objectives: MIMO can substantially improve the information rates. However, the increased system and hardware complexity (in the form of multiple RF chains) that MIMO entails has inhibited its wide spread adoption. MIMO techniques fall under two broad categories: spatial diversity, in which only one stream is transmitted from all the antennas, and spatial multiplexing, in which multiple streams are transmitted. For both these categories, antenna selection, in which only a subset of the signals from the available antenna elements is adaptively chosen and processed, has been the conventional technique of choice to address this problem. However, antenna selection leads to degradation in the performance. Joint RF-baseband design addresses this problem.

Technical Discussion: For spatial diversity, RF-baseband design maximizes the signal to noise ratio (SNR) after baseband signal processing. For spatial multiplexing, the metric that is optimized is the information theoretic rate in bits/sec/Hz. Channel statistics-based RF-baseband design instead maximizes the average output SNR or ergodic capacity. To accommodate practical constraints of current variable phase-shifter technology, sub-optimal solutions that use only phase-shifters and adders is also considered for all the cases and shown to lead to a negligible loss in performance.

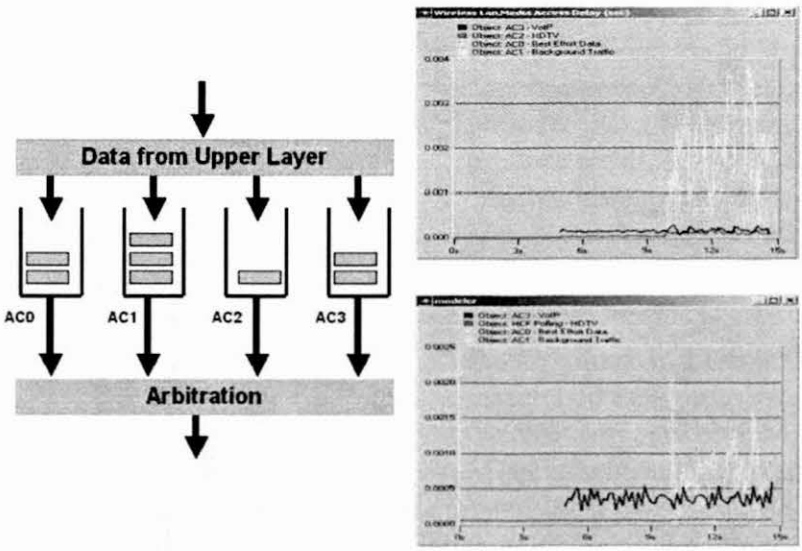
Collaboration: Johosoken

Future Direction: Investigate other benefits of RF pre-processing such as dynamic range enhancements. Adapt RF pre-processing ideas to other transmitter and receiver structures.

Contact: Neelesh Mehta, Andreas F. Molisch
http://www.merl.com/projects/rf-baseband-processing/

Lab: MERL Technology Lab
Project Type: Research

IEEE802.11e MAC for QoS



See Color Figure 7

IEEE802.11e is a new 802.11 Wireless LAN (WLAN) standard which supports applications with Quality of Service (QoS). It is widely considered to play a major role in multimedia wireless home networks. In the project, we built a complete 802.11e simulation platform, and evaluated its performance with a variety of environments and applications.

Background and Objectives: The new IEEE802.11e standard is paving a way for customers to have a multimedia wireless network at home. Many electronics and communication companies have been developing new home electronics and network devices with 802.11e network capability. The objective of the project is to develop 802.11e simulation software, and research the possibility to create a new home network business for MELCO.

Technical Discussion: IEEE802.11e standard adds a new function called HCF (Hybrid Coordination Function) which includes both contention-free and contention-based channel access methods in a single channel access protocol. HCF supports both differentiated and parameterized QoS through prioritized contention-based and controlled contention-free medium access. The 802.11e simulation is developed to simulate 802.11e channel access mechanism based on OPNET. It has a very friendly user interface to simulate various application scenarios.

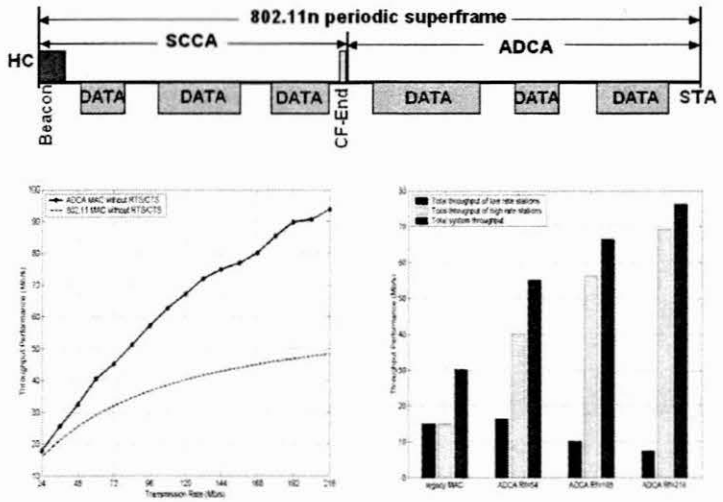
Collaboration: This project is in collaboration with SentanSoken and MDEA.

Future Direction: Implement 802.11e QoS functions in the broadband AV home network.

Contact: Daqing Gu, Jinyun Zhang
http://www.merl.com/projects/ieee802.11emac/

Lab: MERL Technology Lab
Project Type: Advanced Development

High Efficient MAC for Next Generation Wireless LAN



To increase network throughput, high efficient Medium Access Control (MAC) with Quality of Service (QoS) support is needed for next generation high-speed Wireless LAN (WLAN). In the project, we developed a high efficient MAC scheme with QoS consideration for high throughput 802.11n WLAN standard which is under development.

Background and Objectives: With the increase of bandwidth required for new applications, IEEE802.11 Working Group recently established Task Group n (TGn) to develop a high throughput WLAN standard called 802.11n. The new 802.11n standard is a modification to the existing 802.11 standards and can achieve a maximum throughput of at least 100 Mbps. The objective of the project is to develop a high efficient MAC targeting to contribute to 802.11n standard, and increase MELCO's portfolio in wireless communications.

Technical Discussion: With the increase of bandwidth required for new applications, IEEE802.11 Working Group recently established Task Group n (TGn) to develop a high throughput WLAN standard called 802.11n. The new 802.11n standard is a modification to the existing 802.11 standards and can achieve a maximum throughput of at least 100 Mbps. The objective of the project is to develop a high efficient MAC targeting to contribute to 802.11n standard, and increase MELCO's portfolio in wireless communications.

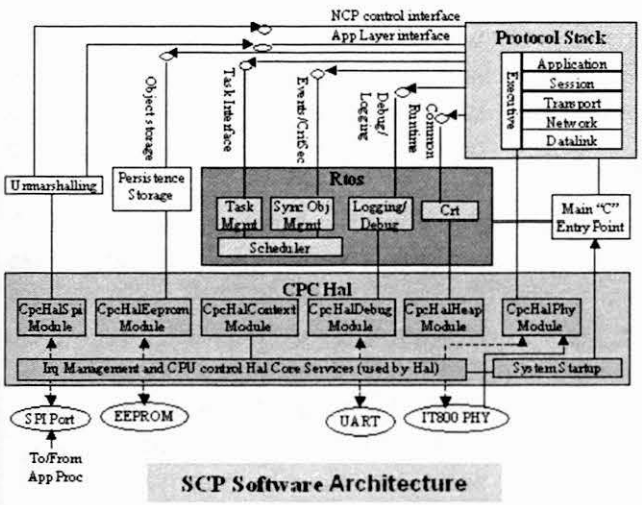
Collaboration: This project is in collaboration with Johosoken.

Future Direction: Improve the proposed MAC scheme, and contribute to 802.11n standardization body.

Contact: Daqing Gu, Jinyun Zhang
http://www.merl.com/projects/highefficientmac/

Lab: MERL Technology Lab
Project Type: Research

Simple Control Protocol for Power Line Communications



SCP Software Architecture

The Simple Control Protocol (SCP) is a non-IP based home networking protocol created by Microsoft Corporation. It is a peer-to-peer, physical medium independent protocol that allows manufacturers to produce small, low-cost, intelligent devices, which communicate with each other securely and robustly. We developed SCP software for low-speed power line communications.

Background and Objectives: MERL has researched SCP protocol and implemented the SCP software onto Mitsubishi/Renesas M16C platform using standard power line as physical communication medium. The hardware related software was written by MERL to control physical layer and interact with the MAC layer. MERL also developed the protocol related software to interact with the upper layers MERL assisted in the hardware and chip development for Renesas. In addition, MERL developed the applications and demos shown at the CEATEC 2002 & 2003, 2003 MERL Road show, UPnP Forums, IEEE CCNC 2004, and Connections 2004. MERL participated in the development of the UPnP/SCP bridge and is now developing a ZigBee/SCP bridge. MERL has been providing technology transfer to Renesas in terms of support, training, and documentation.

Renesas has shipped M16C/6S chip to customers, which include Matsushita Electric Works, Digital Stream Corporation, and Industrial Technology Research Institute of Taiwan. The mass production is expected in later this year. Microsoft, Matsushita Electric Works, Digital Stream, ITRAN, and Renesas agree to establish the SCP user group in Japan for promoting SCP technology and sales.

Technical Discussion: A SCP network consists of a physical network and multiple logical networks. The physical network provides a common communications medium among the devices within logical networks. SCP is a complementary technology to Universal Plug-n-Play (UPnP), which is an IP based home network protocol. The devices in SCP network can be bridged into UPnP network to communicate with other devices in an IP network. SCP devices can also be bridged into other networks. Therefore, SCP technology can be used to develop new low-cost hybrid networking systems such as sensor networks, energy monitoring system, automated meter reading, etc.

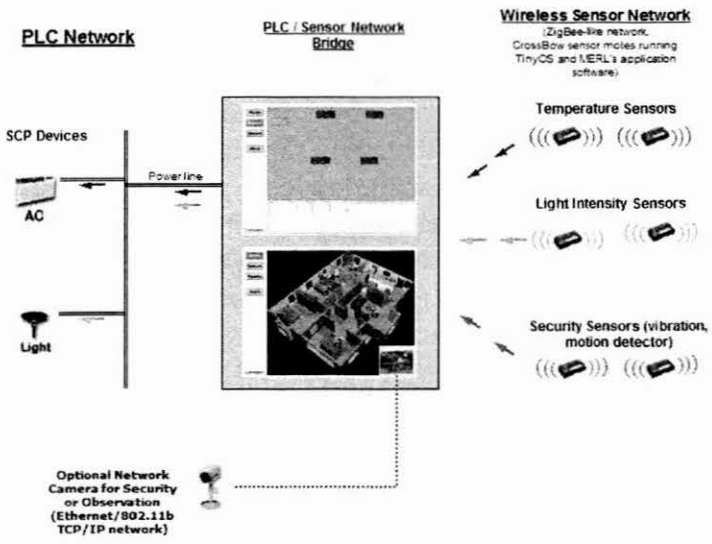
Collaboration: During the development of the SCP technology, MERL has collaborated with Renesas, Johosoken and ITRAN. MERL is now developing a ZigBee/SCP bridge with Renesas.

Future Direction: The SCP is an emerging home networking technology. MERL has developed its SCP/UPnP/ZigBee home networking system and will improve this networking system. MERL is continuously deploying new applications that utilize SCP technology.

Contact: Jianlin Guo
http://www.merl.com/projects/SCP/

Lab: MERL Technology Lab
Project Type: Advanced Development

Simple Control Protocol and ZigBee Bridge for Home Networking



Future homes will utilize advanced wired and wireless networks for connectivity. We want to demonstrate some of the underlying communication technologies that form the basis of such networks. Specifically, we demonstrate our bridging technology for Simple Control Protocol (SCP), a power-line communications network, and ZigBee, a wireless sensors network. Both of these protocols offer a low cost and low bit-rate connectivity solution primarily to provide the controlling functions in a home environment. Both technologies are currently being developed at MERL.

Background and Objectives: Recent developments in power line data transmission and low bit rate radio enables a whole new range of devices to effectively communicate. Typical household devices (e.g., lights, doorbells, appliances, entertainment, security systems) are now on the verge of being seamlessly integrated into home networks by using both wired and wireless technologies. Essentially, the home network no longer needs to be consisting of just a cable modem and wireless access point that links together high end computer and audio visual equipment. Rather, it will consist of many additional low cost devices. This will allow homeowners a complete home network with the capability to control all aspects of their environment as well as introduce new applications for the home.

Technical Discussion: Both the SCP and ZigBee technologies are non-IP based networking protocols. ZigBee is especially suitable for sensors and/or control networks. The SCP, on the other hand, utilizes existing infrastructure. Bridging the two domains will offer the best of both technologies.

Collaboration: We are working closely with Renesas and Johosoken.

Future Direction: The SCP-ZigBee Bridge can be extended to include IP to give it access to the internet. It can function as a gateway to the outside world.

Contact: Jinyun Zhang, Ghulam Bhatti, Georgiy Pekhteryev **Lab:** MERL Technology Lab
http://www.merl.com/projects/scpzigbeebridge/ **Project Type:** Advanced Development

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 databases. Digital Video enriches our everyday lives by enabling various forms of communication and entertainment.

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. The three areas of concentration are described below.

- Compression: In the past, we have focused mainly on encoder optimization of traditional coding schemes. More recently, we have engaged in the design of new schemes such as the coding of multi-view video for 3D rendering and distributed source coding.

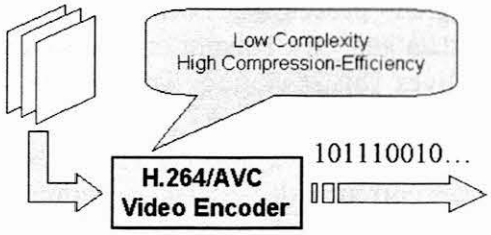
- Distribution: Most of our effort in this area is focused on video transcoding, where MERL has developed award-winning architectures and algorithms for converting video to specific bit-rates and spatio-temporal resolutions. Streaming of JPEG 2000 images sequences based on region-of-interest is also an active area at the moment.

- Analysis: Multimedia storage and retrieval, including video indexing, summarization and audio-visual event detection are active areas of research at MERL. We apply this expertise to integrated systems such as DVD recorders, PVR devices, and surveillance systems.

In addition to the above, active participation and contribution to MPEG and other DTV-related standards have always been a major part of our activity. In the pages that follow, we provide a brief overview of related projects that we have been working on during the past year and have substantial results to report on.

Project Descriptions

H.264/AVC Encoder Optimization



H.264/AVC is the latest video compression standard. At the same video quality, it could achieve about 50% bit-rate saving over MPEG-2. Due to its high compression efficiency, it is expected to have a wide range of applications, including video conferencing, mobile TV broadcasting and high-definition DVD. However, its encoding complexity is extremely high. The goal of this project is to develop low-complexity H.264/AVC coding techniques that can still achieve state-of-the-art video compression.

Background and Objectives: H.264/AVC video coding introduces substantially more coding tools and coding options than earlier standards. Therefore, it takes much more computational complexity to achieve the highest potential coding gain. Our objective is to develop low-complexity video coding techniques that do not compromise video coding quality. We expect these techniques will be used to develop cost-effective H.264/AVC encoder and transcoder products.

Technical Discussion: Rate-distortion optimization using the Lagrange multiplier method is widely used in video coding. In macroblock coding-mode decision, the rate-distortion optimization evaluates the Lagrange cost for each candidate coding-mode for a macroblock and selects the mode that yields the minimum cost. The computation of the optimized mode decision is very intensive. Our approach uses intensive transform-domain processing such that the optimal mode decision is achieved with significantly reduced computational complexity. Specifically, we achieve the computational savings by calculating the distortion and the residual error in the transform-domain. We show that the distortion calculation can be performed efficiently in the transform-domain such that the inverse transform and reconstruction of the pixels can be omitted. We calculate the residual error in the transform-domain by taking advantage of the fact that the transform of several intra prediction signals (DC, Horizontal, and Vertical) can be very efficiently calculated, and their transform coefficients have few non-zero entries. We also use transform-domain processing to directly convert DCT coefficients (used in MPEG-2, MPEG-4 etc.) to H.264/AVC transform (HT) coefficients. Our transform-domain approach outperforms the conventional pixel-domain DCT-to-HT conversion both in complexity and in quality. This technique is useful in transcoding DCT-based video to H.264/AVC format.

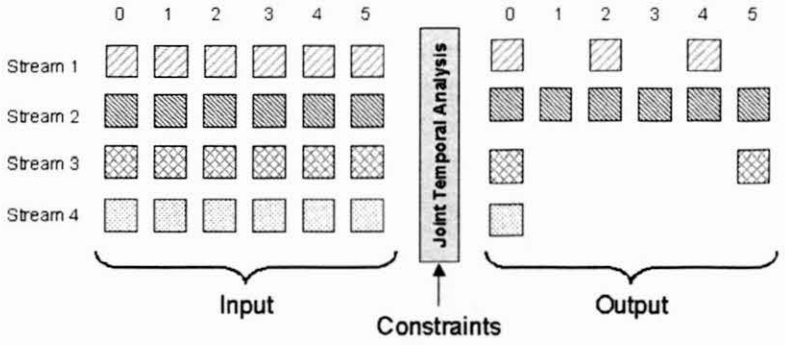
Collaboration: This project is done in collaboration with the Multimedia Information Coding & Transmission Technology Department at Johosoken.

Future Direction: Implement these techniques in H.264/AVC encoder and MPEG-2 to H.264/AVC transcoder. Extend these techniques to interlaced video. Further reduce the complexity of mode decision and motion estimation.

Contact: Jun Xin, Huifang Sun, Anthony Vetro
http://www.merl.com/projects/avc-optimization/

Lab: MERL Technology Lab
Project Type: Research

Variable Rate Sampling of Multiple Image Sequences



This work aims to add functionality to existing digital video recorders of surveillance video. In particular, assuming a multi-camera recording system, we investigate techniques that will dynamically record more frames for active cameras, while keeping the same memory constraints.

Background and Objectives: Recording of surveillance video is a memory intensive operation. Typical surveillance systems include digital recorders that are capable of storing compressed images or video from up to 16 different cameras. Most systems contain a hard-disk drive (HDD) of 120GB or 240GB in size. Given this fixed memory, and the desire to record as much content as possible, a lower temporal rate of the video is typically recorded. In current recorder systems, the lower temporal rate is achieved by simply uniformly sampling the input frames. The major drawback is that frames from inactive inputs are recorded the same as active inputs. In this way, the memory is not being efficiently utilized since inactive inputs are unnecessarily being sampled, and information contained in active sequences is lost due to under-sampling. Clearly, a variable sampling of the input image sequences would overcome these drawbacks.

Technical Discussion: This project addresses the problem of sampling frames from multiple image sequences with the goal to minimize distortion of the sampled sequences subject to a total memory constraint. A joint temporal analysis over a window of all input image sequences is performed to determine the optimal set of frames to be recorded. Compared to uniform sampling, simulation results indicate that significant reductions in distortion could be obtained with the proposed variable sampling approach.

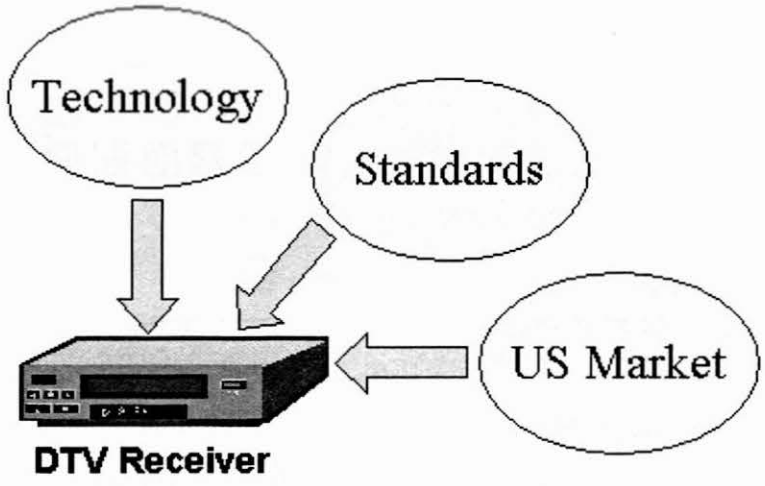
Collaboration: This project is done in collaboration with the Audio-Visual Information Technology Department at Johosoken.

Future Direction: Consider addition metrics to decide the variable sampling of frames and investigate feasibility of algorithm implementation on low-end CPU.

Contact: Anthony Vetro
http://www.merl.com/projects/variable-sampling/

Lab: MERL Technology Lab
Project Type: Research

DTV Receiver: Technology, Standards and LSI



MERL is currently involved in three activities related to the development of DTV receivers including down-decoding technology, post-filtering technology and DTV related standardization in the US. We coordinate closely on these topics with our DTV related business units and the domestic labs of MELCO.

Background and Objectives: We develop technologies that aim to lower the cost of DTV receivers and improve the quality of playback. Additionally, we track standards and regulatory activities so that developed products will maintain compliance with the latest standards and mandates.

Technical Discussion: Down-decoding is essentially a filtering technique that enables decoding of HDTV bitstream with reduced memory and memory bandwidth and is a key feature of a low-cost DTV receiver that receives HDTV input bitsreams, but displays SDTV output resolution. Post-filtering is a noise removal technique that enables improved playback of highly compressed video stored on HDD/DVD and targets such devices that are responsible for storage and playback applications. It is also important to track related standardization activities in the US that impact receiver LSI design. Issues that have come up recently include content protection systems to meet compliance to broadcast flag mandate, an enhanced VSB modulation scheme, new audio-visual codecs such as Windows Media and H.264/AVC for video, as well as the standard for unidirectional receiving devices that has been agreed upon between Cable and CE industries.

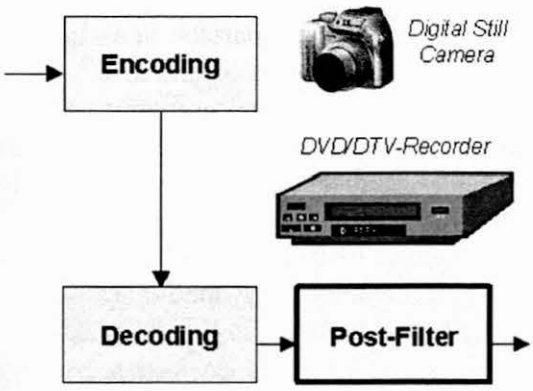
Collaboration: DTV related business units, as well as related departments in corporate research labs

Future Direction: Continue tracking related standardization activities and developing related technologies.

Contact: Anthony Vetro, James Fang, Huifang Sun
http://www.merl.com/projects/dtv-receiver/

Lab: MERL Technology Lab
Project Type: Advanced Development

Post-Filter for Image/Video Artifact Reduction



High compression techniques are required in many imaging and video applications, including digital cameras, HDTV broadcast and DVD. However, visual artifacts, such as block noise, may be present in the decompressed images due to the high compression. Post-filtering techniques are used to suppress these visual artifacts and improve the quality of the image during playback. The algorithm that we have developed is able to effectively preserve the image quality, i.e., keeping sharp edges and details, while reducing the artifacts. The quality is better than other known techniques and the complexity is also much lower.

Background and Objectives: Visual artifacts are normally present in decompressed images due to coarse quantization and coefficient truncation. Blocking and ringing artifacts are the two major coding artifacts caused by high compression. Many post-processing approaches have been proposed to remove the visual artifacts either from the spatial domain or the frequency domain. They attempt to adaptively filter each pixel in the image based on quantization parameter and neighboring information. Since these filtering methods are pixel-by-pixel operations, they inevitably introduce undesired smoothing effects to pixels without artifacts. Classification-based methods have been recently proposed to detect the artifacts before applying the post-filtering. However, these methods mainly concentrate on blocking artifacts, and are less effective in removing ringing artifacts.

Technical Discussion: We have proposed a new adaptive approach for both blocking and ringing artifacts removal. The proposed method is based on edge information extraction and edge preserving filtering. Since coding artifacts are observed as certain patterns to the human visual system, we first apply pattern classification techniques to identify different type of artifacts and then perform the filtering accordingly. Our strategy is as follows: 1) form an edge map based on the local statistics; 2) according to the edge map, detect the blocking artifacts and classify the coding blocks into three (smooth, edge and texture) categories; 3) apply a 1-D low-pass filter to reduce the blocking artifacts and a 2-D fuzzy identity filter to reduce the ringing artifacts. Since the fuzzy identity filter is applied to the edge blocks only and it possesses a good edge preserving property, and the filtered images look sharp and clean. We have also developed practical implementation of this algorithm suitable for LSI design.

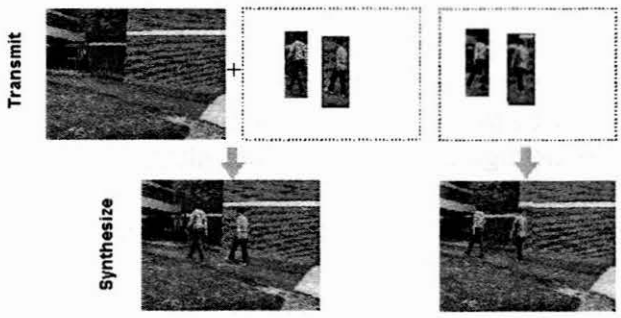
Collaboration: This project is done in collaboration with the Multimedia Processor Department at Johosoken, as well as the University of Delaware.

Future Direction: Continue to research methods that would improve the image quality.

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

Lab: MERL Technology Lab
Project Type: Research

ROI-Based Streaming of JPEG 2000



The JPEG 2000 image coding standard provides excellent compression performance in comparison to earlier image coding standards, and also provides a scalable representation of the coded image or image sequence. With the scalable representation, different spatial-temporal resolutions, levels of quality, and Region-of-Interests (ROIs) may be easily accessed or streamed in a progressive manner. In this project, we consider technology to enable progressive streaming of images and images sequences according to ROI information, which may be obtained in a automatic or semi-automatic way.

Background and Objectives: As an example, consider a Digital Video Recorder that is used for surveillance applications. Assume this box stores multiple image sequences from different camera inputs in a JPEG 2000 format. The stored bitstreams are likely to contain several quality layers; ROI information for each stream may be determined during recording or offline after recording. In any case, if remote access to a particular stream is desired, the image sequence may be streamed progressively according to the ROI(s) and quality layers to overcome limitations in bandwidth. The reordering of information in the originally coded bitstream to the ROI-ordered bitstream that is transmitted over the network is the primary focus of our work.

Technical Discussion: One way to achieve an ROI-based representation of an image is by the well known max-shift method, which essentially alters the image data by scaling coefficients such that ROI coefficients are coded prior to background coefficients. In this way, the ROI is reproduced with higher quality at the receiver. The primary drawback to this method is that the original image data needs to be altered according to the ROI info at the time of encoding. This is a serious drawback if (a) the ROI info is not known at the time of decoding, (b) the image is already encoded and/or (c) the ROI info is provided dynamically from an external source. Our proposed method overcomes these drawbacks, while achieving the progressive streaming functionality based on ROI information.

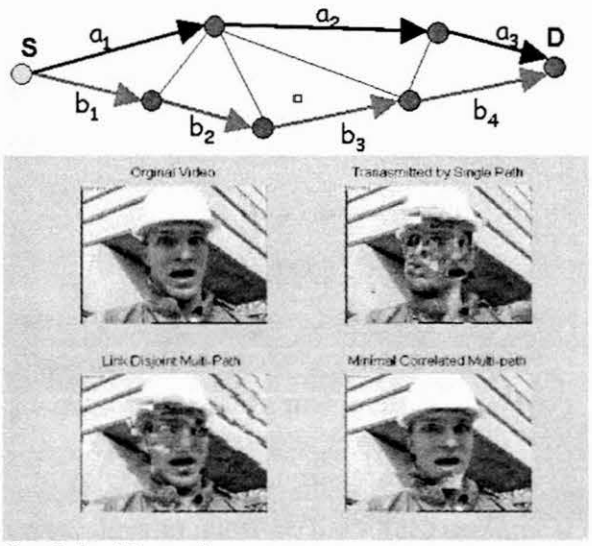
Collaboration: This project is done in collaboration with the Physical Security Group at Sentansoken.

Future Direction: Performance and quality optimization will be investigated further.

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

Lab: MERL Technology Lab
Project Type: Research

QoS-Enhanced Multi-path Overlay Network Support for Media Streaming



See Color Figure 8

Overlay network is regarded as a promising solution for providing high-level network services over today's Internet. This project investigates multi-path overlay network support for audio-video streaming applications. Our approach can provide end-to-end QoS through multi-path transmission on the overlay network.

Background and Objectives: With the fast growing development of Internet, audio-video streaming over Internet is in great demand. Streaming media is characterized by its strict delay constraint and sensitivity to packet loss. The current infrastructure of packet networks has not been sufficiently developed to support such challenging applications. Various techniques have been proposed to improve end-to-end streaming media quality by using special coding technique and path diversity. For example, multiple description coding (MDC) and Fine-Granularity-Scalability (FGS) partitioning are two such schemes. These coding schemes work particularly well on uncorrelated paths in which it is unlikely that the network paths simultaneously suffer from packet losses. While most of the previous work focuses on how to generate multiple related sub-streams, efficient multi-path selection methods to achieve good path diversity have not been fully investigated in the literature. This projects aims to propose QoS-enhanced multi-path selection methods for audio-video streaming applications.

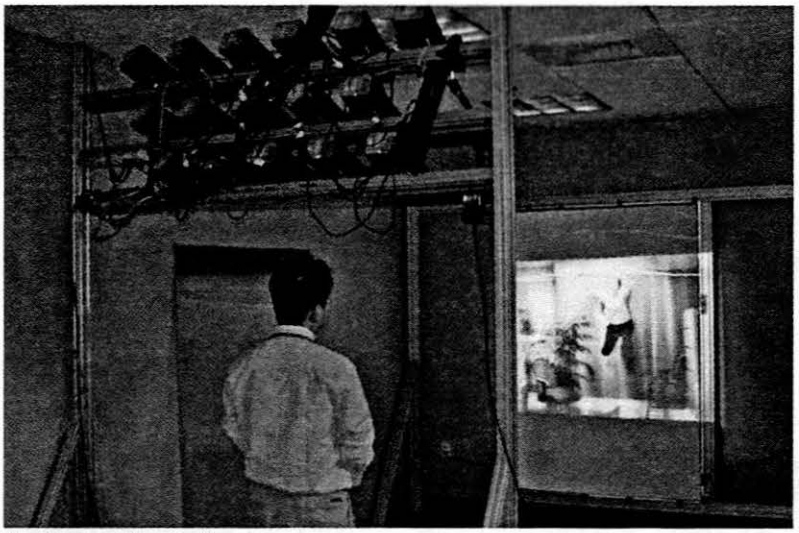
Technical Discussion: An overlay network is a virtual network built over an existing network as substrate. Overlay network brings more flexibility to add higher layer functionalities for today's Internet to support QoS. This project assumes an overlay network architecture with minimal assumptions about the knowledge of the underlying network. We first propose a new QoS metric link correlation and a path correlation model for multi-path selection problem. After discussing the tractability of minimal correlation multi-path selection problem, we present an efficient algorithm called correlation cost routing to select multi-path in overlay networks. The simulation results show that the average peak signal-to-noise ratio (PSNR) of the transmitted multiple descriptions coding (MDC) video using our multi-path selection algorithm improves by up to 3.2 dB over maximally link-disjoint multi-path selection method. Furthermore our new algorithm is more efficient than previous methods since it shares the same complexity with Dijsktra algorithm.

Future Direction: The future work would be implementing the scheme in the Internet and measure its performance in real world using Internet testbed like planetlab.

Contact: Huairong Shao, Chia Shen
http://www.merl.com/projects/multi-path/

Lab: MERL Research Lab
Project Type: Research

3D TV



See Color Figure 9

Three-dimensional TV is expected to be the next revolution in the history of television. It has only recently become feasible to deal with the high processing and bandwidth requirements for real-time acquisition, transmission, and display of high-resolution 3D TV content. We have built a complete end-to-end 3D TV system. Our system performs real-time acquisition, transmission, and 3D display of dynamic scenes.

Background and Objectives: Our system uses existing broadband protocols and compression standards for immediate, real-world 3D TV experiments and market studies. This system can plug into today's digital TV broadcast infrastructure and co-exist in perfect harmony with regular TV. Today, digital broadcast networks carry hundreds of channels and presumably a thousand or more channels after the introduction of MPEG-4. This makes it plausible that a number of them will be dedicated to 3D TV. Similar to HDTV, the introduction of 3D TV can proceed gradually, with one 3D channel at first and more to follow, depending on market demand. Furthermore, our system demonstrates that 3D TV offers a richer, more immersive experience than regular TV. It increases entertainment value and realism without the encumbrance of special glasses.

Technical Discussion: In our system image acquisition consists of an array of hardware-synchronized cameras that capture multiple views of the scene. In order to deal with the high processing and bandwidth requirements, the system uses a fully distributed architecture with clusters of PCs. A multi-projector 3D display with horizontal parallax achieves large, high-resolution output images. The system is scalable in the number of acquired, transmitted, and displayed realtime video streams.

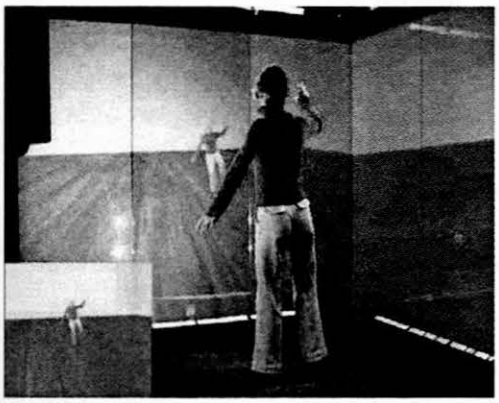
Collaboration: Microlens Technologies, North Carolina, USA.

Future Direction: Improved 3D display, multi-view video coding, computational improvement of the displayed video.

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

Lab: MERL Research Lab
Project Type: Research

Point-Based Graphical Objects for the MPEG-4 SNHC Standard



See Color Figure 10

We have proposed new syntax and semantics for the MPEG-4 standard to allow high-quality rendering of point-based graphical representations. Our contribution was conditionally accepted to be included in the MPEG-4 SNHC (Synthetic Natural Hybrid Coding) Animation Framework eXtension (AFX).

Background and Objectives: In recent years we have witnessed a dramatic increase in the polygonal complexity of computer graphics models. As a result, the average size of a rendered polygon is less than the size of a screen pixel. The processing of many small triangles leads to bandwidth bottlenecks and excessive rasterization requirements. This trend, combined with the associated overhead of managing, processing, and manipulating polygonal-mesh connectivity information, has led many leading researchers to question the future usefulness of polygons as the fundamental graphics primitive.

At the same time, advances in 3D digital photography and scanning technology have spawned systems that acquire both the geometry and appearance of complex, real-world objects. A major application for such 3D range cameras is the ready creation of digital 3D content, including some of the most visually stunning models to date. One of the challenges with these techniques is the huge volume of point samples they generate. These points constitute discrete building blocks of 3D object geometry and appearance much as pixels are the digital elements for images.

For these reasons, point primitives have experienced a major renaissance in recent years. Since then, considerable research has been devoted to the efficient representation, modeling, processing, and rendering of point-sampled geometry. Point primitives can be efficiently rendered and displayed in high quality, and novel algorithms for advanced shape modeling and digital geometry processing are entirely based on point samples. These efforts have spawned a new field called point-based computer graphics. The purpose of our research is to include point-based objects in the MPEG-4 SNHC standard.

Technical Discussion: For a dynamic point based object representation, e.g., 3D video fragments, we extend certain nodes in the existing MPEG-4 AFX format with additional attributes, such as normal and elliptical splat information. For a static point based representation, e.g., surfels, novel geometric attributes are specified through the use of MPEG-4 vector and normal nodes. By using existing nodes for normal data we can directly use the quantization procedure of MPEG-4, if necessary.

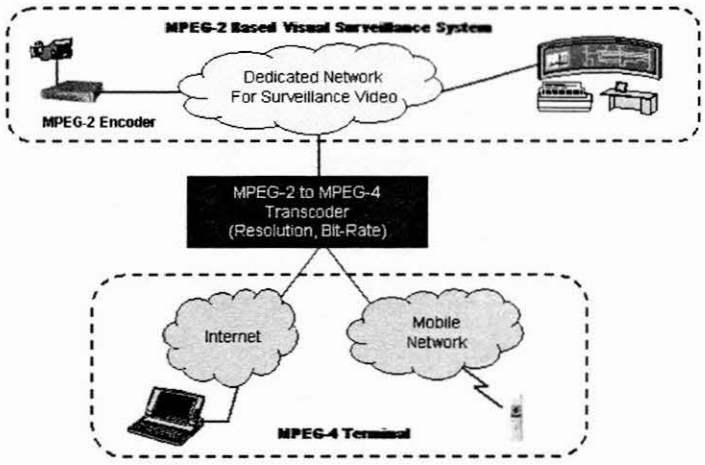
Collaboration: Stefan Wuermlin, Michael Waschbuesch, Eduard Lamboray, and Markus Gross, ETH Zurich.

Future Direction: Implementation of our proposed framework in the MPEG-4 AFX reference software.

Contact: Hanspeter Pfister
http://www.merl.com/projects/mpeg-4-snhc/

Lab: MERL Research Lab
Project Type: Research

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. Conventional cascaded approached decode the original bit stream to the spatial-domain, perform some intermediate processing, and then finally re-encode to a new bitstream. Our proposed architectures simplify this process, while still maintaining the picture quality. Many advanced features have been added, such as the ability to handle adaptive GOP input, improved error robustness in the input and output streams, as well as adaptive field/frame filtering to achieve a reduced spatial resolution.

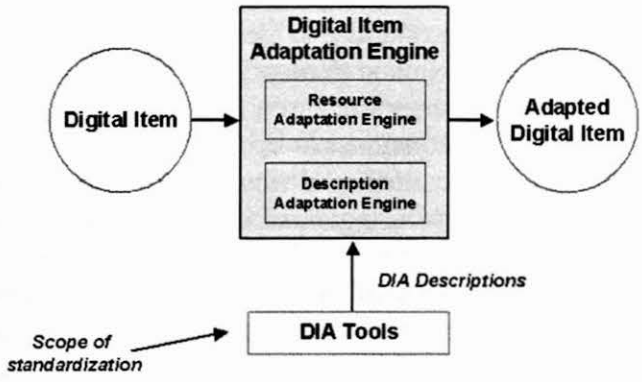
Collaboration: Technical aspects of this project are presently done in collaboration with the Multimedia Information Coding & Transmission Technology Department at Johosoken. Past collaborators include the Image Information Processing Department at Sentansoken and Princeton University.

Future Direction: Upgrade transcoding capability for new application domains.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

MPEG-21 Standards Activity



Universal multimedia access has become the driving concept behind a significant amount of research and standardization activity. It essentially refers to the ability for any type of terminals to access and consume a rich set of multimedia content. Ideally, this is achieved seamlessly over dynamic and heterogeneous networks and devices, independent of location or time, and taking into account a wide variety of possible user preferences. The vision of MPEG-21 is very much in line with this concept and standardizes a number of key components towards achieving this goal.

Background and Objectives: MERL has played an active role in defining the Digital Item Adaptation (DIA) standard, which is Part 7 of the MPEG-21, and has been finalized in December 2003. This standard specifies tools that are designed to assist the adaptation of Digital Items in order to satisfy transmission, storage and consumption constraints, and also provide a means for Quality of Service management by the various Users.

Technical Discussion: Given that adaptation always aims to satisfy a set of constraints, tools that describe the usage environment in a standardized way are essential. As a result, the DIA standard specifies tools that could be used to describe a wide array of user characteristics, terminal capabilities, network characteristics and natural environment characteristics. Among other things, the DIA standard also specified a means to describe the relationship between the constraints on the usage environment, the feasible adaptation operations that would satisfy these constraints and associated utilities (qualities) that result from particular adaptations. These tools provide a rich framework for multimedia adaptation and tie well with our ongoing research related to video transcoding.

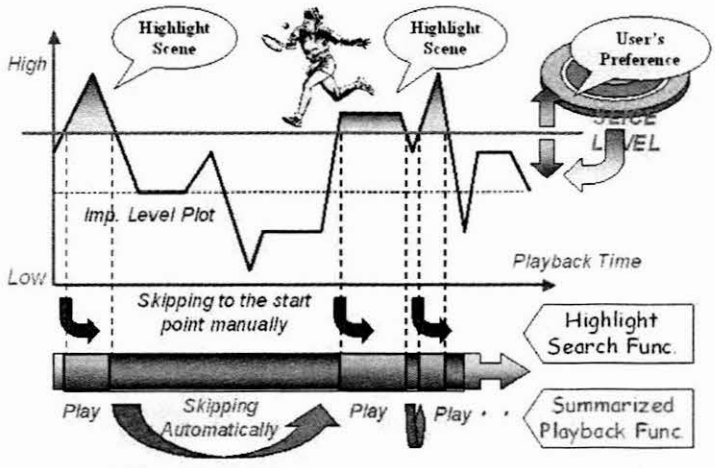
Collaboration: This project is done in collaboration with the Multimedia Information Coding & Transmission Technology Department at Johosoken.

Future Direction: Pursue widespread use of the standardized technology for business applications.

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

Lab: MERL Technology Lab
Project Type: Research

Video Summarization for PVR's



Personal Video Recorders have increasingly large storage capacity extending to almost 200 hours of content. Video Summarization is therefore essential to enable the consumer to skim through the content and view the content in differing detail depending on preference. We have developed a suite of summarization algorithms that are based on rapid audio-visual analysis in the compressed domain, and work well across diverse content genres.

Background and Objectives: 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. Our summarization algorithms are based on compressed domain analysis of the both the audio and the video. Since such analysis is fast, our algorithms have been easy to realize on our target platforms. Our target products include personal video recorders such as DVD recorders.

Technical Discussion: The technical challenge lies on two broad fronts. The first is audio-visual content analysis techniques that enable accurate content summarization over a broad range of content genres. The second is feasibility on our target platforms. Our algorithms therefore have to be computationally simple and robust to the high noise in broadcast video. Furthermore, they should have a small memory overhead. From the user's point of view, there should be ease of use and flexibility. We plan to meet the flexibility requirement by developing scalable summarization algorithms that allow the user to adjust the summary length at will. We will collaborate with our MERL colleagues in developing convenient user interfaces for the PVR application. We are currently working towards a prototype based on an audio-only solution.

Collaboration: Sentan-Soken, Joho-soken.

Future Direction: We will extend our content analysis techniques to generation of a more detailed description of the content.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

Off-the-Destktop Interaction and Display

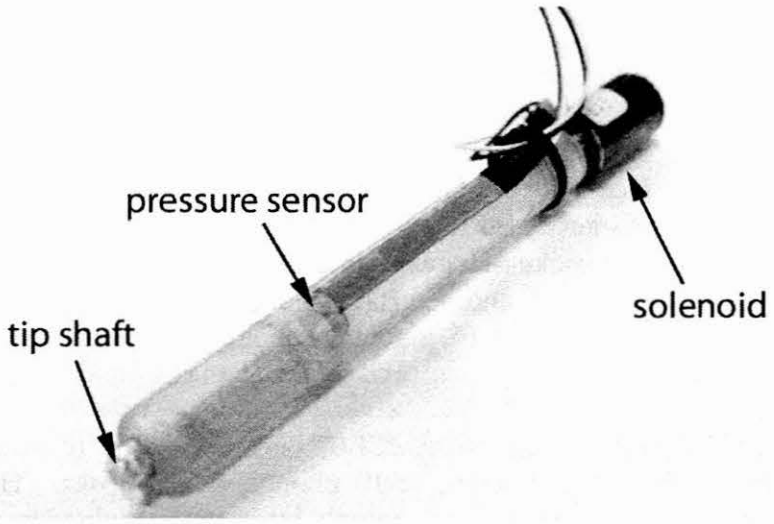
Computing devices are proliferating at all scales and sizes, from huge outdoor electronic signage and ever-larger digital televisions to navigation systems in automobiles and ever-smaller cell phones and projectors. Whereas computer desktop interfaces have become largely standardized, off-the-desktop devices have become the new frontier for human-computer interface research.

At MERL, we are exploring new user interface devices, paradigms, and interaction techniques that go beyond today's single-user mouse and keyboard interfaces. Our research investigates four areas of newly emerging technologies: (1) spoken-language interfaces for automotive and handheld devices, (2) multi-user touch technology and interfaces for shared displays, (3) calibration and interaction for projected displays, and (4) agent technology for simplifying human-computer interaction in an age of increasing complexity and information overload.

In the speech area, MERL's researchers are working with MELCO business units to bring innovative spoken-language interfaces into automobiles, cell phones, and kiosks. The SpokenQuery project and applications have developed new speech interaction analogous to typed search interfaces used on the Internet today. Other speech projects include FormsTalk, the Sphinx4, and ComBadge. As for multi-user shared-displays, MERL plays a leading role in the international research community in its development of the DiamondTouch table and related software technologies. Projected displays have been a strong area of research at MERL for some time, and this work continues with projects in Fiber Optic Projector Callibration, Multi-Projector Imagry, and Handheld Projectors. Finally, MERL continues to explore adding intelligence to interfaces through the projects in DiamondHelp, Human Robot Interaction, and COLLAGEN.

Project Descriptions

Haptic Stylus



When multiple users interact with a single large display, it is often difficult to distinguish one person's actions from those of other users. The problem is that the visual and auditory paths are shared. Our solution is to create individualized haptic sensations via each user's stylus. With an inexpensive mechanism, we can effectively emulate the feel of various switch and buttons.

Background and Objectives: Our work with DiamondTouch has shown the need for true multi-user input devices. We are investigating alternatives, including stylus-based systems. Because there is a physical device associated with each user, we can take advantage of this to provide personal feedback.

Technical Discussion: Our goal is to allow each user to "feel" graphical user interface elements through a stylus. We do this by adding a solenoid to move a small mass along the axis of the stylus away from the tip. The stylus also contains an inexpensive elastomeric force sensor which senses tip pressure. This system allows us to convincingly emulate the feel of various buttons and switches. We dynamically select among the various haptic profiles depending upon which interface element is currently under the stylus. The solenoid control loop runs on a dedicated microprocessor, so the host computer only needs to select the appropriate profile at each moment and the stylus does the rest.

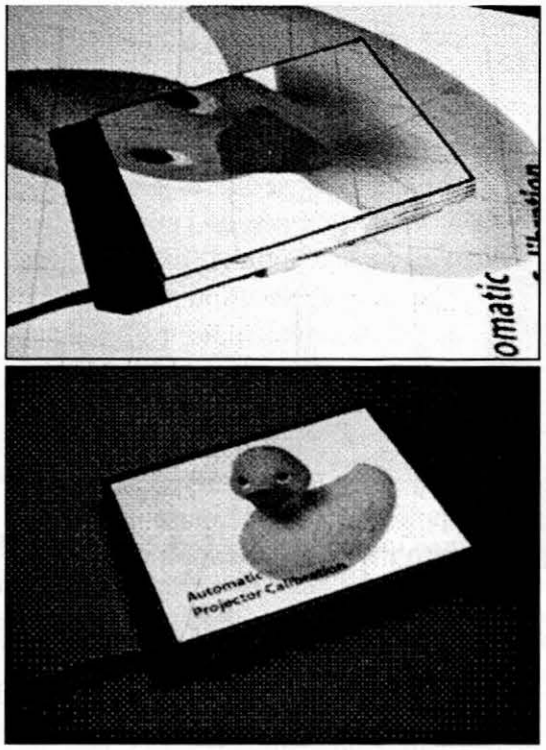
Collaboration: Johnny Lee and Scott Hudson of Carnegie Mellon University are collaborators on this project.

Future Direction: Ultimately, we hope that this haptic feedback system will become an integral part of our projector light pen system which is currently under development. In the short term, we intend to create a version for DiamondTouch to allow application exploration. We are also looking into licensing opportunities with current electronic whiteboard vendors.

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

Lab: MERL Technology Lab
Project Type: Initial Investigation

Fiber Optic Projector Calibration



In many projector-based systems, it is necessary to align the projection with the surface being projected upon. We automate this task via light sensors embedded in the projection surface.

Background and Objectives: MERL has done a large number of projects involving precision computer video projection. Examples include DiamondTouch, Shader Lamps, Interactive Retail Displays, etc. In all of these cases, the projection must be accurately aligned to the surface. In the past, we have used semi-automated procedures where a user must indicate certain target locations, allowing the various geometries to be calculated and the projection suitably warped to match the surface. Unfortunately, this is a tedious process prone to human error. Our new system is fully automated and can quickly and precisely locate key pixel positions. Hopefully, this will greatly increase the practicality of our various projector-based systems.

Technical Discussion: The calibration system uses light sensors which are embedded underneath the projection surface. Binary Gray code patterns are projected onto the surface to discover the precise pixel location of each sensor. For an XGA display, only 20 patterns are required to create a unique sequence at each pixel. To simplify the system, inexpensive fiber optics are used to carry the light from the sensing points to a single detector board. Once the precise pixel locations are known, the projection can then be suitably aligned, either through image warping or mechanical adjustment.

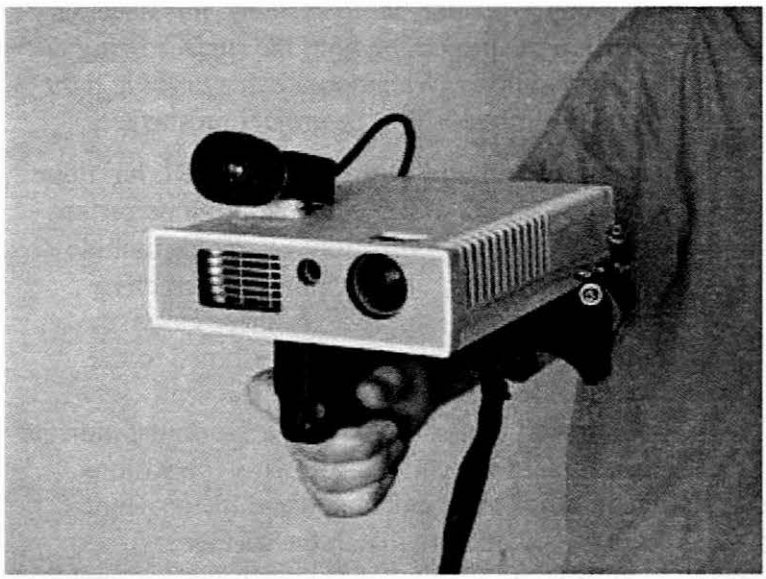
Collaboration: Johnny Lee and Scott Hudson of Carnegie Mellon University are collaborators on this project. Dan Maynes-Aminzade, presently at Stanford University, was also a key contributor.

Future Direction: We are pursuing various product possibilities within MELCO while exploring external licensing opportunities. The ability to rapidly align imagery to physical objects opens up exciting mew possibilities. Our next step will be to greatly enhance the speed of the system from 1 calibration/sec to 100 calibrations/sec using invisible IR patterns. This will allow real-time object tracking, and thus future device displays can become part of an environment rather than part of the device.

Contact: Paul Dietz, Ramesh Raskar, Jeroen van Baar
http://www.merl.com/projects/fiber-cal/

Lab: MERL Technology Lab
Project Type: Research

Handheld Projector



This work is on the use of handheld projectors for projected augmented reality, and on user interaction with the projection. The prototype device in the picture includes a projector, a camera, attached laser pointers for motion recovery, and a hand-grip with mouse buttons for interaction. The main application is projected augmented reality for task guidance, for example guidance of a maintenance task on a piece of equipment. The work also addresses projection of ordinary desktop applications, and mouse interaction with the projected application.

Background and Objectives: Handheld projection has the potential for new and versatile types of display. Projectors have traditionally been fixed devices requiring manual intervention for setup. Recent innovations have replaced the manual setup with automatic calibration via a camera. The next stage in projector evolution will be towards handheld projectors. The goal of this work is to demonstrate the feasibility of a handheld projector, and to develop innovative methods and applications.

Technical Discussion: The main technical requirement for handheld projection is to detect device motion, so that the projection can be stabilized on the display surface e.g. under hand-jitter. We attach laser pointers to the device, and observe the laser spots with the camera, to reliably recover device motion relative to the display surface. The second requirement for the device to be truly useful is a method to allow the user to interact with the projection. We provide a method that allows pointing motion of the projector to control a cursor, plus mouse buttons on the device to support the usual types of interaction. We are investigating three classes of application (i) projected augmented reality, (ii) projected desktop applications, (iii) guided computer vision in which the user employs the device to do a visible selection in the physical world, and then invokes an image processing or computer vision algorithm on the selected area.

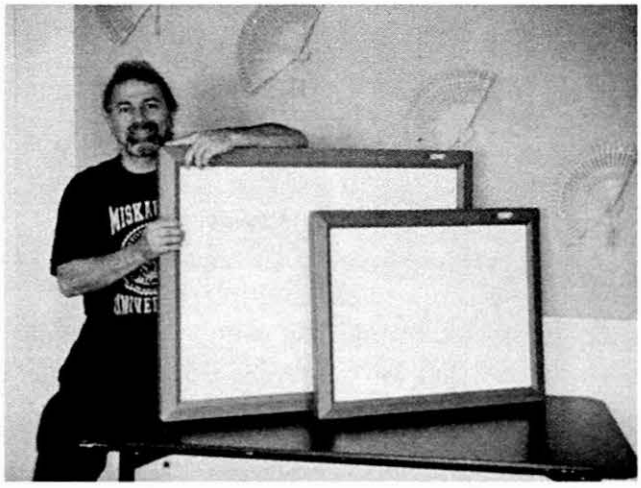
Collaboration: Dr. Shiotani, Sentan-soken.

Future Direction: Better motion recovery, a smaller device.

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

Lab: MERL Research Lab
Project Type: Research

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 prototype versions, 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.

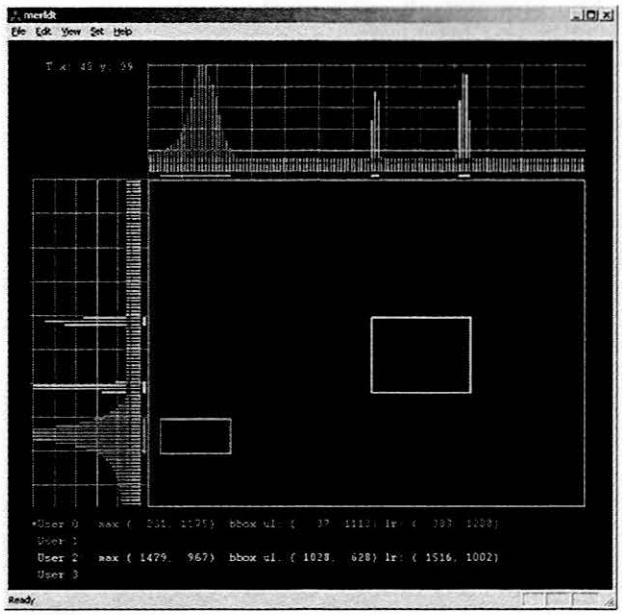
Collaboration: The DiamondTouch hardware project is a collaboration of MERL MRL and MTL; we are working with Johosoken, Kamaden and IDken to make best use of the hardware for securing future contracts for MELCO. Additionally, we are collaborating with many universities in researching further applications.

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
http://www.merl.com/projects/DiamondTouch/

Lab: MERL Research Lab
Project Type: Advanced Development

DiamondTouch Software Development Kit (SDK)



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). 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, and *-- DiamondTouch Applications --* for more details on applications.) Using the SDK, programmers can develop novel, multi-user applications in the language of their choice.

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 libraries to support DT application development; version 2.0 was recently released, which includes a much-improved API. The DiamondTouch hardware periodically produces frames of data indicating the proximity of the user's finger(s) to each antenna. The SDK reads these data frames from the device and provide access to the raw data as well as o 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. A more programmer-friendly API, provides access to more semantically oriented events. The SDK provides support for application development in a variety of languages (C/C++, Java, ActiveX Control) and includes a number of diagnostic (e.g., merldt) and utility applications (e.g., mouse emuation, projector calibration, thresholding, etc).

Collaboration: DiamondTouch is a joint project of MERL's Technology and Research Laboratories. MERL is collaborating with MELCO partners from Kamaden, Johosoken, Shaden, 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, Alan Esenther, Samuel Shipman **Lab:** MERL Technology Lab
http://www.merl.com/projects/dtsdk/ **Project Type:** Advanced Development

DiamondTouch Applications

The goal for this project is to develop novel applications to provide a shared focus of attention for collaborating users, exploiting DiamondTouch, a novel multi-user input device. 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 and entertainment, etc).

Background and Objectives: To date, most 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. This application provides a display containing 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, and thus backwards compatibility for legacy software. We have also implemented a number of entertainment and gaming applications to showcase DiamondTouch. Most recently we have begun to explore multi-display applications (e.g., table plus wall) for DT.

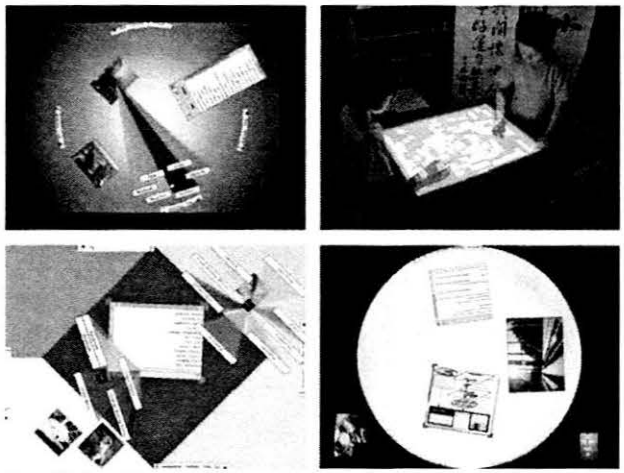
Collaboration: DiamondTouch is a joint project of MERL's Technology and Research Laboratories. MERL is collaborating with MELCO partners from Kamaden, Johosoken, Shaden and Sentansoken. We also have active collaborations with universities who will explore DiamondTouch as a collaborative input technology.

Future Direction: Our MELCO collaborators and their customers will use many of the applications we develop. We will work with them to refine the applications, and develop new functionality. We have also begun exploring external licensing opportunities, and are working with external groups to develop appropriate content and interactions. In the longer term we would like to investigate the use of DiamondTouch for remote collaboration, and continue development of multi-user, multi-display applications.

Contact: Kathy Ryall
http://www.merl.com/projects/DTApplications/

Lab: MERL Technology Lab
Project Type: Advanced Development

DiamondSpin



See Color Figure 11

DiamondSpin is an interactive, and platform independent, Java Tool Kit that allows multiple users to work on a digital tabletop simultaneously, in a truly around-the-table setting. DiamonmdSpin is a core project that has enabled a range of research investigations into large shared displays and off-the-desktop human computer interaction. Currently, MERL's own UbiTable project, and several research projects in other universities are based on DiamondSpin.

Background and Objectives: Large shared interactive surfaces offer great potential for ubiquity of computationally augmented tables and walls. Their affordances to multi-person interactivity, concurrent simultaneous direct manipulability, and individual user identification prompt new ways of thinking in terms of UI design and interaction technique development. This, in turn, requires us to re-examine the conventional metaphor and underlying system infrastructure, both of which have been traditionally geared towards mice and keyboard-based, single-user desktop computers and devices. DiamondSpin toolkit is a research testbed, a sandbox and an evaluation vehicle in our pursuit of understanding large shared interactive surfaces.

Technical Discussion: One of the fundamental research issues in the design of a true tabletop UI is how to afford face-to-face or corner-to-corner multi-user collaboration. We are experimenting with the construction of a real-time Cartesian to Polar coordinate transformation system within Java to afford continuous orientation and arbitrary viewing angles of tabletop documents. DiamondSpin tool kit supports 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 personal and shared regions within the same display space. DiamondSpin also provides a set of tabletop UI functions, including document visualization and movement methods based on Fisheye techniques, Context Rooted Popup menus, multiple menu bars, and popup soft keyboards.

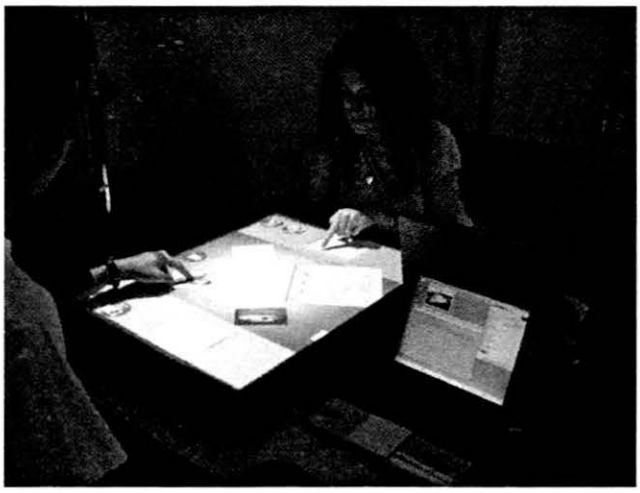
Collaboration: DiamondSpin is distributed to universities with a free license agreement. Current DiamondSpin license holders span from North America to Europe and Japan. We are developing applications using DiamondSpin at MERL and with our university partners.

Future Direction: Seamless interaction amongst multiple inputs from single point interaction and stylus/pen, to bimanual and stylus-hand cooperative gestural input interaction. DiamondSpin based applications, including public area and meeting room tables, photo tables, teacher's tables, as well as educational games.

Contact: Chia Shen, Clifton Forlines
http://www.merl.com/projects/diamondspin/

Lab: MERL Research Lab
Project Type: Research

UbiTable



MERL's UbiTable allows one to walk up to an interactive table, connect laptops, cameras, and other USB devices to the table; so that people can fluidly share, manipulate, exchange, and mark up their contents with each other on a large tabletop surface. At the same time, each user can still maintain explicit control over the accessibility and interactability of his/her own documents displayed on the tabletop.

Background and Objectives: Tables are a natural focal point for both formal and informal gatherings where people work on solutions to collaborative problems, sketch out design ideas, browse personal or web contents, or simply discuss issues of the day. It is common for participants to bring their own documents, sketches, or other data in digital form on mobile personal devices; however, the display and input capabilities of personal devices are primarily designed for use by only one individual at a time. UbiTable enables its users to opportunistically annex the input and display affordance of a large tabletop surface.

Technical Discussion: The key research challenges include (1) interaction techniques for fluid movement and transfer of documents amongst the tabletop and different devices in a multi-device, multi-user, environment, (2) a tabletop interface that allows simultaneous operations by multiple people on the same documents, while a user still maintains control over what operations on his/her own documents are allowed. The UbiTable tabletop functions as a shared workspace. External devices (e.g., laptops, digital cameras, etc.) function as private workspaces for each user, and also as data repositories. The UbiTable shared tabletop can be divided into two types of regions, Personal and Shared, such that the access control and interactibility of documents positioned in each region are distinctly different. Documents that are visibly displayed on the tabletop provide feedback information for both the ownership and current user. UbiTable is implemented using DiamondSpin, a Java Toolkit for developing multi-user tabletop applications. The UbiTable interface is projected onto a DiamondTouch touch input surface that can distinguish which user is touching the surface. Laptops use a wireless link to communicate with UbiTable, while other personal mobile devices can be connected via USB.

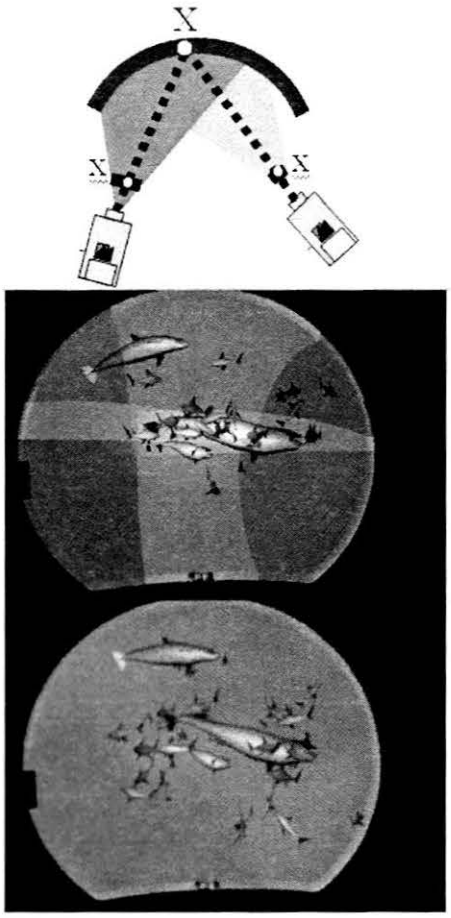
Collaboration: UbiTable is a collaborative research project between MERL Research and MERL Technology Labs.

Future Direction: UbiTable is an ongoing research project. We are expanding the UbiTable concept into collaborative interactions amongst multiple connected interactive surfaces and devices, including electronic whiteboards, in addition to tabletops and USB devices. We are also taking a participatory design approach to develop application concepts on the UbiTable.

Contact: Chia Shen, Clifton Forlines, Kathy Ryall
http://www.merl.com/projects/UbiTable/

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

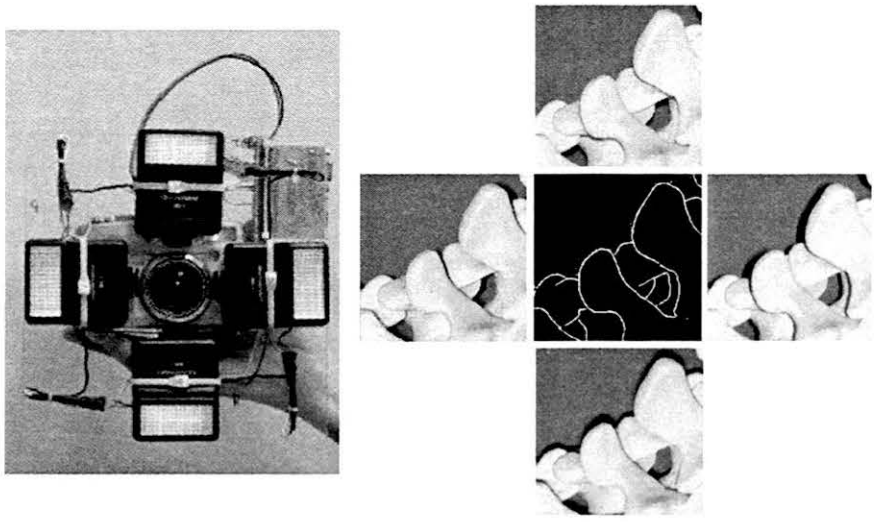
Collaboration: MPC, Johosoken.

Future Direction: Color management, head-tracking, stereo images, calibration support for larger dome screens.

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

Lab: MERL Research Lab
Project Type: Research

Multi-Flash Camera for Shape Boundary Detection



We describe a new type of multi-flash camera that can capture and convey shape features of real-world scenes. We compute the depth discontinuities in a scene with a shadow based technique. We take a very different approach to capturing geometric features of a scene than traditional approaches that require reconstructing a 3D model. This results in a method that is both surprisingly simple and computationally efficient. The entire hardware/software setup can conceivably be packaged into a self-contained device no larger than existing digital cameras.

Background and Objectives: Depth discontinuities corresponding to silhouettes and shape boundaries play an important role in scene understanding in Computer Vision. Unfortunately, majority of the techniques based on stereo cameras or range finders fail at depth discontinuities due to partial occlusion from one of the views. Time of flight based range scanners cannot detect small depth differences. We use an active illumination approach and exploit shadows as a robust cue for depth discontinuity.

Technical Discussion: We use a camera with multiple flashes that are strategically positioned to cast shadows along depth discontinuities in the scene. The projective-geometric relationship of the camera-flash setup is then exploited to detect depth discontinuities and distinguish them from intensity edges due to material discontinuities. We can highlight the detected features, suppress unnecessary details or combine features from multiple images. The resulting images more clearly convey the 3D structure of the imaged scenes. We can also use the depth edges to improve stereo reconstruction, detection change in scenes (for surveillance) or object recognition.

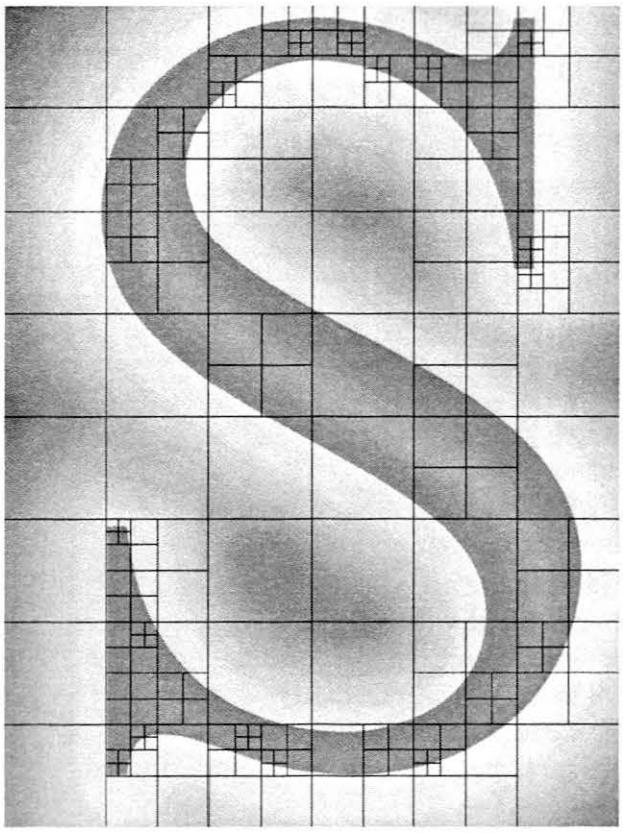
Collaboration: Rogerio Feris and Matthew Turk at UCSB and Jingyi Yu at MIT.

Future Direction: Exploiting depth discontinuities in various computer vision applications.

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

Lab: MERL Research Lab
Project Type: Research

Saffron: High Quality Scalable Type for Digital Displays



The Saffron Type System (Saffron) is a breakthrough approach to rendering high quality type on digital displays. Built on a core of patented Adaptively Sampled Distance Field (ADF) technology, Saffron achieves its superior results without the liabilities of current approaches. Saffron is a key enabling technology for the display of rich textual content on the next generation wireless devices and flat panel displays.

Background and Objectives: Saffron offers the following advantages over existing type systems:

-- Highly legible type even at very small font sizes without the use of labor intensive manual hinting.

-- Unparalleled adaptability for flat panel display technologies including new materials such as OLED.

-- Unique Continuous Stroke Modulation (CSM) feature provides interactive user tuning of type for enhanced viewing comfort and personal preference.

-- Backward compatible with the thousands of outline fonts already available in OpenType, Type 1 and TrueType format.

-- Computationally clean rendering pipeline straightforward to implement in silicon.

-- Supports advanced applications such as pen-based input, 3D type, animation and special effects.

-- Patently distinct from coverage/image based rendering approaches.

Technical Discussion: Saffron provides an alternative font rasterizer that can be integrated in the OS, at the application level, or in embedded systems. It takes font outline descriptions as input, converts them to an internal ADF representation and renders them in real time. Because Saffron rendering is computationally simple and does not use TrueType or Type 1 hinting required by competing technologies, fonts do not need to be special cased; Saffron can be implemented in both custom hardware or accelerated using standard graphics processing units (GPUs).

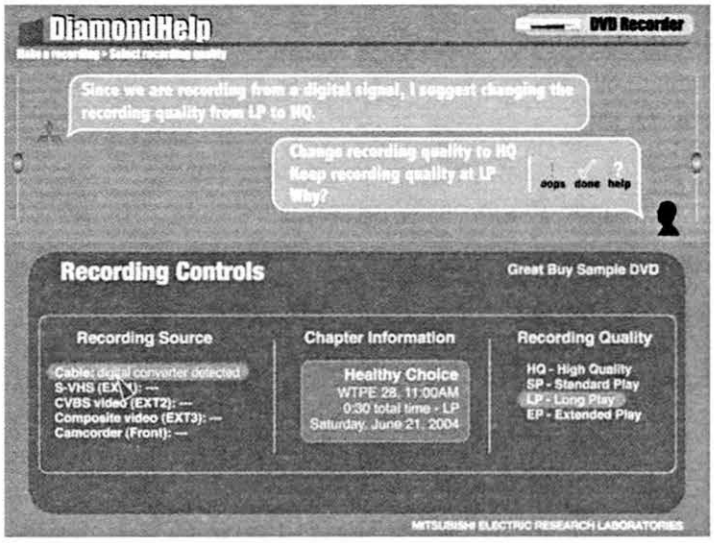
Collaboration: Jun Someya, Manager Image Processing LSI Group, Advanced Technology R&D Center.

Future Direction: Continued collaboration with Jun Someya to develop a Saffron-based ASIC for MELCO products. Continued licensing efforts to deploy Saffron in the marketplace.

Contact: Ron Perry, Sarah Frisken
http://www.merl.com/projects/ADF-Saffron/

Lab: MERL Research Lab
Project Type: Research

DiamondHelp: Collaborative Help for Networked Home Products



The basic idea of DiamondHelp is to move the complex programming and customization features of new digitally-enhanced home products for heating, cooling, laundry, entertainment, etc., to a shared household DiamondHelp "station", which is connected to the appliances through the existing home electrical wiring.

Background and Objectives: Ordinary people already have great difficulty using the advanced features of digitally-enhanced household products, and the problem is getting worse as more features are continually being added. This usability problem cannot be solved using only the tiny displays and limited control buttons typically found on home appliances. However, using a home network to share a larger and more powerful display, we can provide home products with a new type of collaborative interface in which the product actively helps the user, especially with complex features that are only occasionally used.

Technical Discussion: The shared household DiamondHelp station could be either a dedicated unit, or a home PC, wireless tablet, etc. Other technology, such as wireless, infrared, etc., could be used in place of or in combination with power line control to implement the home network.

DiamondHelp provides a consistent interaction style across products through a unique combination of the conversational and direct manipulation interface paradigms. The top half of the screen is like a chat window between the user and DiamondHelp, which is the same for all products. The bottom half of the screen is a direct-manipulation interface to the product's state, which is different for each product.

DiamondHelp is a *mixed initiative* system: it can provide detailed step-by-step instructions and demonstrations when appropriate, but also allows the user to do things by himself when he wants to. DiamondHelp also makes use of *task context*: it knows the user's goal at every point, either because the user explicitly stated it or via automatic goal recognition.

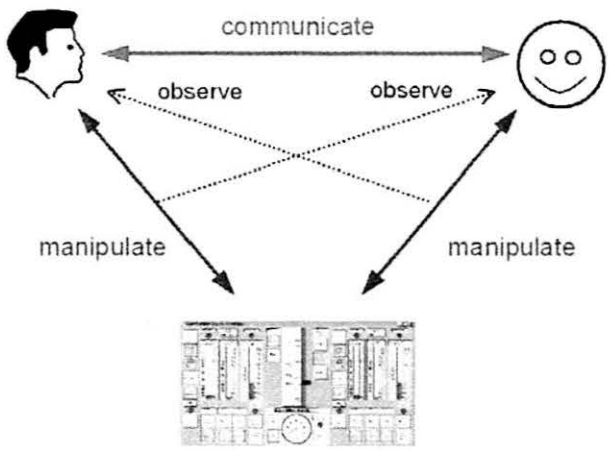
DiamondHelp uses Collagen (see project description) to track task context and to manage the conversational part of the interface.

Future Direction: We are exploring a number of variations and extensions to the basic DiamondHelp design, such as using speech recognition and generation to enhance the interface. We are also investigating the issue of standards, which will make it possible for products from different manufacturers to operate with the same shared DiamondHelp station.

Contact: Charles Rich, Neal Lesh, Candace Sidner
http://www.merl.com/projects/diamondhelp/

Lab: MERL Research Lab
Project Type: Advanced Development

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: As shown in the diagram above, 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.

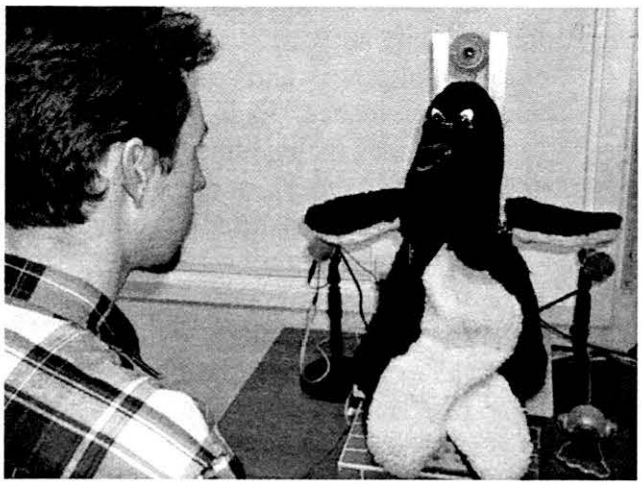
Collaboration: COLLAGEN is currently being used to build prototype systems for a range of applications, including networked home products (see DiamondHelp project description), multimodal mixed-initiative form filling (see FormsTalk project description) and a hosting robot (see Human-Robot Interaction 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

Human-Robot Interaction for Hosting Activities



We are investigating the process of engagement in human-robot interaction. We have studied human-human engagement, developed rules to apply to human-robot interaction, built a functional robot that interacts in collaborative conversations with people, and studied how people interact with our robot.

Background and Objectives: Our objective is to develop collaborative robots that not only hold conversations with people, but use physical gestures and movement to interact with them, thereby "engaging" the human in the interaction. We have focused on hosting settings, where the robot can provide information about the shared environment. Typical hosting settings include museums, stores and homes. In the home, the robot uses its knowledge of the home and its contents to direct people to items they need or tasks they need help performing. In museums and stores, it directs people to locations of interest and can demonstrate objects of interest.

Technical Discussion: Our robot collaborates with people to perform the hosting task of demonstrating the iGlassware system developed at MERL. The robot can seek out a person to interact with for the demonstration. People interacting with the robot do not need any training to interact. In recent work, the robot also recognizes and responds to their head nods, in addition to the previously developed interpretation and production of head gestures to indicate interest in the robot and environment. Over 100 people have completed demonstrations with the robot.

The robot architecture includes components for sensor fusion of information gathered from vision algorithms via a stereoptic camera, and sound from dual microphones, speech recognition, robot control algorithms and collaborative conversation using Collagen (TM). The architecture separates the robot's conversation and sensor fusion/robot control, with sharing of critical information between the two sub-systems.

Studies of people interacting with the robot show that they find the robot's gestures more natural than an unmoving conversational partner, and that they direct their attention more to the robot than the unmoving conversational partner.

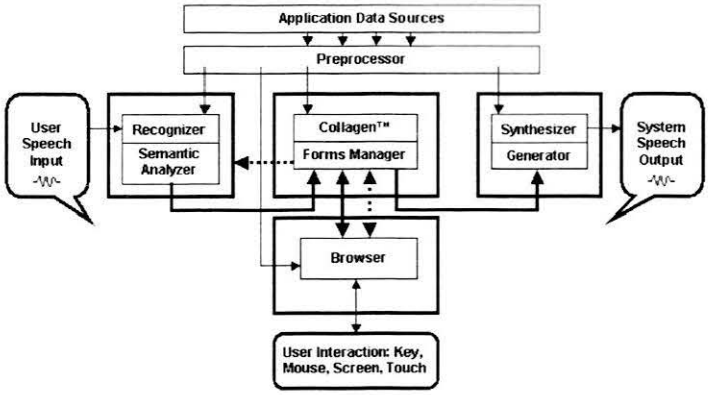
Collaboration: We have collaborated with members of the MIT Vision Group (http://www.ai.mit.edu/projects/vip/projects.htm) to use the Watson vision system and HMM algorithms to interpret head nods of people interacting with the robot.

Future Direction: We will conduct user studies of head nodding during conversation. We will mobilize the robot so that it can approach people to engage in conversation, and use its body, as well as head movements and conversation, to indicate its engagement in an interaction.

Contact: Candace Sidner
http://www.merl.com/projects/hosting/

Lab: MERL Research Lab
Project Type: Research

FormsTalk: Multimodal Mixed-Initiative Form Filling



FormsTalk is middleware for building form-filling applications that support a mixture of speech and non-speech interaction modes. Examples of non-speech modes include touch screen, keyboard and mouse, and telephone keypad. FormsTalk also supports flexible, mixed-initiative interaction, in which either the user or the system can take the lead, depending on circumstances. Our current application is to provide remote access to industrial plant control data via cell phone.

Background and Objectives: Form-filling is a common framework for many different kinds of applications. A multimodal interface 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 content of the forms, and deployment on different platforms (e.g., PCs, PDAs, phones, kiosks) requires a minimum of additional effort.

Technical Discussion: FormsTalk is built on top of DiamondTalk, 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 description) as its dialogue manager, which provides its mixed-initiative capabilities. 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, which is embodied in the FormsTalk middleware. Our first target application was public kiosks, such as for purchasing admission and movie tickets at a science museum, and used standard PC web browser technology. Our current cell phone application uses special purpose data display protocols.

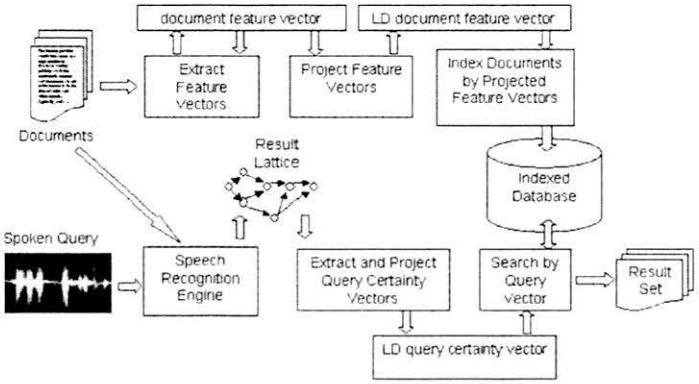
Collaboration: FormsTalk is the user-interface portion of the "Broadband Web Services" project led by the System Technology Dept. of SentanSoken. We are also members of the W3C Multimodal Interaction and W3C Voice Browser working groups, which are developing relevant standards.

Future Direction: As well as looking for other applications of FormsTalk, we also plan to explore adding SpokenQuery capabilities (see project description) to the architecture.

Contact: Bent Schmidt-Nielsen
http://www.merl.com/projects/FormsTalk/

Lab: MERL Technology Lab
Project Type: Advanced Development

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 and Objectives: 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. SpokenQuery solves several problems associated with searching for information using a standard spoken user interface. With SpokenQuery, there is no grammar to memorize, no tedious menu tree to navigate, and the system is more robust to misrecognition and out-of-vocabulary speech. Usability tests show SpokenQuery offers significant advantages over menus for these tasks.

Technical Discussion: SpokenQuery is a post-processor for a speech recognition engine. It is not tied to any particular speech recognizer, and has already been ported to several different engines. SpokenQuery is also not tied to any language. It treats the target information as a simple "bag of words", and does not make use of any language specific semantics or natural language processing.

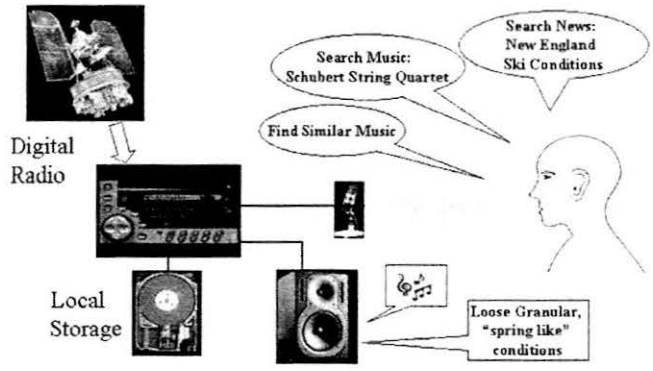
Collaboration: AdvisoryAgent (Sentansoken), FormsTalk (Sentansoken), KnowledgeProvider (Johosoken), AdaptiveAgent (IDken), MediaFinder (MERL).

Future Direction: Improve accuracy performance, reduce memory and processor footprint, and 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 tens of thousands of songs. Unfortunately, no current product offers any acceptable way to navigate such a large collection. Most current products provide just a tiny display of up to about 10 songs, and "next" and "previous" buttons. The problem is particularly bad in the automobile where it is vital the driver keep his/her eyes on the road and hands on the steering wheel. Finally, 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.

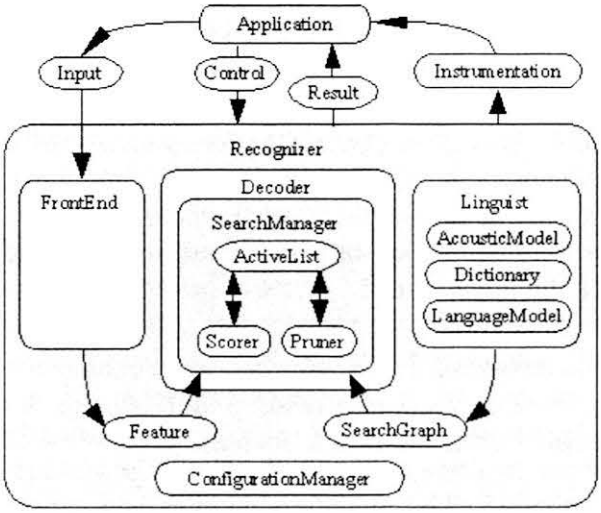
Collaboration: AdvisoryAgent (Sentansoken), AdaptiveAgent (Idken), MediaFinder MERL), and FormsTalk (Sentansoken).

Future Direction: Produce prototype products and services for MELCO Business Units.

Contact: Peter Wolf, Joseph Woelfel
http://www.merl.com/projects/MediaFinder/

Lab: MERL Technology Lab
Project Type: Advanced Development

Sphinx4 Development



Speech recognizers are getting increasingly commoditized - it is easy to purchase a speech recognizer for any project. Yet, commercial recognizers are becoming increasingly opaque, providing little exposure of their inner framework, and little control to the user. This makes it extremely difficult for researchers and developers to utilize them for innovative product designs or cutting-edge techniques.

The latest release of Sphinx4 can be downloaded from cmusphinx.sourceforge.net.

Background and Objectives: Speech recognition is a successful and growing area. However, for the average developer, speech recognizers themselves remain mysterious bots that must be purchased in binary form from vendors, and whose limited APIs are the only control provided to the developer. They also cost a lot of money. Finally, speech recognizers are highly constrained - the typical recognizer is good at one job, but incapable of others. E.g., it may be able to recognize CFGs, but not n-gram LMs, or vice versa.

Technical Discussion: In the Sphinx4 project we have developed a highly modular open source state-of-art speech recognition system. The system is open source, and hence free. It carries a BSD license, i.e. there are no encumbrances. The system is designed to allow complete access to all modules, to the knowledgeable developer. The APIs are extensive. Also, it has several modules that enable recognition using CFGs, FSAs, N-gram LMs and other linguistic structures, all within the same system. The system also includes several new features, such as the unusual bushderby search system and a token stack decoder. It can also perform recognition with multiple feature streams (i.e. multi-modal speech recognition), with tying at various levels, e.g. phoneme, word or utterance.

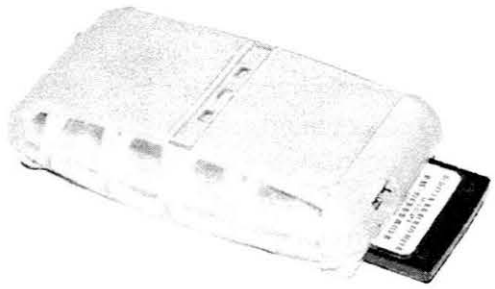
Collaboration: Rita Singh (CMU), William Walker, Paul Lamere, Philip Kwok (SUN).

Future Direction: Several new features are being added to the system, in response to requests from users. Also, we hope to build a trainer that can train acoustic models that fully utilize the capabilities of the decoder.

Contact: Bhiksha Raj
http://www.merl.com/projects/sphinx4/

Lab: MERL Research Lab
Project Type: Advanced Development

ComBadge



The ComBadge is a two-way voice messaging device with a simple spoken user interface. This 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: This project set out to develop a communication device that would be controlled solely via speech and would be less expensive to produce and to operate than a cellular phone. This would create new sales opportunities for such devices.

Technical Discussion: ComBadge 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 compressed and, therefore, 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.

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 closely with the speech applications group in MERL MTL.

Future Direction: We are extending ComBadge to perform group messaging by using a spoken distribution list name to deliver the same message to multiple users. Delivering voice messages to machines in addition to people allows automated services to be developed. Initially, we are looking at using ComBadge to control household devices and as a voice portal to traffic, weather, appointments, and stock quotes. An important new research area being investigated is using ComBadge in a mesh-connected environment in which no infrastructure is required. In such a system, a voice message is delivered to a destination ComBadge by forwarding the message through other intermediary ComBadges.

Contact: Jamie Frankel
http://www.merl.com/projects/ComBadge/

Lab: MERL Research Lab
Project Type: Research

Sensor and Data Systems

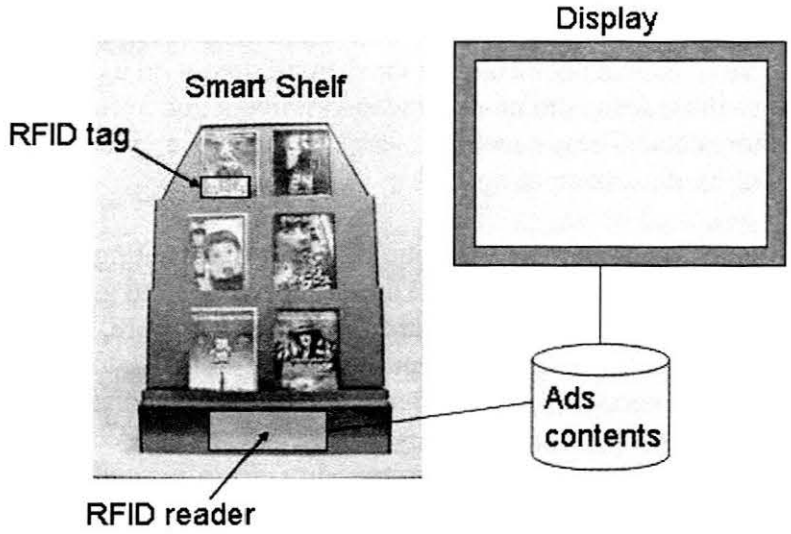
Until recently, computer applications have largely presented the user with human-generated content that has been processed according to fixed or human-generated algorithms. Now it is becoming possible for a computer system to collect data from the local environment on its own, process that data adaptively and then use those results to modify the environment and to further refine the system's ability to analyze more data. These capabilities are driving new applications in sensor networking, data mining and ubiquitous computing.

MERL's work in "Sensor and Data Systems" is creating new technologies for this exciting area, ranging from fundamental ideas to support the technology itself to applications meant to grow new businesses and enhance current ones. Projects span the gamut from sensor hardware, which can automatically and cheaply collect data, to new systems for communicating that data, to algorithms for finding features in the data and correlating it all. Because cost of infrastructure is a limiting consideration for most applications of these technologies, MERL's researchers have kept this in mind so that the new applications will be able to use commodity electronic and computing hardware, and be cheaply deployed.

Our work this year includes important advances in sensitive, inexpensive ways to monitor the environment. These include chemical and physical sensors, which can process their sensed data and supply it in digital form, and yet are cheap enough to be disposable. Powerful new capabilities in sound recognition and analysis allow standoff monitoring of systems including machines and road traffic. RFID technology is now coming of age, and MERL is developing technologies that will take advantage of it in the usual fields such as supply chain management, as well as in new areas such as location-aware services. New advances in dimensionality reduction and intelligent, multi-dimensional data summarization allow data to be analyzed, automatically and succinctly, so that a computer system can take advantage of only the relevant information it needs. We have also been working on ways for humans and computers to cooperate on various optimization processes. These techniques are useful in information-cluttered applications such as data mining and the design of complex systems such as antennas. MERL has also created highly efficient algorithms for elevator scheduling. These have been designed to take advantage of outside data so that a building sensor network can contribute to the efficiency of the entire elevator system.

Project Descriptions

RFID Application



We have been developing an RFID-based "Smart Shelf" that detects the presence or absence of each item on the shelf. When an item is removed from the shelf for examination, relevant advertising is displayed.

Background and Objectives: The RFID market is rapidly expanding in applications such as supply chain management, logistics, and other retail systems. Adoption of global standards for UHF RFID is expected to increase interest in these systems. Although the technology promises to give us the ability to identify the location of items without line of sight, the reality is that the adoption of RFID is problematic. In addition to technical issues like RF interference, it is unclear how to create profitable systems which directly benefit consumers.

To better understand and overcome these problems, we are integrating RFID into retail application concept systems. This integrated system will serve as a realistic test bed for UHF RFID, which is not yet licensed for use in Japan.

Technical Discussion: The smart shelf system uses RFID tags which are attached to the items and RFID readers which are embedded underneath the shelf. The readers periodically interrogate the tags so that the system can manage real-time shelf inventory. Furthermore, the system can detect the removal of an item when a customer picks it up. The "pick-up" event changes the content of a computer-driven display to show advertising and information related to the item. Thus, the system provides benefits beyond simple inventory management by attempting to directly influence sales. In order to detect the presence of the tags on the shelf accurately, we measured actual rate of successful reads on each shelf, as well as basic read performance in a free field with a single tag or multiple tags.

Collaboration: Data Management Technology Department of Johosoken, MDIS, MDIT

Future Direction: We are pursuing pilot installations and business possibilities with our collaborators. There are also ideas for integrating other MERL technologies into the application concept system.

Contact: Mamoru Kato, Paul Dietz
http://www.merl.com/projects/rfidapplication/

Lab: MERL Technology Lab
Project Type: Research

RFIG: Photosensing RFID for Location Aware Services

We have developed wireless tags that can be located using coded illumination. The idea can be easily adopted for passive RFID allowing a range of geometric operations. We augment each tag with a photo-sensor to significantly extend the current functionality and support radio frequency identity and geometry (RFIG) discovery. The ability to address and wirelessly access distributed photosensors creates a unique opportunity. We recover geometric information, such as 3D location of tags or shape history of tagged objects, and exploit the associated geometric operations to bring the RF tags into the realm of computer vision and computer graphics.

Background and Objectives: Wireless tags such as RFID (radio frequency identification) are becoming ubiquitous in inventory management. However, it is difficult to precisely locate the tags using a handheld device. We overcome this problem by using a hybrid of radio frequency and optical communication with the tag.

Technical Discussion: The handheld device consists of a RF reader plus a data projector. The handheld device is first aimed casually in the direction of a tagged object. The handheld device sends an RF signal to synchronize the tags, followed by illumination with a sequence of binary patterns such that each projector pixel emits a unique temporal Gray-code. The tag records the Gray-code that is incident on its photo-sensor, and then makes an RF transmission of its identity plus the recorded Gray-code back to the RF-reader. The computed (x,y) projector pixel coordinate is turned on to visually indicate the location of the tag. Multiple tags can be attached to a single rigid object to compute the relative pose (3D location and orientation) of the object and to compare geometry of multiple objects.

Collaboration: Dr. Shiotani, Sentansoken.

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

Lab: MERL Research Lab
Project Type: Research

Sound-Based Traffic Incident Detection

Accidents in traffic are notoriously hard to detect using automated means. This is especially the case when using computer vision where the lighting, weather and accident setup conditions can vary dramatically and hinder proper detection. However the sounds emitted by various anomalies in traffic are usually fairly consistent and can be used to detect accidents fairly easily. In this project we analyzed traffic sounds and came up with consistent and accurate recognition results for anomalous events such as accidents.

Background and Objectives: Sound-based traffic incident detection is a project that employs our sound recognition framework to detect accidents, near-accidents and traffic anomalies in monitored city intersections. Our objectives were to develop a lightweight and cost effective real-time monitoring system that can accurately detect traffic anomalies automatically. Because of the nature of sound this is an easier task than video detection and the hardware requirements are much less demanding, suggesting a cheap solution.

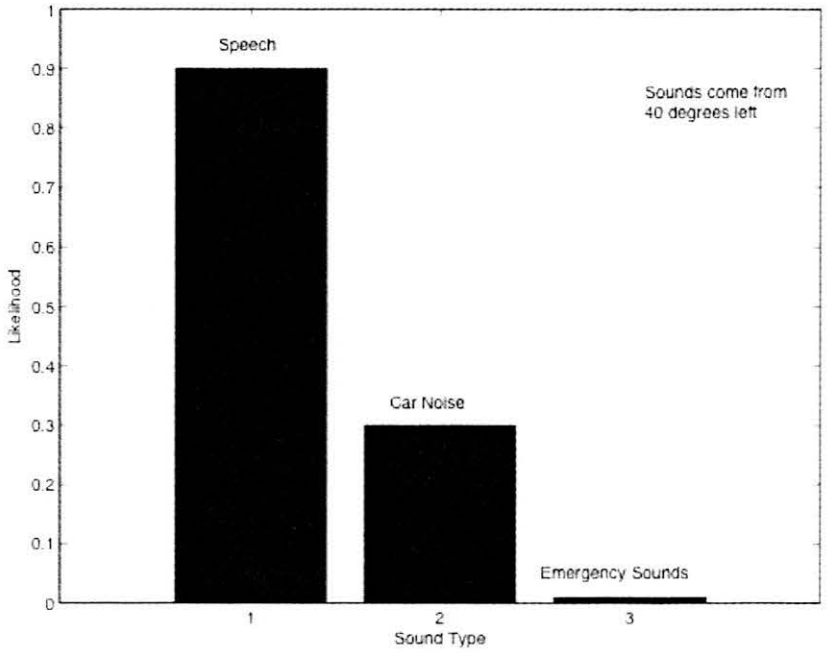
Technical Discussion: We have employed three different approaches to this detection problem. The first one is a supervised learning methodology, where classifiers are trained to traffic sounds such as car crashes, brakes, and tire squeals. These classifiers are then applied to the incoming signal and raise the appropriate event flags according to the detected sounds. The second methodology is unsupervised learning that attempts to detect anomalous events in the sound stream. Due to the rarity of car accidents they are an anomalous sound class which is automatically isolated from the recordings by this system. Finally we developed a user interface which allows a human operator to quickly browse through hours of recordings and easily find sections which need more attention. All of these systems provide accurate results and have successfully detected accidents from recording provided to us by the TRIMARC intelligent transportation project in KY, USA. The training data we used was provided by Mitsubishi Electric Engineering in Nagasaki and are data gathered from the Japan National Police.

Collaboration: Mitsubishi Electric Engineering, Intelligent Transport System Business Group of MELCO.

Contact: Paris Smaragdis, Ajay Divakaran
http://www.merl.com/projects/traffic/

Lab: MERL Research Lab
Project Type: Research

Sound Recognition



Sound recognition is a project with the goal to enable machines to listen and understand their surrounding auditory environment. We anticipate this technology to 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.

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. Our objective is to provide lightweight and cost-effective monitoring capabilities using sound.

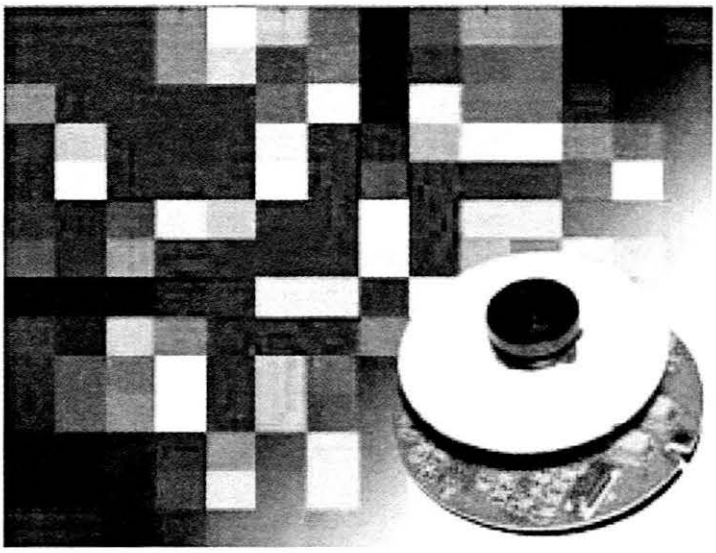
Technical Discussion: The sound recognition project is an umbrella for various computational frameworks. We have developed a wide variety of methodologies that can perform many tasks such as recognizing known classes of sounds, isolating unknown sounds, detect anomalous behavior, localize sounds in space and perform scene analysis and segmentation of a sonic environment. Our algorithms are in the forefront of machine learning and signal processing and provide a fairly wide margin from competition. Parts of our technology have also been included in the MPEG-7 standard for indexing and extraction of audio content.

Collaboration: Iwamoto-san, Hashimoto-san, Sentansoken.

Contact: Paris Smaragdis, Ajay Divakaran, Bhiksha Raj
http://www.merl.com/projects/SoundRecognition/

Lab: MERL Research Lab
Project Type: Research

Pedestrian Flow in Buildings



This project focused 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 are 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.

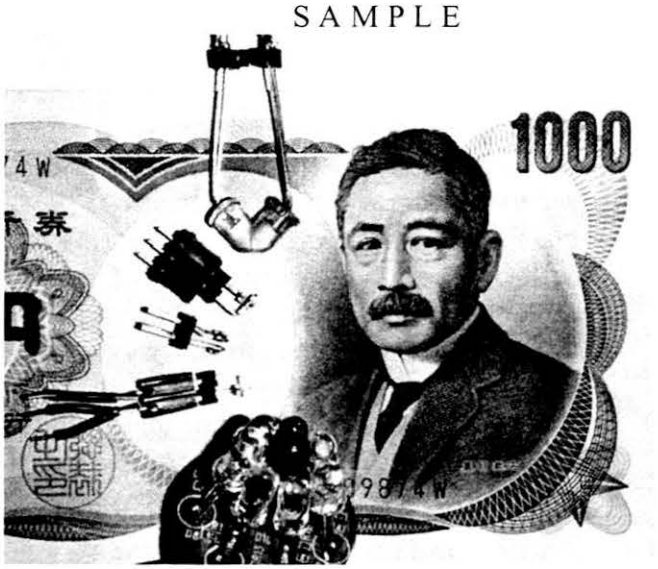
Technical Discussion: This pilot study proved that it is possible to accomplish sensor localization, behavior classification, and pattern discovery from ultra-low resolution sensor networks without the need to track people, or invade their privacy. The primary lesson from the pilot study is that the "1 bit per square meter per second" regime is useful for human activity recognition and is ripe for continued research work. An interesting secondary lesson is that for behavior understanding at the building scale, context is extremely important: in a careful study our behavior recognition rate improved with area covered, even for constant information bandwidth, all the way up to the edge of the 175 square meter test area.

Collaboration: MTL

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

Lab: MERL Research Lab
Project Type: Initial Investigation

LED-Based Sensors



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 and directly producing a PWM signal (no power-hungry A/D converter needed). The system holds great promise for following color changes in chemochromic sensing materials, and surface-state and contamination-detection on a wide range of surfaces.

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 (~ 5 parts per billion). In addition to the sensing capabilities of this system, it holds the added bonus of offering short-range wireless communication of data.

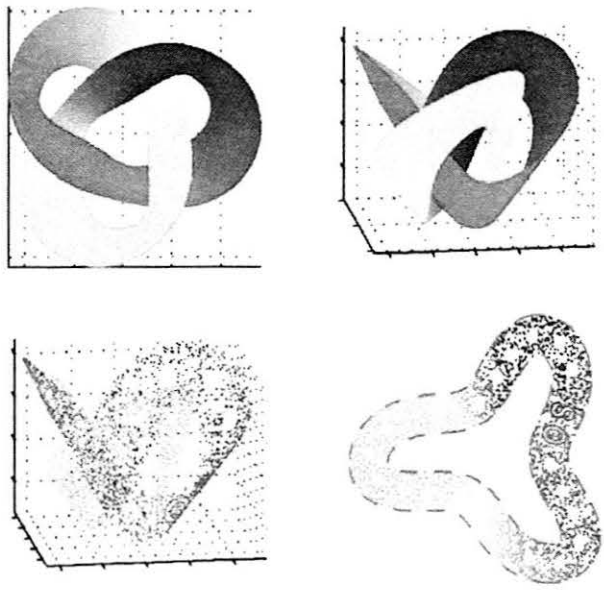
Collaboration: This is a joint effort of MERL, Sentansoken, Tokyo Institute of Technology in Tokyo, and Dublin City University, Ireland.

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

Dimensionality Reduction



See Color Figure 12

We are developing manifold models of high-dimensional signals that capture much more information than classic dimensionality reduction methods, yielding superior performance in compression, classification and regression tasks. This technology applies to a vast array of multimedia applications, and allows us to manipulate and edit data in new ways. The image shows 3D faces synthesized from a 15-dimensional manifold. The underlined faces are real; the rest are interpolated and extrapolated.

Background and Objectives: Linear dimensionality reduction is used extensively in signal processing, statistics, machine learning, machine perception, and data mining. It is a core component of technologies as diverse as face recognition, web searching, visual target tracking, audio source separation, and image compression. Yet linear models are fairly poor data approximations, because most kinds of data have a substantial nonlinear component. Our goal is to devise nonlinear dimensionality reducers that give us control over the linear and nonlinear components of the data.

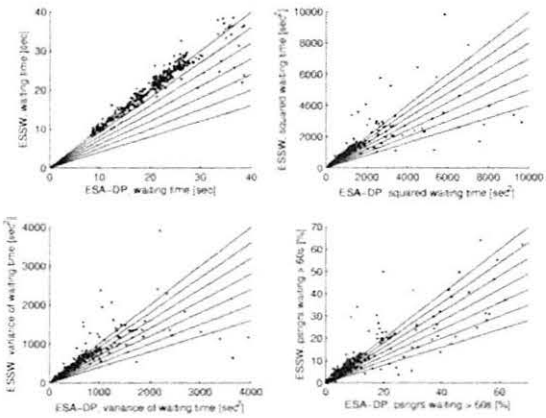
Technical Discussion: We illustrate with an example of faces. Although facial data typically has millions of degrees of freedom (e.g., face images have 10^6 pixels), faces probably have a relatively small number of degrees of freedom. The manifold of faces, which collects all possible faces, is a thus low-dimensional surface in image space. We seek a function that relates the manifold's intrinsic coordinate system to face images. Because the manifold is not known, it must be estimated from sample data. We have developed estimators of the manifold dimensionality, coordinate system, and mapping functions that give much more accurate data compression and reconstruction than linear models, while using fewer dimensions. On the theoretical side, we have shown that there are classes of extrinsically curved manifolds that can be recovered from finite samples with perfect (not approximate) isometry.

Future Direction: We are working on making these methods "industrial strength" so that they can replace principal components analysis in MELCO's signal processing and data mining products.

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

Lab: MERL Research Lab
Project Type: Research

Group Elevator Scheduling



Elevator passengers expect fast and efficient elevator service, and it is the job of the elevator scheduler to dispatch cars so as to balance the simultaneous needs of all passengers currently present. While passengers are prepared to wait for a reasonable amount of time, their patience disappears quickly if waits become excessive, e.g. more than a minute. Consequently, rather than minimizing only the expected waiting time of passengers, their aversion to excessive waits must also be encoded in the objective function of the scheduler. We have developed an efficient computational algorithm that can minimize an arbitrary objective function expressing any desired attitude to waiting times.

Background and Objectives: Group elevator control is a hard industrial problem characterized by huge state spaces and significant uncertainty. In past work, we have shown that the uncertainty in the destinations of existing passengers can be marginalized efficiently by means of a fast dynamic programming algorithm, thus reducing significantly the expected waiting times of passengers. However, elevator passengers typically express an aversion to risk (excessive waits), which can be encoded as an additional control objective by means of a supra-linear function of waiting time. Computing the expectation of a non-linear function is not a simple problem, given the substantial uncertainty in the system.

Technical Discussion: Since any given continuous utility function of waiting time can be approximated arbitrarily well by Chebyshev polynomials of a suitable degree, the computation of expected utility can be reduced to the computation of the first non-centered (raw) statistical moments (expected powers) of waiting time. We developed an algorithm for efficient computation of all such statistical moments that is only linear in the desired degree of the approximating polynomial. When the resulting algorithm was used to minimize squared waiting times, thus penalizing excessive waits, the fraction of passengers waiting more than 60 seconds for service was reduced significantly, thus improving the quality of service.

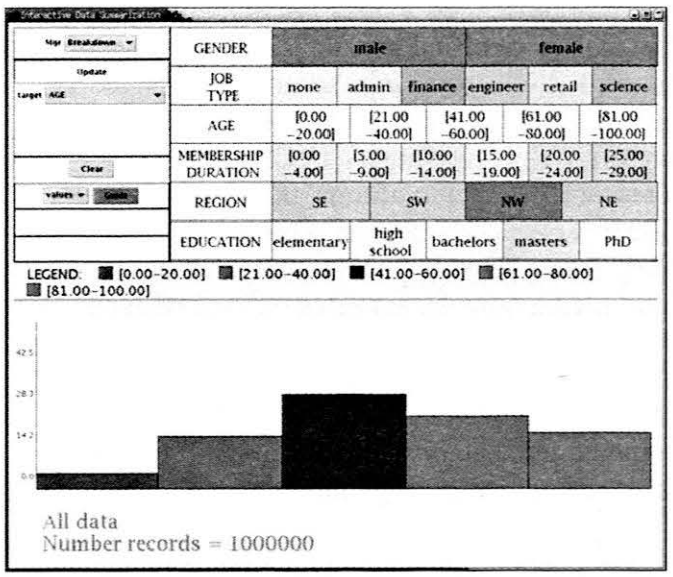
Collaboration: Sentan Soken.

Future Direction: When the scheduling system is not obliged to make an assignment immediately after a new passenger has arrived and requested service, but is allowed to delay and reconsider decisions until just before the passenger is to be picked up, the added flexibility and reduced uncertainty usually result in significant savings in waiting time. However, no scheduler has ever been able to fully benefit from this flexibility, because of the combinatorial explosion resulting from the need to constantly reconsider all existing assignments. Still, we hope that the structure of the problem might allow for an efficient branch-and-bound algorithm that can successfully replace the heuristic solutions currently employed in industry.

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

Lab: MERL Research Lab
Project Type: Research

Intelligent Multi-Dimensional Data Summarization



Summarizing or exploring large multidimensional datasets often requires extensive investigation by a user to identify overall trends and important exceptions to them. While many visualization tools help a user produce a single aggregation of the data at a time, they require the user to explore the dataset manually. Our idea is to have the computer analyze all aggregations exhaustively and inform the user about where further investigation is warranted. Our prototype tool allows the user to quickly view different types of aggregations of different subsets of the data. The key novel feature is to guide users to aggregations that are the most different than the current aggregation, in order to help them more quickly find patterns and trends that are interesting to them.

Background and Objectives: Several business units in MELCO have expressed interest in this project from a demonstration of the initial prototype. We are in close collaboration with the Data Management Technology Department at Johosoken to jointly refine the current tool based on feedback from the business units and to support a trial examination of it with a customer of Koden.

Technical Discussion: While summarization is highly subjective and context-dependent, the fundamental insight we leverage is that when a person investigates a series of visualizations of related aggregations, what she learns is based on what changes or does not change between the visualizations; the difference between visualizations is something we can compute. Our techniques provide the user with visual cues as to where they will find change. Summarization is closely related to compression, machine learning, and data mining. The closest connection is to data mining. We have presented a formal framework that encompasses both our summarization task and the conventional data mining task called association mining. Our prototype allows users to interactively construct summaries of multidimensional data that are composed of a variety of graphs such as barcharts, correlations, and simple averages.

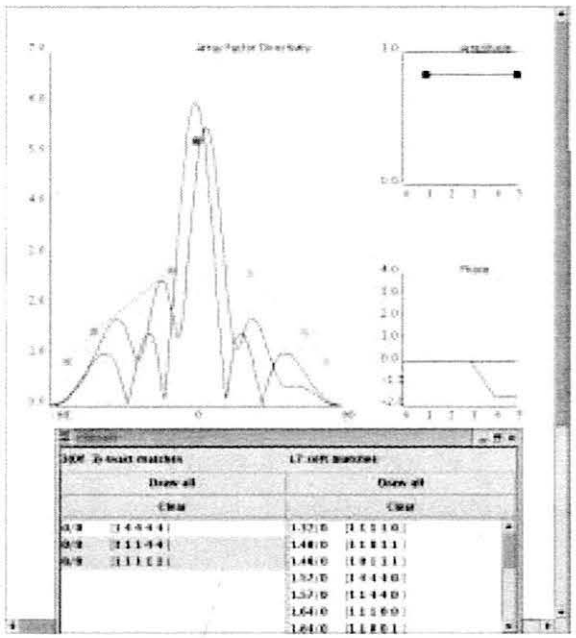
Collaboration: This project is joint work with the Data Management Technology Department at Johosoken with input from a consultant from Harvard University.

Future Direction: Our future efforts include connecting the prototype to DIAPRISM, generally improving the usability of the current features of the prototype, as well as adding new elements to the prototype such as a fish-eye view to handle data sets with large numbers of dimensions and fast median computation.

Contact: Kathy Ryall, Neal Lesh, Daniel Nikovski
http://www.merl.com/projects/imds/

Lab: MERL Technology Lab
Project Type: Initial Investigation

Human-Guided Antenna Design



Optimization-based approaches to antenna design have enjoyed limited success. The task is often computationally intractable and it is 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 is often impossible to consider all relevant design criteria in the optimization process. In this project we propose that the computer be used differently, leaving the task of choosing a final design from the computer-generated sampling to the human user, who can apply experience and judgment to recognize and then refine the most useful antenna design. Thus the "generation" of the candidate set, and "visualizing" the set are separated into two tasks.

Technical Discussion: We have designed and developed QueryLines to provide an interactive query system that employs "approximate query and visual browsing" that can be used to search and analyze two-dimensional data-sets. Antenna performance data (e.g., radiation patterns) and other data sets involving linear ordered sequences (e.g., finance, weather, census data) are especially well-suited for visual query techniques; the graphical nature of the input and output methods support users in their information-seeking tasks, in this case designing an antenna with particular performance characteristics. QueryLines provide both constraints and preferences, and query results are categorized into matches and soft matches, which aid users in refining their queries and understanding the data. We have integrated QueryLines into a larger system in which multiple graphs may be displayed in a unified viewing window (e.g., one graph for gain, a second for amplitude, and a third for phase). A user may query in any or all of the graphs.

Collaboration: This project is a joint effort between MERL Technology and Research Laboratories in collaboration with Johosoken and with sponsorship from Denshihon.

Future Direction: In the upcoming year we will be focusing our efforts on visualizing and interactively browsing large datasets of phased-antenna arrays with non-uniform spacing, and exploring the utility of our QueryLine technique for other domains.

Contact: Kathy Ryall
http://www.merl.com/projects/antenna/

Lab: MERL Technology Lab
Project Type: Advanced Development

Color Figures

Figure 1 – see page 50

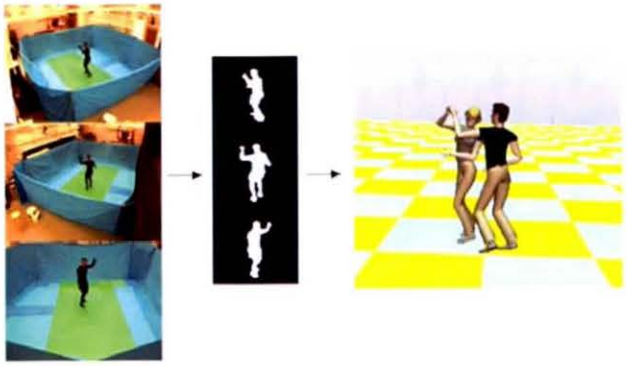


Figure 2 – see page 51

Figure 3 – see page 52

Figure 4 – see page 54

Figure 5 – see page 55

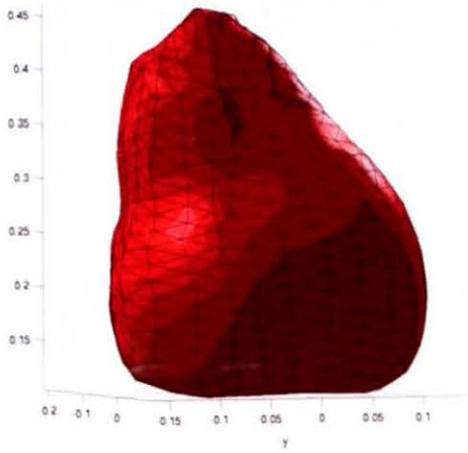


Figure 6 – see page 62

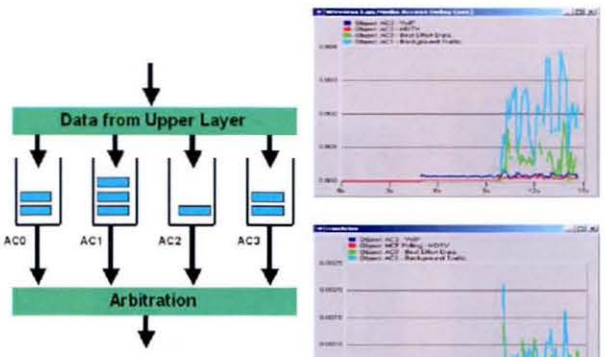


Figure 7 – see page 75

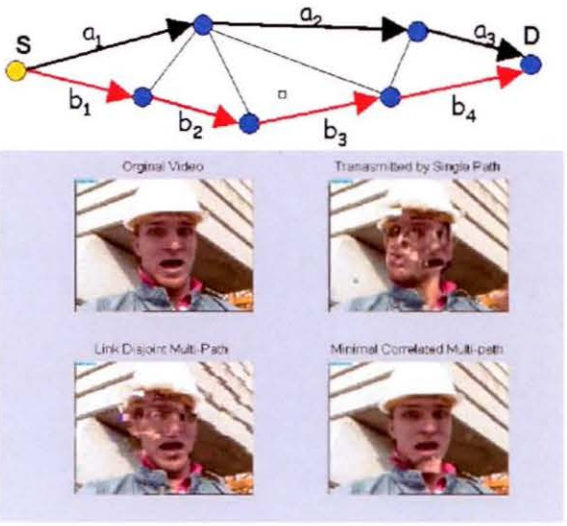


Figure 8 – see page 85

Figure 9 – see page 86

Figure 10 – see page 87

Figure 11 – see page 98

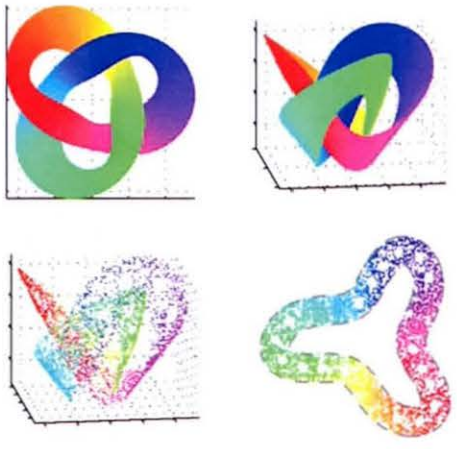


Figure 12 – see page 118