Mitsubishi Electric Research Laboratories (MERL)

Annual Report

July 2005 through June 2006

TR2006-00

(Published July 2006)

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 subsidiary of the corporate research and development organization of Mitsubishi Electric Corporation (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 & wireless transmission technology & networking.
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 data analysis.

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, Stanford, Georgia Tech, Harvard, Columbia, Tufts, the University of Toronto, Imperial College London, ETH Zurich and Dublin City University. 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 realize the full market potential 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 facility in Cambridge, Massachusetts and collaborate closely to achieve groundbreaking results. The seven 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
  Dr. Masatoshi Kameyama (EVP, CFO & CLO)
- Research Lab - *Basic & Applied Research, staff 22*
  Dr. Joe Marks (VP & Director)
  Dr. Joeseph Katz (VP & Deputy Director)
- Technology Lab - *Applied Research & Advanced Development, staff 42*
  Dr. Kent Wittenburg (VP & Director)
  Dr. Huifang Sun (VP, Deputy Director & Research Fellow)
- 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.

Masatoshi Kameyama *Ph.D., Tokyo Institute of Technology, 2005*
Executive Vice President, Chief Financial Officer & Chief Liaison Officer

Masatoshi Kameyama joined MELCO in 1975. He was a Visiting Research Scientist at the MIT Media Lab from 1985 to 1986. At MELCO's Information Technology R&D Center (Johosoken) he developed display systems and graphics accelerators for workstations, PCs and mobile products. He also led a project that created very large "display walls" using multiple rear projection modules. Before coming to MERL in 2004, he was the general manager of Johosoken's Multimedia Laboratory.

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

Prior to Joining MERL in 1994, Joe Marks worked at Bolt Beranek & Newman and Digital Equipment Corporation's Cambridge Research Laboratory. As a researcher at MERL, Joe's primary focus was on computer graphics, user interfaces, and heuristic optimization. Joe became Associate Director of MERL's Research Lab in 1999 and Director in 2000.

Joeseph Katz *Ph.D., California Institute of Technology, 1981*
Vice President, Deputy Director Research Lab

After working at Caltech's Jet Propulsion Laboratory for a number of years, Joseph Katz went to Symbol Technologies, where as Senior VP of R&D he participated in, initiated, and led projects in a wide range of technologies, including barcode/RFID data capture, optics, imaging, signal processing, computing, networking, security, biometrics, and communications. He joined MERL's management in 2004.

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 focused on Human-Computer Interaction (HCI) technologies. 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, Deputy Director Technology Lab & Research Fellow

After four years as a Professor at Fairleigh Dickinson University, Huifang Sun moved to the Sarnoff Research Laboratory in 1990 becoming Technology Leader for Digital Video Communication. In 1995, Huifang joined MERL as the leader of our video efforts, becoming a Deputy Lab Director in 1997. In recognition of his 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 and 7 years at Active Control eXperts Inc. beginning as Director of Sales and Marketing and ending as Vice President, Core and New Business Unit. Adam came to MERL in June of 2000 to lead our Marketing and Business Development effort.

Achieving Effective Business Impact

As a research lab, MERL's defining activity is research. However, from the perspective of our parent company MELCO, this is useless if it does not eventually benefit MELCO. Therefore, MERL must work with MELCO to make sure that this happens.

To eventually impact MELCO with our research, MERL must cross two divides: one between research ideas and product ready technology, and the other between MERL and the MELCO organization seeking to use MERL's technology.

It is next to impossible to cross both of these divides at once. While there are significant differences in work approaches, researchers who are rubbing shoulders with advanced developers in the same building can collaborate effectively with them. Similarly, while there are significant problems due to distance and differences in organizational allegiance, advanced developers in one organization can collaborate effectively with advanced developers in a distant organization. However, it is very difficult for effective collaboration to occur between a researcher and an advanced developer in a distant organization. The same issues apply when advanced developers collaborate with developers.

If MERL only worked on research, then it would be hobbled in it's attempts to provide benefit to MELCO, because it could collaborate effectively only with the relatively few people at MELCO who are also doing research. Because MERL also works on advanced development and development, it can collaborate effectively with everyone at MELCO. When a technology is born at MERL, MERL can (and must) carry it far enough toward product to effectively work with any group in MELCO.

Working with MELCO is a matter of two-way communication. It is not just a question of ideas originating at MERL and ending up in MELCO. Information and ideas that come from MELCO by means of collaborative work are equally important. Ideas and needs from MELCO make important contributions to individual projects. In addition, they are a vital source of information that MERL can use when deciding which projects to pursue.

Balancing the Time Horizons of Projects

It is vital for any research lab to have projects spanning a wide range of time frames. MERL's goal is to spend 1/3 of its effort on short-range projects, for which MERL is within 1 year of final delivery to MELCO. In addition, MERL seeks to spend 1/3 of its effort on middle range projects that are 2 to 3 years from delivering to MELCO. Finally, we believe that it is essential to spend 1/3 of MERL's efforts on long-range projects that are 4 or more years from completion.

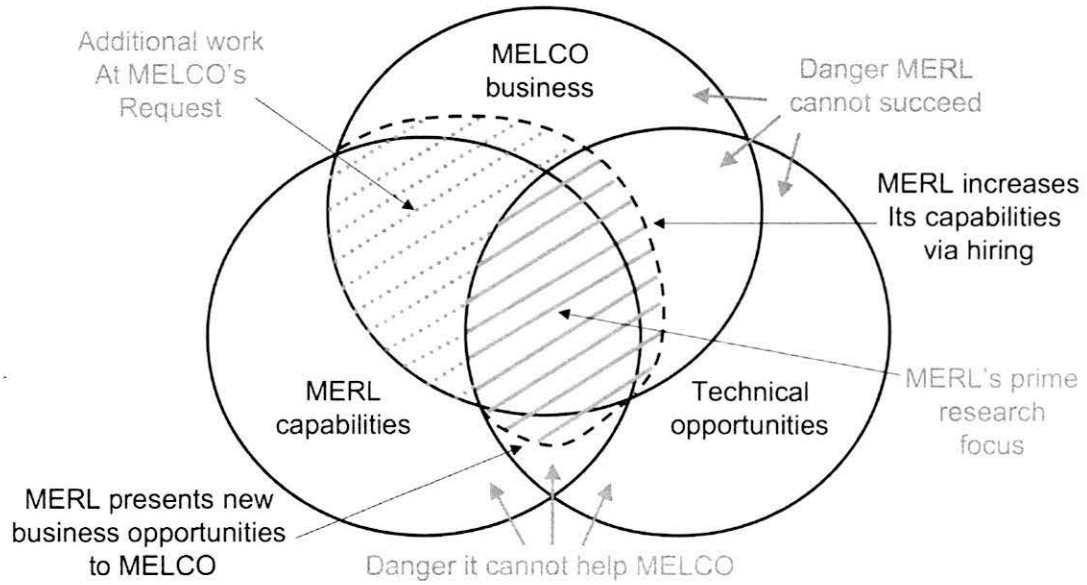
None of these time frames are more important than the others. Rather, all are essential. In a given year, if a lab does not have projects that are delivering to the parent company in that year, the lab is not fulfilling its mission. If it does not also have maturing projects that will soon be delivering to the parent, then it soon will not be fulfilling its mission. Lastly, if it does not have long-range projects that are developing the key ideas that will support contributions in the future, then it eventually will not be fulfilling its mission. MERL continually monitors its activities to make sure they do not get over-weighted to any one time frame.

Project Evaluation

The most important decision at any research lab is which projects to pursue. It is only to be expected that many research projects that appear promising will nevertheless fail. However, it is essential that a research lab not waste resources on projects that do not appear promising.

At MERL, three criteria are used to evaluate whether a project is promising: (1) Is the project closely enough related to one of MELCO's businesses that it can become a real business success for MELCO if it works? (2) Will work on the project provide opportunities for realistically achievable technical advances that will give MELCO competitive advantages? (3) Does MERL have appropriate people and sufficient resources to complete the project successfully?

The Venn diagram below illustrates the kind of projects that MERL seeks to support most and those that it avoids. Our prime goal is to seek out projects that satisfy all three goals, because we believe that that is where MERL can make the greatest contribution to MELCO. At the same time, we provide assistance to MELCO when asked in any area where we are capable of doing so. As projects mature, they often change from ones focused on technical opportunities to ones focused on supporting the productization of the technology within MELCO.



In contrast, we are careful to avoid projects beyond MERL's capabilities even if they are strongly related to MELCO's business, because there is no advantage to MELCO to have MERL pursue projects that will not succeed. However, by means of consultants, visiting researches and new staff hires, we continually extend MERL's capabilities so that we can usefully pursue as wide a range of projects relevant to MELCO business as possible.

We are equally careful to avoid the temptation of working on projects that are not strongly related to MELCO's business even if they represent great technical opportunities. However, when we believe that a technical opportunity can be converted into a big business opportunity, we do not hesitate to present this opportunity to MELCO.

MERL has a great deal of freedom to decide what it believes is best to do. With this comes an equal amount of responsibility to ensure that what we do strongly benefits MELCO's business.

Mitsubishi Electric

Number 156 on Fortune magazine's most recent list of the world's 500 largest corporations, Mitsubishi Electric Corporation (MELCO) has $32 billion in annual sales, $662 million in profits and nearly 100,000 employees in 35 countries.

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 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
Public Utility Systems Government Systems, Transportation	(Kobe, Itami)	Shakaihon
Communication Systems Wired & Wireless 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
Energy & Industrial Systems Power Equipment, Plant Control	(Kobe, Nagasaki)	Denryokuhon
Semiconductors Optical and Radio Frequency Semiconductors	(Kita Itami)	Hanpon

Together, these ten 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 Trading Company the Mitsubishi-Tokyo Financial Group, Mitsubishi Heavy Industries, Mitsubishi Chemical and Mitsubishi Motors, (all five of which are also on the Fortune Global 500 list—Numbers 149, 217, 221, 288, & 304 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 in order of revenue.

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 70 on IEEE Spectrum's most recent list of the top R&D spenders, 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. K. Kyuma (Director), Mr. A. Morita (GM), 18 people (Tokyo) Managing MELCO's R&D	Hatsuhon
Advanced Technology R&D Center (ATC) Research & Advanced Development: Dr. S. Yamamoto (GM), 916 people (Itami) Materials, Semiconductor Devices, Electrical & Mechanical Engineering	Sentansoken
Information Technology R&D Center (ITC) Advanced Development: Mr. R. Nishii (GM), 838 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	IDken
Mitsubishi Electric Research Laboratories (MERL) Research & Advanced Development: Dr. R. Waters (CEO), 82 people (MA) Computer Vision, Communications, Video coding, HCI, Sensor networks	MERL
Mitsubishi Electric Information Technology Centre Europe (ITE) Advanced Development: Mr. K. Oshima (CEO), 56 people (France & England) Wireless Communications, Digital Audio & Video	ITE

Awards and Commendations

The high caliber of MERL's research and researchers 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, 2005 through June 30, 2006.

Current Technical Society Fellows and Senior Members

Five (8%) of MERL's researchers are fellows of professional societies:

Dr. Joseph Katz, Fellow Institute of Electrical and Electronic Engineers
Dr. Joseph Katz, Fellow Optical Society of America
Dr. Andreas F. Molisch, 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
Dr. Huifang Sun, Fellow Institute of Electrical and Electronic Engineers

A further Sixteen (26%) of MERL's technical staff are Senior Members of the Institute of Electrical and Electronic Engineers (IEEE).

Best Paper Awards

Molisch, A.F., "Ultrawideband propagation channels - theory, measurement and models", *IEEE Transactions on Vehicular Technology*, pp 1528-1545 (2005). (Neal Shepherd Memorial Best Propagation Paper Award, which recognizes the best paper on propagation published in Transactions on Vehicular Technology during the year.)

Morency, L.-P., Sidner, C., Lee, C., and Darrell, T. "Contextual recognition of head gestures", *Proceedings of the Seventh International Conference on Multimodal Interfaces (ICMI'05),* October 2005. (Outstanding Paper Award.)

Tse, E., Greenberg, S., Shen, C., and Forlines, C. "Multimodal Multiplayer Tabletop Gaming", in *Proceedings of the Third International Workshop on Pervasive Gaming Applications - (PerGames)* (Dublin, Ireland) May 2006. (Best paper award.)

Wren, R. and Tapia, E.M., "Towards Scalable Activity Recognition for Sensor Networks", in *Proceedings of the 2nd International Workshop on Location- and Context-Awareness (LoCA 2006)* (Dublin, Ireland) May 2006. (Best presentation award.)

Technology Awards from Outside Organizations

Japan's prestigious audio-visual magazine HiVI rated MELCO's 'Raku-Reko' DVD recorder as the best analog DVR of 2005 and singled out MERL's Highlight Playback feature (see the business impacts section of this report) as a key reason for the selection.

The Dome Multi-Projection Imaging System from Mitsubishi Precision, which is based on MERL's multi-projector mosaicing technology (see page 21 of MERL's 7/04-6/05 Annual Report) was named one of the ten major products of 2005 by the Japanese Nikkan newspaper.

Awards from MELCO

In April 2006, MERL staff received a MELCO Excellent Patent Award for their patent "Interactive Wireless Tag Location and Identification System".

In April 2006, MERL staff received a MELCO Excellent Patent Award for their patent "Multimedia Event Detection and Summarization".

In April 2006, MERL staff received a MELCO Excellent Patent Award for their patent "Method and System for Displaying Images on Curved Surface".

In April 2006, MERL staff in collaboration with staff from MELCO's Sentansoken research laboratory received a MELCO Excellent Patent Award for their patent "Summarized and Play Back Equipment and Method".

In May 2006, MERL staff in collaboration with staff from MELCO's Sentansoken research laboratory received a Sentansoken Laboratory R&D award for their work on "Realization of Highlight Playback Functionality for DVD Recorder".

In May 2006, MERL staff in collaboration with staff from MELCO's Sentansoken laboratory received a MELCO Corporate R&D award for their work on "Face Recognition Technology and Applications".

In May 2006, MERL staff in collaboration with staff from MELCO's Johosoken research laboratory and Shahon received a MELCO Corporate R&D award for their work on "Development of Voice Interface for Car Navigation".

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 first system using it was delivered to the customer as the date of impact.

A different way that MERL can impact MELCO is by making standard contributions. This may or may not lead to direct revenue via licensing. However, 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.

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.

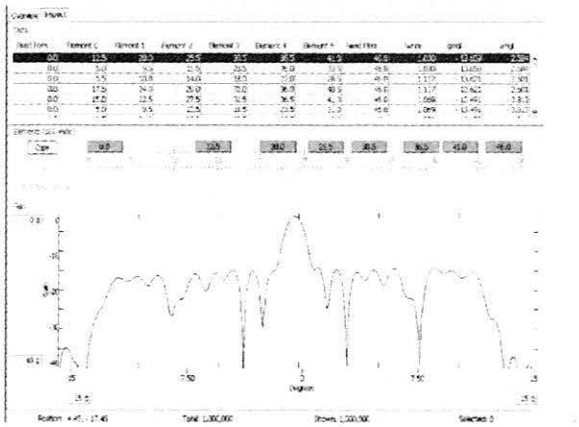
The following subsections detail what MERL's impact on MELCO has been in the 12 months spanned by this report. In addition, they summarize how this impact was achieved. It is worthy of note that there are several distinct models of how impact can be achieved ranging from work specifically requested by MELCO to finding an application in MELCO for a technology developed independently by MERL. In addition, the typical time from the inception of a project at MERL until actual impact on MELCO is 3 years, with some projects taking twice as long.

Linear Phased Array Antenna Design Tool

In the summer of 2005, MELCO's Denshihon business unit started using MERL's linear phased array antenna design tool. A linear antenna array is composed of several antennas arranged in a line. The same signal is transmitted/received on each antenna; however, delays are introduced so that the signal on one antenna is out of phase with the signals on other antennas. The characteristics of the array as a whole depend on the exact positions of the antennas and the phase shifts between them in very complex ways. MERL's tool simplifies the task of exploring the space of design options and designing an array that optimizes the specific properties that are needed for a particular antenna application.

MERL's work on antenna design tools is part of a wider project we call Human-Guided Search (HuGS). This work combines human insight with computer-based optimization to create results that are better than can be created by either alone. We are working on extended our work to more complex antenna arrays and on a tool that generates near-optimal results in much less time.

Details: MERL's work on antenna design tools has two quite distinct parts: work on interfaces for human guided searching in design spaces and work on antenna simulation software.



The work on human guided searching began at MRL in 1995 with an approach called Design Galleries, which allows a person to rapidly explore a complex design space. This combines a back end that can simulate a wide range of alternatives and can categorize them, but does not perform any optimization, with a front end that makes it possible for a user to rapidly search through a large space of alternatives.

In 1997 the name HuGS was coined and the focus of continuing work switched to combining back ends that are capable of partial optimization with front-ends that allow a user to interact in deep but intuitive ways with the optimization. This approach is important for tasks that are so complex that full computer optimization is not possible. It is also important in situations where the optimization criteria are so dependent on human intuition that they cannot be fully stated in computer-understandable form.

The work on HuGS was always a small project, but it has been pursued over many years, because discussions with people at MELCO have revealed a series of potential applications of HuGS to MELCO business. In late 2001, a conversation with engineers at Denshihon revealed that antenna design required particularly complex optimization and might be appropriate for HuGS.

The following year, MRL and MTL researchers working with a summer student intern created an initial prototype antenna design tool. This tool was based on a standard antenna simulation program as its back end and supported the interactive optimization of a particular kind of antennas called Yagi-Uda antennas. Software generated random configurations of such antennas in a manner intended to produce a broad variance in several performance parameters. An antenna designer could then constrain the measured parameters to areas of specific interest, and inspect a range of appropriate antenna designs.

Engineers at Denshihon and the Antenna Department at MELCO's Johosoken laboratory were intrigued by the prototype antenna design tool and asked MERL to implement a more complete system. MTL researchers took prime responsibility for implementing an improved prototype during 2003.

As collaboration between MERL and the Johosoken Antenna Department grew stronger, they proposed a follow-on project to find configurations of low side-lobe level, linear phased-array antennas with quantized phase shifts and antenna positions. This was an important change in the project because it moved the project from experimenting with suggestive prototypes to working on an actual problem relevant to MELCO business.

In the continuous (non-quantized) case, it is a simple task to find a phased-array configuration, which will produce a given response. However, if the parameters of the phased array are constrained, by allowing only limited possible phase shifts and/or limited possible antenna

positions, then the problem becomes vastly more difficult. Existing optimization techniques, such as simulated annealing or genetic algorithms, can handle this type of problem, but they provide the designer with little or no insight into how changes in parameters affect the final result and therefore little or no information that designers can use to make engineering tradeoffs in antenna design.

In 2004, MERL embarked on this new project with MTL taking the lead on developing an appropriate HuGS front end and MRL taking the lead on developing an appropriate antenna simulator.

The key property that makes this problem hard (quantization) opens up a brute force approach to antenna simulation. Given sufficient computational power (in this case MERL's multi-PC course-grained parallel supercomputer) one can enumerate every possible antenna array and compute their characteristics.

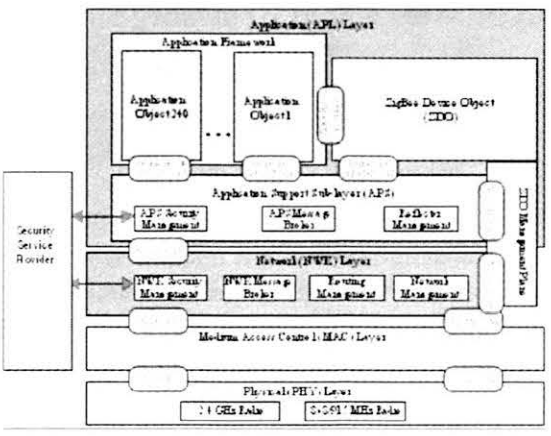
If a designer's constraints on the final array response could be concisely specified in mathematical form, one could then merely select whichever antenna is best. However, most interesting problems involve hard-to-specify trade-offs between various parameters. As a solution to this problem, MERL developed a browser that allows an antenna designer to visualize the space of possible antenna designs and perform graphical queries.

An interesting result of the project was the design of an antenna showing that even given the particular quantization constraints imposed, it is possible for a uniformly illuminated, partially filled linear phased array antenna to outperform the 13.3 dB theoretical maximum side-lobe suppression of a uniformly illuminated, completely filled linear phased array.

Because exhaustive searches were performed during the research, it was possible to gather statistics for the entire set of possible sub-array placement configurations. MERL has since used those statistics to create and verify algorithms that can generate array configurations that are only near optimal, but can be determined in far less time than can be done using exhaustive search.

Protocol Stack for ZigBee Chip

In the third quarter of 2005, Renesas began shipping engineering samples of their ZigBee interface chip M30624GPGP. ZigBee is a wireless communications standard intended for home automation and sensor network applications. The heart of the chip is an M16C processor core running software that implements the ZigBee protocol. MERL provided the network layer (NWK), application support sub-layer (APS), and ZigBee device object code (ZDO) for the protocol stack running on the chip.



In addition to working on the ZigBee standard and its implementation, MERL is working on applications of ZigBee in home automation and sensor networks. Beyond this, MERL is involved with a wide range of other wireless communication standards and applications.

Details: ZigBee is a wireless multi-hop ad-hoc networking protocol defined by the ZigBee Alliance. It supports Low cost, short range, low bit-rate, and low power consumption sensor

networks with features such as self-organization & self healing of networks, automatic device & service discovery, multi-hop cost-aware routing, reliable and secure communication, and application interoperability. The potential applications for this technology include home automation, building automation, industrial monitoring, automatic meter reading, surveillance, asset management, and environmental monitoring.

MELCO became a founding member of the ZigBee Alliance in 2002. At the request of MELCO, MTL researchers have been actively involved in ZigBee research and have gotten some of the technology they developed included in the ZigBee standard.

The basic wireless communication in ZigBee conforms to the IEEE 802.15.4 standard for short range, low bit-rate communication. The novel aspects of ZigBee are confined to the upper layers of the protocol.

The NWK layer provides the functionalities for assigning an address to each device, forming ad hoc networks, self organizing & self healing of the network, routing discovery & routing maintenance, data broadcast and multicast.

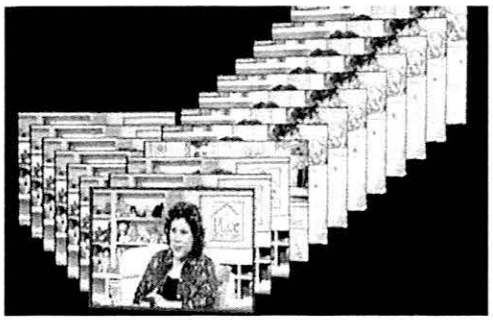
The APS is an interface between the network layer and applications. It provides the 'acknowledge' and 'retry' services for the multi-hop data delivery, indirect addressing via a binding table and security management.

The ZDO represents a base class of functionality that provides an interface between applications and the APS. It controls the network formation and provides device and service discovery, node management services, binding management and device configuration.

Based on requests from MELCO and Renesas, MTL started software implementation of the ZigBee NWK protocol in 2004. Following that, we were requested to develop APS and ZDO software to complete the ZigBee stack. All three pieces of software have been certified in Japan & the US by the ZigBee Alliance authorized testing company, TUV, and were delivered to Melco & Renesas in 2005.

TimeShuttle Video Browsing Interface for MDEA DTV

In September 2005, MELCO's US TV subsidiary MDEA began shipping a new video browsing technique in their high end Digital TVs that have Personal Video Recorder (PVR) capability. This technique, which MDEA calls TimeShuttle, was developed by MERL and provides a temporal-context enhanced view of a stored video that helps the viewer get to a desired spot in a video faster and more accurately than can be done using conventional fast-forward or rewind.



MELCO has a long history of video related products from video transmission equipment to TV sets; and MERL has a long history of working with MELCO in these areas. In addition, MERL has a long history of working on advanced human computer interfaces of various kinds. MTL researchers from these two groups worked together to develop TimeShuttle.

Details: Prior to joining MERL in 2001 an MTL researcher had done work on what is known as "rapid serial visual presentation": presenting a set of images such as the pictures in a photo

collection one after another. The typical way such images are shown is by flashing them one at a time all in the same spot. The researcher showed that other approaches that associate different positions with different images and provide some context for what images are coming next can be more effective.

After joining MERL, the researcher explored ways to apply his image presentation approach to TV interfaces. An initial goal was creating an improved interface for channel browsing. With the help of others at MTL, a prototype was constructed that shows the content on multiple channels somewhat like billboards beside a highway in a layout we called TimeTunnel. The images overlap so that the closest image is seen fully, taking up much of the screen but part of the left edge of each subsequent image is also seen. As you move from one image to the next, all the partial images move toward you. This provides context for what is coming next and makes it easier to get to the channel you want to reach.

This was demonstrated to multiple people in MELCO starting in early 2002 and generated interest, but no direct connection to products. A key problem was that TimeTunnel used 3D rendering to display the images being shown, which was fine for a demonstration, but unacceptably costly for inclusion in a consumer electronics device. Another problem was that while a high end TV has multiple channel tuners (to support picture in picture displays) it only has a small number of tuners and tuning to a new channel takes several seconds. As a result, it is not clear that it is practical to obtain the image data that TimeTunnel needs for channel browsing.

Work on TimeTunnel continued with a focus on improved efficiency by switching to 2D rendering only and using it as an interface to stored video where the needed images have already been captured. The interface was refined with collaboration between MTL and MRL researchers and changed so that it showed both past and future video frames in a V shaped pattern with the current frame the center of the screen at the point of the V. By means of user studies it was shown that viewers can find a particular location in a video more accurately using TimeTunnel than using conventional fast-forwarding and rewinding.

The revised TimeTunnel prototype was shown to MDEA in early 2004. MDEA decided to include it in their PVR-equipped TV sets to be sold in the fall of 2005. To make this possible, MERL had to rewrite TimeTunnel from the ground up paying very close attention to efficiency so that it could run on the embedded processor being used in MDEA's TVs without the need for any additional hardware. The main issue was that TimeTunnel has to access and display many stored images on the screen at once. This places strong demands on bandwidth to the TV's PVR disk drive and on the TV's embedded processor.

Through extensive use of built-in features in the TV's processor and a few compromises, an acceptable software design was devised during 2004. In a final push during January and February 2005, MTL implemented the TimeTunnel interface on the product platform and integrated it into the product prototype through an intensive on-site effort at MDEA in close collaboration with MDEA engineers. MERL's video browsing approach was first demonstrated by MDEA in April 2005 and was released into the market in fall 2005 under the trademark TimeShuttle. It is the first interface of its kind to be introduced into the consumer TV market.

Intelligent Highlight Playback

In September 2005, Lihon's Kyoto works began shipping a new DVD recorder (the model DVR-HE50W/10W). This recorder contains an interface concept provided by MERL as the primary advertised feature. For sports programming in particular, MERL's "intelligent fast forwarding" makes it easy to scan through recorded content, skipping from one key play to the next.

MELCO has a long history of video related products from video transmission equipment to TV sets. One of the labs that became part of MTL was founded in 1993 by the Audio/Visual business unit and a number of researchers originally from this lab have been involved with video ever since. These researchers work both on projects specifically requested by MELCO and on more speculative work.

Details: In 1998-99, MTL worked on video indexing in the context of the MPEG-7 standard, culminating in the acceptance by MPEG-7 of a "Motion Activity Descriptor" that can be very efficiently computed directly from compressed video.

Building on this foundation, MTL did speculative work on automatic video summarization starting in 2000 and continuing into 2002. The general goal of video summarization is to locate a small subset of a video that can serve as a summary of the rest.

MTL's initial work in this area took the traditional approach of breaking video into segments and selecting 'key frames' to represent each segment. However, over time, MTL's focus shifted to what is perhaps better described as intelligent fast forwarding. In particular, MTL created a prototype system featuring variable speed playback where 'interesting' segments are played back at normal speed while other segments are played back faster than real time. In this system, 'interesting' was defined primarily in terms of MTL's motion activity descriptor, with segments featuring highly varied motion being considered more interesting.

MTL demonstrated its summarization work to MELCO people on a number of occasions staring in mid-2001. One of the first people to see MTL's prototype was from the DVD recorder design group at MELCO's Sentansoken laboratory. He was interested in it as a way to interact with user-recorded content.

When purchasing a DVD, people are accustomed to receiving indexing information such as a table of contents that divides the recorded material into 'chapters'. However, when a DVD recorder writes content, this is typically stored without any indexing information at all. MTL's work held the promise of (at least partially) filling this gap.

Starting in the last half of 2001, the DVD recorder group began to push the process forward, providing funding for work at MERL and specific requests relating to the look and feel of MTL's prototype. In late 2002, the Sentansoken group presented MERL's demo to the home entertainment products business group (in LIhon) and started to get buy-in from them.

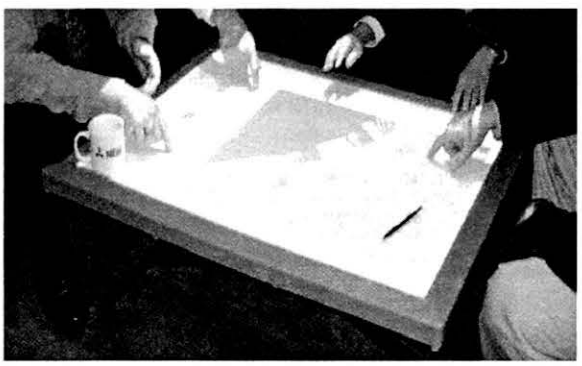
In parallel with the image-based work above, MTL began to experiment with using audio features for summarization. This was done using code developed by an MRL researcher as part of his work on an audio contribution to MPEG-7. This work got off to a slow start in 2001, but by 2002 had yielded tantalizing results. Driven by the needs of their specific application, the MTL researchers eventually re-designed the audio analysis algorithms with the help of a second MRL researcher. They went on to devise a classifier training technique that yielded the high classification accuracies that were required for practical applications. A key aspect of the algorithm is that it enables scalable summarization i.e. generation of summaries of any desired length.

Everything began to come together in 2003, with a strong push toward productization. Sentansoken sent one of its DVD researchers to MERL for 6 months to work closely with MERL. While studying exactly how MERL's work could be included in a DVD recorder, it was discovered that the planned DVD recorders did not have enough processing power to support video analysis---not even highly efficient analysis based on motion vectors. Fortunately, it was discovered that satisfactory results could be obtained by using very efficient-to-compute audio features alone.

Close collaboration with Sentansoken continued through 2004 and into 2005, leading to the joint creation of product-ready summarization code. This was included in the product released in late 2005, which was the first commercial personal video recorder with highlights playback capability. A key feature of the interface is that the user can display a plot showing the estimated level of 'importance' of each part of a stored video and select a 'slice level' that determine how 'important' a section of video needs to be in order to be shown during intelligent fast forwarding.

The product has received critical acclaim in the Japanese press, with the prestigious Japanese magazine HiVi rating it as the best buy in its category.

DiamondTouch Development Kit



In November 2005, MERL began selling a prototype development kit of its DiamondTouch hardware. DiamondTouch is a touch surface that can distinguish multiple simultaneous touches by multiple users. MERL's goal in selling DiamondTouch development kits is to obtain design wins from system integrators and develop a volume market for DiamondTouch. Once this has been done, future manufacture and sales of DiamondTouch will be turned over to some other part of MELCO.

Off-the-desktop computing and next generation interfaces are an important area of research at MERL. MERL is refining the underlying DiamondTouch hardware (exploring transparent and untethered versions) and developing several new, related technologies including light-pen-like approaches (to enable multi-user interaction with arbitrary projected displays). In addition, MERL is working on applications that use DiamondTouch such as DiamondSpace, which supports multiple users and multiple interaction surfaces.

Details: Various kinds of touch surfaces have been in use for decades, most prominently in the form of the touch screens seen in many special-purpose kiosk-like computer interfaces. These surfaces are effective, but share some key limitations. To start with, they typically can only accurately detect the position of a single touch. If they are touched in two places at once, this is reported as a single touch in a position that is a blend of the two positions. In addition, they have no way to tell who is touching them at any given moment.

The research that lead to DiamondTouch started in 2000 with the goal of creating a device that could be simultaneously touched by multiple people and could determine who is touching where so that a group of people could simultaneously interact with a computer system while gathered around a single interaction surface. This goal was posed by MRL researchers and solved by a close collaboration between MTL and MRL researchers. An initial prototype was created in only a few months, utilizing what has turned out to be a very robust approach.

Embedded within the DiamondTouch surface are two antenna arrays formed of narrow lines, one horizontal and one vertical. Extremely low-power electric signals are strobed across these arrays. The table also has a set of receivers, which are connected to the users. When a user touches the table, the user provides a capacitatively coupled pathway that allows the strobed signals to reach the user's receiver. By analyzing the timing of the received signals, the receiver can determine which antennas the user is touching. If a user touches in two places, the receiver is not capable of determining both positions completely, but much more information can be obtained than with standard touch surfaces. This allows useful interaction featuring multi-touch gestures. Users are unambiguously identified by means of the receivers they are connected to.

There are two limitations in the current version of DiamondTouch. The antenna arrays are not transparent and therefore cannot be placed between a user and a computer display. Wires are used to connect the users to the receivers. We are working on ways to eliminate these limitations, but the limitations are not as restrictive as one might think.

We typically orient DiamondTouch horizontally like a table and use front projection from above to show computer data on the table. The simple construction of DiamondTouch makes it easy to make the surface as large as desired, with an inherently high level of accuracy. The receivers are typically connected by wires to chairs set up around a DiamondTouch table. This allows the connections to the users to be made by simply having the users sit in the chairs. Users find interacting with their hands on the table useful and engaging.

While DiamondTouch is essentially a hardware invention, considerable software is needed for it to be useful. In particular, MERL wrote software drivers so that DiamondTouch can be used with the Microsoft Windows and Linux operating systems. In addition, MERL has created a Software Development Kit that provides support for the development of DiamondTouch applications in a variety of languages (C/C++, Java, ActiveX Control). Finally, MERL has developed a number of sample applications to illustrate DiamondTouch's potential. These software efforts spanned many years and consumed much more effort than the hardware design itself.

DiamondTouch was first demonstrated publicly in the Fall of 2001, and was demonstrated to MELCO in Japan in August 2002. It attracted immediate interest from MELCO, but it was not clear how to introduce it into products. Working together with the prototyping department of MELCO's Johosoken laboratory, demonstration tables were constructed with DiamondTouch

built in. Several MELCO business units, including Denshihon, began showing DiamondTouch to customers.

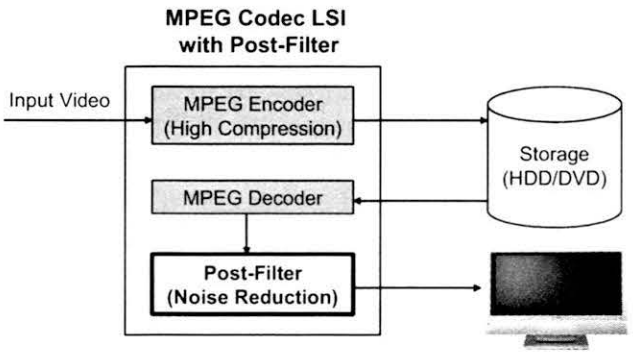
In 2003 Denshihon used DiamondTouch as a key part of a bid for a government contract. In collaboration with Johosoken, MERL delivered DiamondTouch for use in this system in December of 2004.

At the same time that opportunities were being pursued for DiamondTouch to be a subcomponent used within MELCO's traditional lines of business, MERL initiated an effort to determine whether DiamondTouch by itself could become a MELCO product. To generate interest, MERL has given prototype DiamondTouch surfaces to over 30 universities and other organizations so that they can experiment with them in their research.

As the final step in this commercialization effort, MERL worked in the first three quarters of 2005 to redesign DiamondTouch for improved manufacturability and then began prototype sales. Like all the other work on DiamondTouch, this effort was a close collaboration between MRL and MTL. We hope that this will lead to volume sales and that DiamondTouch will become a MELCO product in its own right.

Post-Filtering in DVD Recorder Chip

In the fourth quarter of 2005, Renesas began selling an MPEG-2 codec LSI that includes post-filtering technology developed at MERL. The codec is used in the DVD recorder of a major consumer electronics manufacturer. The chip employs high-compression technology developed by MELCO for efficient storage. Post-filtering techniques are then used during playback to suppress visual artifacts that are introduced during the compression thereby improving the quality of the video output.



Video compression and analysis is a traditional area of research at MTL. This includes work on transcoding between one compression standard and another and down-conversion of images from high to low resolution. We expect that in the future, MERL's post-filtering algorithm will also be used in the image processing chips for MELCO's digital TV products.

Details: MTL and MELCO's Johosoken laboratory have worked together on various issues relating to high-compression video coding for a number of years. In 2003, Johosoken asked MTL to investigate the problem of compression artifact reduction.

Advanced video coding methods discard some of the least important information in an image. This is necessary to achieve high levels of compression, but causes errors when the video images are reconstructed. Two key problems are blocking artifacts and ringing artifacts.

Blocking artifacts are due to the fact that each video image is broken into an array of blocks, which are compressed separately. Separate coarse quantization of the data in each block produces discontinuous intensity values at the block boundaries. This produces a visible blockiness in the reconstructed image. Adaptive low-pass filtering can deblock images quite

well with low complexity. However, it is difficult to determine when (and how strongly) these filters should be applied.

Ringing artifacts occur as the result of the quantization of high frequency components of images containing edges. In the absence of accurate high frequency components, ringing of intensity values (i.e., alternating light and dark bands) parallel to edges occurs. Most techniques that remove these artifacts have the undesirable side effect of blurring edges.

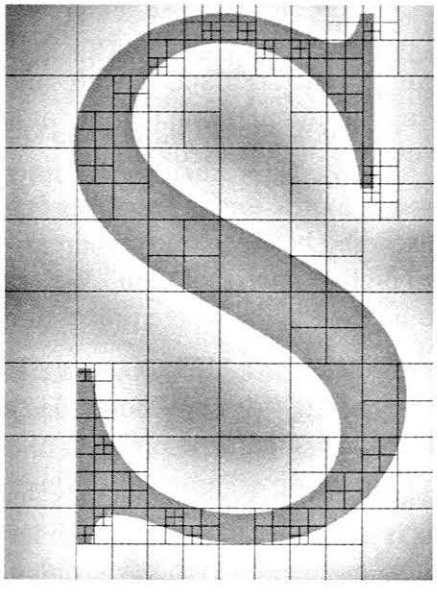
MERL's post-filtering algorithms are unique in attacking both blocking and ringing at the same time and doing so in computationally inexpensive ways. MERL developed a low-complexity block artifact detection method that does a good job of controlling adaptive low-pass filtering. In addition, MERL developed a fuzzy filtering technique that can remove ringing artifacts without blurring edges. To control this filtering process, a novel edge-based classification algorithm is used to adaptively determine the parameters of the fuzzy filter, which is then applied to specific portions of the image.

MERL developed its post-filtering algorithms during 2003 and early 2004. Extensive quality tests were conducted by MELCO and algorithmic refinements were made in close collaboration with Johosoken. The algorithm design was fixed in September of 2004. It appeared in a professional HDTV codec from MELCO's Koriyama works not long thereafter.

The 9/05 version of the algorithm was designed to be implemented in software and used floating-point numbers and exponential functions. To be appropriate for a chip, the algorithm had to be changed to use fixed-point numbers and avoid exponential functions. This change was made by MTL in close collaboration with Johosoken during 2005; and revised code was delivered to Renesas for conversion into part of their DVD recorder chip.

Saffron Digital Typography

In October 2005, Monotype Imaging Inc. signed an exclusive license to sell products based on MERL's Saffron digital type renderer. Monotype is the world's largest seller of computer fonts and font-rendering algorithms, other than those that are sold bundled with computer operating systems. Saffron displays scalable type more esthetically and more efficiently than any other method. Its use should allow Monotype to significantly expand its business, which will lead to significant rewards to MELCO in the form of royalties.



In close collaboration with MELCO's Sentansoken laboratory, MERL is also seeking uses of Saffron in MELCO products. A key feature of Saffron is that unlike other type renderers, it can be directly implemented in an integrated circuit. Efforts are underway to include it in chips used in MELCO products.

Details: Saffron is the outgrowth of work at MRL that goes back to the late 1990s. At that time, 3D computer graphics was a prime area of research at MERL. One of many graphics projects was a long-range fundamental research project to explore new ways of representing 3D shapes

using implicit functions. This resulted in the invention of Adaptively sampled Distance Fields (ADFs).

The most common way of representing 3D graphical objects is to use a mesh of polygons that represent the positions of points on the surface of the object. This is used in a wide range of applications and is facilitated by special-purpose 3D graphics processor chips.

Instead of explicitly representing points in a 2D manifold embedded in 3D space, ADFs represent the distance from each point in space to the nearest point on the manifold. This implicit representation of the manifold is not as efficient as polygonal representations for some tasks, but much more efficient for others.

For instance, if you merely want to enumerate points on the surface and don't care what order they are enumerated in, a polygon representation is better. However, if you want to determine the distance to the surface from a given point in 3D space, then ADFs are better. In addition, many solid modeling tasks such as combining objects together, deforming objects, or checking for collisions can also be done much better with ADFs.

This makes ADFs very promising as a basis for 3D modeling tools. However, 3D modeling tools are not a business that MELCO is involved in. In addition, polygonal representations are very strongly entrenched as the basis for these tools. Therefore, it is not clear whether one could successfully introduce tools with a radically different internal representation.

In 2003, one of the two MRL researchers working on ADFs came to the conclusion that a 2D version of ADFs could be beneficially applied to the rendering of computer type. It was decided to focus on this application both because it was more relevant to MELCO and because it seemed easier to have significant impact on the world.

For many years computers have displayed type (letters and other characters) by using stored bit maps specifying which pixels should be black and which white. For representing type at a small number of fixed sizes, this works very well and is still the primary way of displaying type in devices such as cell phones and electronic equipment. However, if you wish to display type at a wide range of sizes (so called scalable type) bit maps work terribly, because there is no way to compute an esthetic bit map at one size from a bit map at another size.

To display scalable type, 'outline fonts' such as Post Script fonts were introduced. Instead of directly specifying which pixels should be black and which white, an outline font uses parametric curves to define the shape of each letter. From these shapes, one can then compute which pixels should have what values in a process called "rendering". Rendering is traditionally done by determining what percentage of the area corresponding to each pixel a given letter covers. If the pixel is totally covered then it is totally black; if it is half covered it is half black; and so on.

Rendering an outline font is much more complex than using a bitmap but produces beautiful results over a wide range of type sizes. However, as the type size becomes small, problems begin to arise. When there are not very many pixels available to display a character, parts of the character begin to merge together into blobs and thin parts of the character can disappear altogether, causing the character to be unreadable. In contrast, a hand-tuned bitmap can be fully readable at small sizes.

To deal with this problem, the designers of outline fonts introduced the concept of "hints"---procedural instructions associated with each character that specify how to shrink the character

without losing key features that are required for legibility. These hints work effectively. However, they add yet more complexity to the rendering process and designing the hints is a difficult task that adds greatly to the cost of designing a quality font.

Instead of representing characters as outlines, Saffron represents them as 2D distance maps. It then does rendering not based on coverage but rather based on the distance from the center of the pixel to the edge of the character. Experimentation has shown that this approach to rendering behaves much better than coverage-based rendering, as type sizes get small. As a result, hints are no longer required, which greatly simplifies the process of designing a font. In addition, distance-based rendering is much simpler than coverage based rendering---so simple that it can be represented in silicon rather than software.

All in all, Saffron is better than traditional outline font rendering in every way but one---the memory required to represent a character as a distance map is considerably greatly than the memory required to represent a hinted outline. However, by using on-the-fly conversion from outlines to ADFs combined with caching, this memory problem can be dispensed with in any situation where there is sufficient cache space available.

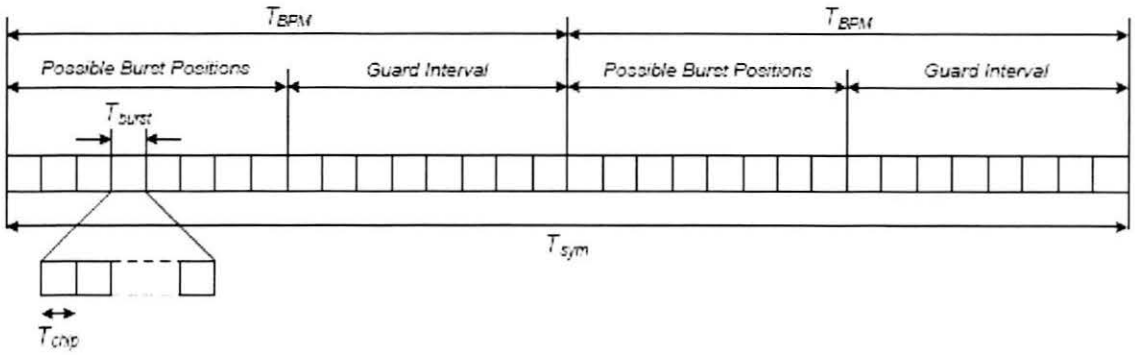
Saffron was first demonstrated in the latter part of 2003 and its quality advantages over standard type rendering were readily apparent to all. In collaboration with MELCO's Sentansoken laboratory, an effort started to put Saffron into chips that are used by MELCO products. Close collaboration between MERL and Sentansoken began in early 2004 and is working toward the design of an appropriate chip.

At the same time, MERL contacted various outside companies that could make use of Saffron. Strong interest was expressed by Macromedia, which licensed it for use in their Flash player in 2004. Once a license was signed with Macromedia, MERL researchers began a crash effort to create an improved version of Saffron for use in the Flash player. This was successfully completed in December of 2004. Macromedia to integrate Saffron into Flash and released it as part of Flash 8 in the summer of 2005.

The 2004 version of Saffron was designed to run on PCs and required a significant amount of computation and large amounts of memory. As such it was not really appropriate for use in embedded devices such as cell phones. During 2005 an entirely new version of Saffron was developed that overcomes both of these limitations. While operating in the same logical way as prior versions of Saffron, it operates directly on outlines without creating ADFs as an explicit intermediate representation.

Monotype was greatly attracted to the potential of the new Saffron for embedded devices and licensed it for general use in their products. In the middle of 2005, MRL researchers began the process of creating final code for delivery to Monotype. This process should be done by the fall of 2006.

2006 will also see the completing of a chip for Melco that implements the same version of Saffron that is being used by Macromedia. Future work will create a chip implementing the much-improved version created for Monotype. We hope that the direct impact of Saffron on MELCO through these chips will be at least as great is the indirect impact due to Royalties from Monotype.



802.15.4a Wireless Communication Standard Contributions

In December of 2005 the IEEE 802.15.4a task group released its initial draft for an Ultra-WideBand (UWB) Low-Rate Wireless Personal Area Network (LR-WPAN). This draft (P802.15.4-Amendmend-a/D2) includes 5 features that were contributed by MERL: hybrid BPSK/PPM modulation, Private Ranging, Adaptive Frame Duration, Spectrum shaping and Scaling Factor for Range Estimate Confidence Interval.

MTL is actively contributing to a wide range of wireless communications standards in addition to IEEE 802.15.4a such as: High Data Rate Ultra-WideBand (IEEE 802.15.3a), High data rate Local Area Networks (IEEE 802.11n), ZigBee, and future cell phone standards (3GPP). MTL is also active in various efforts to make products relating to these standards.

Details: Traditional wireless communication methods use a relatively high level of energy in a relatively narrow band of frequencies to punch through noise. In contrast, UWB uses a low level of energy spread over an extremely wide of range of frequencies. It overcomes noise because its wide band coherency is totally unlike noise. UWB has a number of unique features that give it enormous commercial potential: it is virtually impossible to jam, it can be used in conjunction with traditional wireless communication without interference, it can pass through walls and other obstacles that block other forms of wireless communication and due to its very high frequency components, it can be used to determine highly accurate range information.

UWB communication was developed initially for military applications and is only recently moving into the commercial arena. Due to UWB's promising applications and business potential, MTL started some initial UWB research in 2001. The research and technology development were accelerated after approval for a large UWB project was obtained from MELCO CR&D in 2003.

Since January 2004, MTL has been attending the IEEE 802.15.4a Task Group and developing proposals. The goal of 802.15.4a is to provide robust performance for sensor and control applications while leveraging the unique capability of UWB waveforms to support precision ranging between devices. MERL has taken an active role in the development of the standard, made numerous contributions, and has held key leadership positions in the task group. These positions include: Chair of the channel modeling subgroup, Chair of the UWB pulse modulation and coding subgroup and Vice chair of the Ranging subgroup.

MERL's hybrid BPSK/PPM modulation scheme supports symbol transmission that can be simultaneously received by two kinds of receivers. In the scheme, each symbol interval is divided into two halves. A train of pulses (called a burst) that encodes a bit of data is transmitted

in either the first or the second half of the symbol. Non-coherent receivers decode the bit by detecting which half of the symbol the burst is in. Coherent receivers decode the bit by observing the polarities of the individual pulses. Coherent receivers are more expensive, but more robust and can receive signals at longer ranges and enable higher data rates as opposed to non-coherent ones.

A key feature of UWB communication is that it can be used to determine very accurate inter-node distances in a UWB network. This is important for providing location-based services in the network such as tracking assets that are tagged with UWB transceivers. Range information is determined and communicated by sending special data packets between the nodes in a UWB network. MERL's Private Ranging scheme uses dynamic waveform selection as a privacy tool that protects these special packets from interference and eavesdropping.

One of the key parameters in UWB communication is the Pulse Repetition Interval (the delay between successive data pulses). Using a longer interval reduces interference between pulses. Using a shorter interval allows increased data rates. To support a range of applications, IEEE 802.15.4a supports a number of different pulse repetition intervals. MERL's Adaptive Frame Duration scheme allows a network of nodes to select a repetition interval that achieves the best data rate given the environment the nodes face.

Another challenge of UWB communication is that the spectrum of a UWB signal (the amount of radiated energy at each frequency within the signal) must fit within an FCC-specified 'mask' (a maximum allowed radiated energy level at each frequency). Unfortunately, many conventional impulse radio UWB systems produce relatively large spikes of energy at certain frequencies. To keep these spikes within the mask, the total power of the signal must be significantly reduced. MERL introduced the use of polarity randomization, which eliminates the spikes and leads to general smoothing of the signal spectrum. This allows greater total radiated power and thus greater range/reliability of transmission. In addition, MERL introduced the concept of using a linear combination of basis pulses to enable the detailed shaping of the UWB spectrum so as to avoid frequencies that contain interferences.

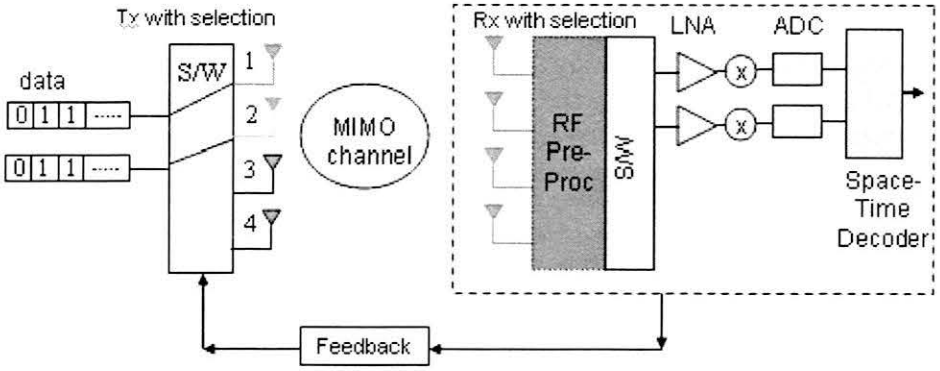
In IEEE 802.15.4a the accuracy of ranging information is specified using two metrics: confidence level and confidence interval. MERL's Scaling Factor for Range Estimate Confidence Interval introduces two bits in a range data packet that scale the confidence interval so that it can represent a wider dynamic range.

The IEEE 802.15.4a standard is on track to be completed in the first half of 2007. As a result, MERL's work on the standard itself is drawing to a close. However, MTL has an active project creating business opportunities for MELCO with systems based on IEEE 802.15.4a.

Multiple Antenna Technology in IEEE 802.11n

In January 2006, two key contributions from MERL, Antenna Selection for MIMO and Space-Time Block Coding With Unequal Modulation were accepted into the standard draft of IEEE 802.11n, which is intended to be a high-speed extension of the popular WiFi standard. 802.11n increases WiFi's data rate from the current 54 Mbit/s to 600 Mbit/s, by using multiple-antenna technology. MERL's contributions to this standard help to decrease the implementation costs and increase transmission quality.

MTL is actively contributing to a wide range of wireless communications standards in addition to IEEE 802.11n such as Low Data Rate Ultra-WideBand (IEEE 802.15.4a), High Data Rate Ultra-WideBand (IEEE 802.15.3a), ZigBee, and future cell phone standards (3GPP). MTL is also active in various efforts to make products relating to these standards.



Antenna Selection in MIMO Systems

Details: Wireless Local Area Networks (LANs) are one of the big wireless success stories of the last decade. Due to the proliferation of laptops, and the desire of users to connect to the Internet from a variety of different locations, high-speed wireless LANs are in great demand. The data rate provided by those LANs has been increasing dramatically over the years, from 1 Mbit/s in the first standard (IEEE 802.11), to 11 Mbit/s in 802.11b (often known as WiFi), to 54 Mbit/s in the current high-speed standard. However, new applications demand even higher transfer speeds; for this reason, a 600 Mbit/s system (standardized as IEEE 802.11n) is currently being developed.

The goal of IEEE 802.11n is to achieve this high data rate by using multiple-antenna technologies, also known as MIMO (Multiple-Input Multiple-Output) systems. In the simplest form of MIMO, a high-rate data stream is split up into several, lower-rate data streams, which are then transmitted from different transmit antenna elements. The propagation from the transmitter to the receiver mixes the signals together. At the receiver, the signal is received at multiple antenna elements, and sophisticated signal processing is used to separate the low-rate data streams from each other. They are then recombined to recover the original high-rate data stream. Because low-rate streams are sent in parallel, the data rate of the total system can be extremely high without excessive power consumption.

Because it was clear that MIMO would be central to developments in several different areas of wireless communication, including cell phones as well as WLANs, MTL began research on MIMO in 2002. This put MERL in a good position to participate in 802.11n whose main technical development began in 2004. In close collaboration with MELCO's Johosoken laboratory, MTL suggested a number of key technologies in both the physical (PHY) and Medium Access Control (MAC) layers of 802.11n.

During the course of standard development, two major industry alliances emerged: TGnSync (Intel, Agere, ...) and WWiSE (Texas Instruments, Broadcom, ...). In December 2004, MELCO/MERL became a part of the TGnSync alliance. In the following months, negotiations between the two alliances sought to find a compromise system. Within the negotiation group, a new alliance, EWC (Enhanced Wireless Consortium) emerged; MERL was one of the early

members of this new consortium in addition to Intel, Broadcom, Atheros, and others. This new consortium successfully led the development of a compromise specification, which was adopted by TGnSync as a whole and then by IEEE 802.11n. Final refinements are currently being done within the framework of 802.11n, but no major changes are expected.

One of the key problems of MIMO systems is hardware complexity. Each receive antenna element requires a dedicated radio frequency (RF) chain that comprises a low noise amplifier, a frequency down-converter, and an analog to digital converter; and analogously for each transmit antenna a dedicated power amplifier, a frequency up-converter, and a digital to analog converter. In addition, processing the signals received in many MIMO systems requires sophisticated signal processing whose complexity increases rapidly (sometimes even exponentially) with the number of transmit and receive antennas.

Antenna selection reduces the hardware complexity of transmitters and receivers by using fewer RF chains than the number of antenna elements. A subset of the available antenna elements is adaptively chosen and only signals from the chosen antennas are processed further by the available RF chains. MERL began studying antenna selection in 2002 and is acknowledged as a leader in this field of MIMO.

Following proposals made by MERL, the 802.11n standard now contains all the necessary provisions to include antenna selection in a high-speed wireless LAN system. In particular, special messages are present that make it possible to select the best antennas to use (through a so-called antenna training protocol).

Space-time block coding (STBC) is a technique that has been widely used to increase signal robustness when two transmit antennas are available. By transmitting slightly modified copies of the same signal from the two antennas, the probability that the signal gets corrupted on the way to the receiver is drastically reduced.

MERL's Space-Time Block Coding With Unequal Modulation scheme makes it possible to send data streams robustly using space-time coding, and to do so at a higher data rate using multiple space-time code blocks. Its ability to fine tune the transmission rates of the individual streams allows one to optimize the total throughput of the system.

Face Detection and Object-Tracking

In February 2006, Lihon's Kyoto works began selling the DX-PC55PRO software package, which allows a PC to significantly extend the capabilities of their Digital Video Recorder (DVR) products for surveillance such as the TL5000U. The DX-PC55PRO supports sophisticated searching of video data based on features that are automatically extracted using two technologies from MERL: face detection and object tracking. In addition to traditional DVR searching based on time and alarm conditions, these features allow searching based on the faces that are

visible and on the pattern of movement of objects.

MERL has a long history of working in the area of Computer Vision at MRL. MELCO has a long history of products in the area surveillance and access control---Lihon's Kyoto works developed the world's first time-lapse video recorder. In 2000, MERL focused its vision research on the observation of people in video with the goal of developing technology for MELCO's various surveillance and access-control products. This led in 2001 to the staffing of a computer vision application group in MTL and a large surveillance-related project starting in 2002. This in turn has lead to a range of technologies in the areas of face detection, tracking and face recognition.

Details: 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.

In some ways, detecting a face is not that hard. Faces look very different from other things and given a picture, which either is or is not completely filled with a face, it is not hard to write a program that will determine whether it is a face in a fraction of a second. However, that is not the question that is usually being asked. Rather, you are given a picture and asked to determine whether there is a (perhaps quite small) face somewhere in it. In order to do that, you have to answer the simpler question above not once, but tens of thousands of times at tens of thousands of different places and sizes in the image as a whole. To get good performance, you now have to be able to answer the simple question not in a fraction of a second, but in only a few microseconds and you have to be extremely accurate in each answer because you have to answer such a huge number of questions. This is not easy at all.

The MERL 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. The classifier is also asymmetric in the sense that while it takes a moderate amount of time to be sure that a face is a face, it can determine that something that looks nothing like a face is not a face in an extremely small amount of time. This is important because when searching for faces in an image, almost every sub-image is in fact not a face. Due to the two properties above, the face detector is fast enough that it can locate faces in video in real time and has become the world's most commonly used method of face detection.

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

This improved face detector was demonstrated to MELCO in the summer of 2002. During 2003 through 2005, MTL continued to refine the face detector producing a highly accurate and robust module that can locate 98% of the faces in images while falsely identifying non-face objects as faces only one time in a million. In 2004 the demonstrated a 'face browsing' system that can use the face shown in a video as index into the video.

A collaborative effort began with MELCO's Johosoken laboratory in 2002, which lead to the face detector being used in a MELCO cell phone in 2004. As described in more detail below, a collaborative effort also got underway with MELCO's Sentansoken laboratory in 2002, which led to the DX-PC55PRO.

MERL's work on tracking has its roots in MERL's video-related research. MERL's first tracker was implemented in early 2002 based on video object segmentation technology that had been developed as part of his PhD thesis by a newly hired MTL researcher. From this simple start, the tracker evolved into a fully competent tracker by the end of 2003 and more recently into a tracker comparable to the best in the world. In particular, it was extended to deal with low-frame-rate video in 2004 and moving backgrounds in 2005. In 2004, a new MRL researcher arrived with strong tracking expertise. Together, MTL and MRL are continuing to make world-leading progress in tracking research. Current work includes an investigation of tracking in IR as well as visible images.

In the simplest case tracking is deceptively easy. If you have a fixed camera looking at a fixed background and wish to track a single moving object, this can be done almost trivially. All you have to do is subtract the background image from the image captured at any given moment and all that is left is the object to be tracked. The task is reduced to one of following a moving blob on a totally black background, which is trivial.

However, things are seldom that simple. To start with, backgrounds are never really static. Illumination and shadows change with the time of day. Clouds move by. Wind blows trees and makes waves in water. Semi-static objects, such as parked cars, move around from time to time. To deal with illumination changes, the MERL tracker uses a layered model of the static background using a Bayesian update formulation and compares it with the current frame. For more challenging scenes where the background is in motion (e.g., with sea waves or blowing trees) the MERL tracker switches to frequency domain analysis and intrinsic image decomposition.

In addition, you are seldom just looking at one moving object that is always in front of the background. It can be occasionally hidden behind other objects. Due to changes in shape or orientation, the appearance of a moving object can alter radically over time. Two objects can move together and then move apart, in which case it can be very hard to determine whether they passed each other or rebounded from each other. All this is complicated by the fact that if you are using low frame rate video (as is often the case in surveillance situations) an object can move a long way between frames. To deal with these issues, the MERL tracker also uses statistical rather than static models of the moving objects based on color and orientation information.

In effect, MERL's tracker employs on-the-fly generated discriminators that have knowledge both of the background and a particular object. Given a search window located where an object is predicted to be based on its past motion, a discriminator in effect pulls the center of the window toward the center of the object and pushes it away from the background. These discriminators are efficient enough that the MERL tracker can follow multiple fast moving objects even in slow-frame-rate video.

In 2003, Sentansoken began a major project to develop technology for MELCO's surveillance and access control projects. MERL has collaborated closely with this project from the start and MERL's face detection, face recognition, and tracking technology play an important role in the project. During 2004 and 2005 Sentansoken has taken the lead in getting MERL's technology

into MELCO products and has had overall system responsibility. MERL has continued intensive work on making its component technologies more capable and more robust along with researching future technologies and applications. The DX-PC55PRO is the first of fruit of this collaboration. Several more will follow in the coming year.

In the DX-PC55PRO, face detection is used to find faces in video. The software stores examples of the most clearly visible facial image of each person in the video, and presents them to the user as thumbnail images that summarize the persons in the video as in MERL's prototype face browser. The users can then drag and drop images to narrow down a search and thereby identify people more efficiently. Object tracking supports a product feature called 'behavioral pattern search'. This feature is based on analysis of the movement of people and vehicles in the recorded video and storage of their movement trajectories. The trajectory data allows rapid search and retrieval of video containing people matching a particular movement pattern. Both face-based and trajectory-based searches can be limited by the conventional time-lapse-video-recorder search criteria such as time, alarms and other bookmarks.

Bidirectional Reflectance Distribution Function (BRDF) Data

In April of 2005, Microsoft Corp licensed bidirectional reflectance distribution function (BRDF) data for a collection of materials from MERL. BRDF data for a material is a very detailed description of the way light reflects from the material. Microsoft intends to use this data to create more realistic looking objects in computer games.

MERL's work on BRDFs was a small step in a series of projects exploring different kinds of representations for 3D objects in computer graphics. Key results from this larger series of projects include work on volume graphics representations, which led to a high speed volume graphics rendering chip, work on Adaptively sampled Distance Fields (ADFs) which led to the work on Saffron described above, and work on gathering very detailed 3D information about the appearance of human faces as a basis for improved face recognition.

Details: MERL's work on BRDFs was carried out in 2002 in a collaboration between two MRL researchers, an MIT professor and an intern from MIT. The group extended the theory of BRDFs, designed a practical device for measuring isotropic BRDFs, and collected BRDF data for more than 150 different materials. The student continued work on BRDFs after returning to MIT and made it the basis of his PhD thesis in 2003.

Objects appear as they are to our eyes because light falls on them and is then reflected off to our eyes. The reflective process changes the color and intensity of the light giving the object its characteristic appearance. In the simplest model of this process white light falling on an object leads to reflected light that has some particular color and is equally reflected in all directions. However, real world situations are typically more complex. Light is not equally reflected in every direction and it is typically not exactly the same color in each direction.

The simple uniform-reflection model is used as the basis for most work in computer graphics. This causes all the objects in a computer animated movie like "Toy Story" to look like they are

made of plastic, because plastic is one of the few things in the real world that comes close to obeying the uniform-reflection model. To generate more realistic looking images you have to have a more realistic model for how light is reflected.

A class of functions called Bidirectional Reflectance Distribution Functions (BRDFs) characterizes a process where light transformation and reflection occurs non-uniformly at an idealized surface point. Traditionally, physically or empirically inspired analytic reflection models are used. These BRDF models are only approximations of reflectance of real materials. Furthermore, most analytic reflection models are limited to describing only particular subclasses of materials---A given model can represent only the phenomena for which it is designed.

The MRL researchers developed a generative model for isotropic BRDFs based on measured reflectance data. This results in a model that is much easier to use and control than previous analytic models in which the meaning of parameters is often non-intuitive. Our model is also more realistic, because it is based on real data from real materials.

For various angles of incident light on a material such as a piece of wood, leather, granite, etc., MERL's BRDF data observation device records exactly what color and intensity of light is reflected in each direction. This detailed information can then be used in conjunction with MERL's data driven BRDF formalism as the basis for very realistic graphic rendering of the material.

MERL's collection of BRDF data has proven to be a valuable aid to research on computer graphics; and MERL has made it available to nearly 100 researchers around the world over the past 3 years. Other companies may follow Microsoft's example and purchase the collection for commercial use.

The most important follow-on of this work at MERL is research on how to acquire and represent very accurate information about the appearance of a human face. This face data goes beyond BRDFs and makes it possible to very accurately model how the appearance of a human face changes when the direction of incident light changes. We believe that this will be important as a basis for face recognition that is much more reliable than current face recognition systems.

Incremental Imputative Singular Value Decomposition

In April of 2005, Mars Inc. licensed MERL's Incremental Imputative Singular Value Decomposition (IISVD) software. IISVD makes it possible to do rapid forecasting based on very large collections of data. Mars intends to use it for improving their internal production scheduling based on better forecasts of future demand based on past demand.

MERL's work on IISVD is one example of a range of work at MERL on data analysis and forecasting. This includes work ranging from traffic flow prediction to representing the underlying data for face recognition.

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 with high dimensional data to analyze the fundamental structure of the data. For instance, 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 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 a 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 MELCO people in the fall of 2001. It was seen by a researcher at MELCO's Johosoken laboratory who thought that it would fit in well with 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 that was included in the Diaprism product line in 2003.

The candy manufacturer Mars inc. does periodic forecasting as a basis for production scheduling. Traditionally this has been a very length process taking large amounts of computer time due to the large amounts of data involved. Using IISVD, Mars will be able to do this forecasting much more quickly thereby obtaining more timely information. Similar applications exist in many places within MELCO.

Technical Staff

The most important assets of MERL are its people. The following pages present the capabilities and interests of MERL's current technical staff members. Additional information about their work can be found in the publications list and the project descriptions in this report. Complete information can be found in people's individual web pages at "http://www.merl.com/people".

Amit Agrawal *Ph.D., University of Maryland, 2006*
Visiting Scientist MRL

Prior to his graduate studies, Agrawal worked as a DSP engineer at Hughes Software Systems, India, for one year. His research interests are in computer vision, image processing and computational photography. Current projects include motion photography, flash photography, surface reconstruction from gradient fields, high dynamic range imaging and image editing under variable illumination using gradient domain methods.

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. He also had a faculty position at the Interdisciplinary Center, Herzlya, Israel. He also 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.

Luigi Baccari *B.S., University of Massachusetts of Lowell*
System & Network Administrator

Baccari has 23 years of experience in the System and Network Administrations field. For the 6 years prior to joining MERL he worked at HP/Compaq's Cambridge Research Labs providing System and Network. Previous to that he worked for Force Computers, Lycos and Digital Equipment Corp. as Data Center Manger and in various System/Network Support roles.

John C. Barnwell III
Research Assistant MRL

John Barnwell is a former Software Engineer developing configuration and database systems for the aircraft manufacturing, food processing, large truck manufacturing and computer manufacturing industries. His current personal interests include amateur radio, CNC control systems and mechanical and electrical design.

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*
Principal 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.

Stephen Burgess *B.S., Bath University, 1980*
Principal Technical Staff MTL

Steve Burgess has extensive and diverse experience in ASIC/FPGA and digital board hardware design. He was a member of the ASIC design team and principal board designer for VolumePro, MELCO's real-time volume rendering hardware for PCs. Steve was also a hardware development team leader for MELCO's PXB1-E CPU. Currently he is working on the development of a chip for the Saffron type renderer.

Eric Chan *M.S., Massachusetts Institute of Technology, 2005*
Visiting Scientist MRL

Chan's research is focused on graphics architectures, shading languages, and real-time rendering techniques. At MIT, he developed efficient methods for rendering hard and soft shadows. Previously at Stanford, he wrote compiler back ends for the NV30 and R300 fragment architectures. At MERL he is working on software implementations of the Saffron type renderer.

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.

Paul H. Dietz *Ph.D., Carnegie Mellon University, 1995*
Senior Research Scientist MRL

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*
Senior 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

Chunjie Duan *M.S. Colorado State University, 1998*
Visiting Principal Technical Staff

Prior to joining MERL, Duan worked for Alcatel, Qualcomm and Ericsson and other telecomm companies for over 10 years. His research interests are in wireless and optical communications, digital signal processing and VLSI/CAD technology. He is currently working on Ultra-Wideband system development and LSI implementation.

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).

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.

Sinan Gezici *Ph.D., Princeton University, 2006*
Visiting Technical Staff MTL

Gezici's main research interests are in the fields of signal detection, estimation and optimization theory, and their applications to wireless communication systems. Currently, he has a particular interest in timing estimation, performance analysis and receiver design for ultra-wideband systems.

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.

Yuri A. Ivanov *Ph.D., Massachusetts Institute of Technology, 2001*
Principal Technical Staff MTL

Ivanov's main research interests lie in the area of Computer Vision, Machine Learning and Data Mining. In particular, he is interested in dynamic observations - video sequences, sounds, gestures, actions and events.

Ronald L. Johnson *B.A., Louisiana State University, 1990*
Manager Computational & Network Services

Johnson received his B.A. in English Literature from Louisiana State University in 1990. Ronald began work for MERL as the manager of Computational and Network Services in September of 2004, and supports all of the Lab's computational and network efforts. Ronald has 14 years experience in Systems and Network Administration along with 4 years of software development experience.

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

Jones joined MERL in 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.

Keisuke Kojima *Ph.D., University of Tokyo, 1990*
Senior Research Scientist MRL

Kojima spent 8 years in Melco's Sentansoken research lab and 9 years at AT&T Bell Labs. He has been involved in the research and development of semiconductor lasers, optical communication modules, and optical communication and sensor systems. At MERL, he is engaged in the research of security systems and sensor technologies.

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

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.

Yohei Matsuura *M.Sc., University of Tsukuba, 1997*
Visiting Researcher MTL

Yohei Matsuura received his B.S. degree in Information science from Tsukuba University. Then he became a member of Jyoho-soken in 1997, he started to develop high reliability and high performance computing system on UNIX. His current interest technology is sensor network and data mining that are related to electric power management.

Wojciech Matusik *Ph.D., Massachusetts Institute of Technology, 2003*
Research 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.

Janet McAndless
Technical Associate MRL

McAndless formerly held a variety of tech-related positions including management of peer review processes, web development, film and television post production, technical writing/documentation, and technical-support. For many years she has been involved with the planning of the annual SIGGRAPH conferences.

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 the 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.

Clifton D. Mueller *J.D., Boston College Law School, 2005*
Patent Agent

Clifton Mueller received his Bachelor of Science degree from M.I.T. in 1997 with a major in Chemistry and minors in Biology and Music Composition. Prior to starting law school in 2002, he was a research associate at PRAECIS Pharmaceuticals, Inc. in Cambridge, MA working on therapies for Alzheimer's disease, rheumatoid arthritis and HIV.

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.

Barton Nicholls *Northeastern University*
Systems & Network Administration

Nicholls is a member of MERL's Computer Network Services Group. He supports primarily UNIX and some Windows client and infrastructure software and hardware, and networking services for MERL. He comes to us from Verizon's Network Operations Management Group, and before that Information Technology at Art Technology Group.

Daniel N. Nikovski *Ph.D., Carnegie Mellon University, 2002*
Principal Technical Staff MTL

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*
Principal 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*
Distinguished 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.

Fatih M. Porikli *Ph.D., Polytechnic University, 2002*
Principal 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).

Jeffrey S. Proctor *B.S., West Chester University, 1985*
Visiting Scientist MRL

Proctor's experience includes full custom, standard cell, ASIC, and gate array design for CPU's, caches, I/O channels, and high speed interconnects (which most recently consisted of a 70,000 SPARC system at Sun). His current project is an FPGA implementation of a replicated horizontal group shuffled belief propagation low density parity check decoder.

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 lab.

Ramesh Raskar *Ph.D., University of North Carolina at Chapel Hill, 2002*
Senior 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.

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*
Principal 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*
Team Leader / 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.

Hugh Secker-Walker *M.S., Massachusetts Institute of Technology, 1989*
Principal Technical Staff MTL

Hugh came to MERL after 11 years working on Dragon speech-recognition technology. At Dragon and its successors he worked on many aspects of the core speech-recognition algorithms, engine design, and productized delivery for a large-vocabulary speech recognizer used both for research and commercially. Technically he is interested in languages, rapid prototyping, and technology transfer.

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

Shen's current research focuses on shared interactive surfaces. She led the MidART project, which MELCO has incorporated into several distributed industrial plant control systems. MidART is a real-time middle-ware 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 interests include real-time analysis of video and audio content, and real-time and distributed operating systems. He has worked on the Video Summarization, TimeTunnel, DiamondTouch, Open Community, and Network Replication projects, and on smaller efforts related to fingerprint recognition, MPEG-7, and interactive surroundings.

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.

Koon Hoo Teo *Ph.D., University of Alberta 1990*
Senior Principal Technical Staff MTL

Teo was with Nortel for about 15 years where he was actively involved in the research and implementation issues of a number of 3G and 4G wireless systems including Wireless Mesh Networks and WiMAX systems. His current research interests include Cognitive Radio, location tracking using Ultra Wideband technology, and Wireless Mesh and Multi-Hop Systems.

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, focusing on 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*
Principal 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*
Senior Team Leader / Senior Principal Technical Staff MTL

Vetro joined MERL in 1996. His research interests are related to the encoding and transport of multimedia content. He has been an active participant in MPEG standards for several years. Dr. Vetro has contributed several technologies to MELCO/Renesas products, including MPEG-2/4 transcoding for surveillance, post-filtering for artifact reduction and video down-decoding for a low-cost DTV receiver chip.

Benjamin W. Vigoda *Ph.D., Massachusetts Institute of Technology*
Research Scientist MRL

Before joining MERL, Vigoda was a Ph.D. student and post-doctoral associate at MIT. Ben works at the intersection of machine learning and integrated circuit design. At MERL, he is principal investigator of the RF Analog Logic MERL/DARPA/MIT joint project to implement statistical inference and signal processing capabilities with RF analog circuits.

Gene V. Vinokur *M.S., Boston University, 2003*
Associate Patent Agent

Prior to joining MERL, Vinokur spent four years at Putnam Investments developing software applications for the financial industry. He has been a licensed Patent Agent since 2003. He joined MERL's Patent Department in 2006.

Garrett Weinberg *B.A., Yale University, 2000*
Member Technical Staff MTL

Before coming to MERL, Weinberg designed and internationalized automotive speech user interfaces at Dragon Systems, and was a chief architect and implementer of enterprise solutions for two Boston-area startups specializing, respectively, in Digital Rights Management and portfolio management. At MERL, he is leading the effort to port SpokenQuery technologies to various mobile and embedded platforms.

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 MTL

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.

Sehoon Yea *Ph.D., Rensselaer Polytechnic Institute, 2006*
Member Technical Staff MTL

From 1996 to 2001, Yea was a Research Engineer at the Institute for Advanced Engineering in Korea, working on control systems such as industrial robots and servo-drivers. In the summer of 2004, he was an Intern with Sarnoff Corporation, Princeton, NJ. Since joining MERL in January 2006 he has work has focused on digital image and video compression, enhancement and communication.

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, his particular 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).

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, She worked for Nortel Networks for 10 years where she held engineering and management positions in the areas of VLSI design, advanced wireless technology development and wireless & optical networks. She has a broad technical background, specializing in system design and real-time embedded software for wireless communications.

Shengjie Zhao *Ph.D., Texas A&M University, 2004*
Visiting Technical Staff MTL

Before joining MELR, Zhao worked with Wireless Innovation Open Laboratories, Lucent Technologies and NEC Laboratories America, Inc. on system and algorithm designs of the PHY and MAC layer of WiMAX (802.16e). His current interests include system and algorithm design, base band signal processing, system level and link level simulations of wireless channels and networks.

Recent Major Publications

The following lists the 140 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 refereed conference proceeding, or some other significant publication such as a book.

An asterisk (*) appears before the 43 (30.7%) 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 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 CVPR 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. 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.

* Moghaddam, B.; Weiss, Y.; Avidan, S., "Generalized Spectral Bounds for Sparse LDA", *International Conference on Machine Learning (ICML)*, June 2006 (TR2006-046)

Ngan, A.; Durand, F.; Matusik, W., "Image-driven Navigation of Analytical BRDF Models", *Eurographics Symposium on Rendering (EGSR)*, June 2006 (TR2006-061)

McGuire, M.; Matusik, W.; Yerazunis, W.S., "Practical, Real-time Studio Matting using Dual Imagers", *Eurographics Symposium on Rendering (EGSR)*, June 2006 (TR2006-062)

Zwicker, M.; Matusik, W.; Durand, F.; Pfister, H., "Antialiasing for Automultiscopic 3D Displays", *Eurographics Symposium on Rendering (EGSR)*, June 2006 (TR2006-063)

* Agrawal, A.; Raskar, R.; Chellappa, R., "Edge Suppression by Gradient Field Transformation using Cross-Projection Tensors", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2006 (TR2006-058)

* Porikli, F.M.; Tuzel, O., "Covariance Tracking using Model Update Based on Lie Algebra", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2006 (TR2005-127)

Bashir, F.; Porikli, F.M., "Performance Evaluation of Object Detection and Tracking Systems", *IEEE International Workshop on Performance Evaluation of Tracking and*

Surveillance (PETS), June 2006 (TR2006-041)

* Porikli, F.M.; Tuzel, O., "Covariance Tracker", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2006 (TR2006-042)

Goubet, E.; Katz, J.; Porikli, F.M., "Pedestrian Tracking Using Thermal Infrared Imaging", *SPIE Conference Infrared Technology and Applications XXXII*, Vol. 6206, pp. 797-808, June 2006 (TR2005-126)

Guvenc, I.; Sahinoglu, Z.; Orlik, P.V., "TOA Estimation for IR-UWB Transceiver Types: AWGN Channel Analysis", *IEEE International Conference on Communications (ICC)*, June 2006 (TR2006-023)

Mehta, N.B.; Molisch, A.F.; Wu, J.; Zhang, J.Y., "Approximating the Sum of Correlated Lognormal or Lognormal-Rice Random Variables", *IEEE International Conference on Communications (ICC)*, June 2006 (TR2006-010)

Orlik, P.V.; Zhao, S.; Molisch, A.F., "A Hybrid UWB Modulation Design Compatible for Both Coherent and Transmit-Reference Receivers", *IEEE International Conference on Communications (ICC)*, June 2006 (TR2005-087)

Nikolova, E.; Brand, M.E.; Karger, D.R., "Optimal Route Planning under Uncertainty", *International Conference on Automated Planning and Scheduling (ICAPS)*, June 2006 (TR2006-060)

Wigdor, D.; Shen, C.; Forlines, C.L.; Balakrishnan, R., "Table-Centric Interactive Spaces for Real-Time Collaboration", *Advanced Visual Interfaces (AVI)*, May 2006 (TR2006-012)

Wittenburg, K.B.; Lanning, T.; Schwenke, D.L.; Shubin, H.; Vetro, A., "The Prospects for Unrestricted Speech Input for TV Content Search", *ACM Advanced Visual Interfaces (AVI)*, May 2006 (TR2006-045)

Esenther, A.W.; Ryall, K., "Fluid DTMouse: Better Mouse Support for Touch-Based Interactions", *Advanced Visual Interfaces (AVI)*, May 2006 (TR2006-001)

Tse, E.; Shen, C.; Greenberg, S.; Forlines, C.L., "Enabling Interaction with Single User Applications through Speech and Gestures on a Multi-User Tabletop", *Advanced Visual Interfaces (AVI)*, May 2006 (TR2005-130)

Radhakrishnan, R.; Divakaran, A., "Generative Process Tracking for Audio Analysis", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2006 (TR2006-053)

Raj, B.; Shashanka, M.V.S.; Smaragdis, P., "Latent Dirichlet Decomposition for Single Channel Speaker Separation", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2006 (TR2006-064)

Shashanka, M.V.S.; Smaragdis, P., "Secure Sound Classification: Gaussian Mixture Models", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2006 (TR2006-065)

* Wren, C.R.; Tapia, E.M., "Toward Scalable Activity Recognition for Sensor Networks", *International Workshop on Location- and Context-Awareness (LoCA)*, May 2006 (TR2006-011)

Reynolds, C.J.; Wren, C.R., "Worse is Better for Ambient Sensing", *Pervasive 2006: Workshop on Privacy Trust and Identity Issues for Ambient Intelligence*, May 2006 (TR2006-005)

Ivanov, Y.A.; Wren, C.R., "Toward Spatial Queries for Spatial Surveillance Tasks", *Pervasive 2006: Workshop Pervasive Technology Applied Real-World Experiences with RFID and Sensor Networks (PTA)*, May 2006 (TR2006-051)

* Avidan, S.; Butman, M., "Blind Vision", *European Conference on Computer Vision (ECCV)*, May 2006 (TR2006-006)

* Agrawal, A.; Raskar, R.; Chellappa, R., "What is the Range of Surface Reconstructions from a Gradient Field?", *European Conference on Computer Vision (ECCV)*, May 2006 (TR2006-021)

* Avidan, S., "SpatialBoost: Adding Spatial Reasoning to AdaBoost", *European Conference on Computer Vision (ECCV)*, May 2006 (TR2006-014)

* Tuzel, O.; Porikli, F.M.; Meer, P., "Region Covariance: A Fast Descriptor for Detection and Classification", *European Conference on Computer Vision (ECCV)*, May 2006 (TR2005-111)

* Tse, E.; Greenberg, S.; Shen, C.; Forlines, C.L., "Multimodal Multiplayer Tabletop Gaming", *International Workshop on Pervasive Gaming Applications (PerGames)*, May 2006 (TR2006-009)

Brand, M.E., "Fast Low-Rank Modifications of the Thin Singular Value Decomposition", *Linear Algebra and Its Applications*, Vol. 415, Issue 1, pp. 20-30, May 2006 (Elsevier Science, TR2006-059)

Martinian, E.; Behrens, A.; Xin, J.; Vetro, A., "View Synthesis for Multiview Video Compression", *Picture Coding Symposium (PCS)*, April 2006 (TR2006-035)

Xin, J.; Vetro, A., "Fast Mode Decision for Intra-only H.264/AVC Coding", *Picture Coding Symposium (PCS)*, April 2006 (TR2006-034)

Nikovski, D.N.; Kulev, V., "Induction of Compact Decision Trees for Personalized Recommendation", *ACM Symposium on Applied Computing (SAC)*, April 2006 (TR2006-036)

* Wigdor, D.; Shen, C.; Forlines, C.L.; Balakrishnan, R., "Effects of Display Position and Control Space Orientation on User Preference and Performance", *ACM Conference on Human*

Factors in Computing Systems (CHI), ISBN: 1-59593-372-7, pp. 309-318, April 2006 (TR2006-015)

* Shen, C.; Esenther, A.W.; Forlines, C.L.; Ryall, K., "Three Modes of Multi-Surface Interaction and Visualization", *ACM Conference on Human Factors in Computing Systems (CHI)*, April 2006 (TR2006-025)

* Cheng, H.; Zhang, X.M.; Shi, Y.Q.; Vetro, A.; Sun, H., "Constant Quality Rate Allocation for FGS Coding Using Composite R-D Analysis", *IEEE Transactions on Multimedia*, ISSN: 1520-9210, Vol. 8, Issue 2, pp. 405-407, April 2006 (TR2006-019)

* Guvenc, I.; Sahinoglu, Z.; Orlik, P.V., "TOA Estimation for IR-UWB Systems with Different Transceiver Types", *IEEE Transactions on Microwave Theory and Techniques*, ISSN: 0018-9480, Vol. 54, Issue 4, Part 2, pp. 1876-1886, April 2006 (TR2006-023)

* Nakache, Y-P; Molisch, A.F., "Spectral Shaping of UWB Signals for Time-Hopping Impuse Radio", *IEEE Journal on Selected Areas in Communications*, ISSN: 0733-8716, Vol. 24, Issue 4, Part 1, pp. 738-744, April 2006 (TR2005-120)

* Gezici, S.; Sahinoglu, Z.; Kobayashi, H.; Poor, H.V., "Ultra-Wideband Impulse Radio Systems with Multiple Pulse Types", *IEEE Journal on Selected Areas in Communications*, ISSN: 0733-8716, Vol. 24, Issue 4, Part 1, pp. 892-898, April 2006 (TR2005-027)

Zhang, H.; Molisch, A.F.; Gu, D.; Wang, D.; Zhang, J., "Antenna Selection in High-Throughput Wireless LAN", *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (ISBMSB)*, April 2006 (TR2006-057)

Tao, Z.; Panwa, S.; Gu, D.; Zhang, J., "Performance Analysis and a Proposed Improvement for the IEEE 802.15.4 Contention Access Period", *IEEE Wireless Communications and Networking Conference (WCNC)*, April 2006 (TR2006-056)

Martinian, E.; Wainwright, M., "Low Density Codes Achieve the Rate-Distortion Bound", *Data Compression Conference (DCC)*, ISSN: 1068-0314, pp. 153-162, March 2006 (TR2005-125)

* Shafi, M.; Zhang, M.; Moustakas, A.L.; Smith, P.J.; Molisch, A.F.; Tufvesson, F.; Simon, S.H., "Polarized MIMO Channels in 3-D: Models, Measurements and Mutual Information", *IEEE Journal on Selected Areas in Communications*, ISSN: 0733-8716, Vol. 24, Issue 3, pp. 514-427, March 2006 (TR2006-020)

Zhang, J.; Fossorier, M.; Gu, D.; Zhang, J., "Two-Dimensional Correction for Min-Sum Decoding of Irregular LDPC Codes", *IEEE Communication Letters*, ISSN: 1089-7798, Vol. 10, Issue 3, pp. 180-182, March 2006 (TR2006-055)

Green, P.; Kautz, J.; Matusik, W.; Durand, F., "View-Dependent Precomputed Light Transport Using Nonlinear Gaussian Function Approximations", *Symposium on Interactive 3D Graphics (i3D)*, ISBN: 1-59593-295-X, pp. 7-14, March 2006 (TR2006-028)

* Moghaddam, B.; Weiss, Y.; Avidan, S., "Spectral Bounds for Sparse PCA: Exact & Greedy Algorithms", *Advances in Neural Information Processing Systems (NIPS)*, Y. Weiss & B. Scholkopf & J. Platt (Eds.), December 2005 (TR2006-007)

Sidner, C.L.; Lee, C.H.; Morency, L-P; Forlines, C.L., "The Effect of Head-Nod Recognition in Human-Robot Conversation", *ACM SIGCHI/SIGART Conference on Human-Robot Interaction (HRI)*, ISBN: 1-59593-294-1, pp. 290-296, March 2006 (TR2005-129)

Guo, J.; Zhang, J.; Chen, W., "A Dynamic QoS Algorithm for Wireless AV Home Networks", *International Conference Advanced Communication Technology (ICACT)*, Vol. 3, pp. 1526-1531, February 2006 (TR2005-140)

* Otsuka, I.; Radharkishnan, R.; Siracusa, M.; Divakaran, A.; Mishima, H., "An Enhanced Video Summarization System Using Audio Features for a Personal Video Recorder", *IEEE Transactions on Consumer Electronics*, ISSN: 0098-3063, Vol. 52, Issue 1, pp. 168-172, February 2006 (TR2006-024)

Kim, C.; Xin, J.; Vetro, A.; Jay Kuo C.C., "Complexity Scalable Motion Estimation for H.264/AVC", *SPIE Conference Visual Communications and Image Processing*, Vol. 6077, pp. 109-120, January 2006 (TR2006-004)

Divakaran, A.; Radhakrishnan, R.; Peker, K.A., "Blind Summarization: Content-Adaptive Video Summarization Using Time-Series Analysis", *SPIE Conference Multimedia Content Analysis, Management and Retrieval*, Vol. 6073, pp. 6-10, January 2006 (TR2006-026)

Xing, L.; Yu, H.; Huang, Q.; Ye, Q.; Divakaran, A., "Subjective Evaluation Criterion for Selecting Affective Features and Modeling Highlights", *SPIE Conference Multimedia Content Analysis, Management and Retrieval*, Vol. 6073, pp. 188-195, January 2006 (TR2006-027)

Forlines, C.L.; Peker, K.A.; Divakaran, A., "Subjective Assessment of Consumer Video Summarization", *SPIE Conference Multimedia Content Analysis, Management and Retrieval*, Vol. 6073, pp. 170-177, January 2006 (TR2006-002)

Annen, T.; Matusik, W.; Pfister, H.; Seidel, H-P; Zwicker, M., "Distributed Rendering for Multiview Parallax Displays", *SPIE Conference Stereoscopic Displays and Virtual Reality Systems XIII*, Vol. 6055, pp. 231-240, January 2006 (TR2006-031)

Rich, C.; Sidner, C.L.; Lesh, N.; Garland, A.; Booth, S.; Chimani, M., "DiamondHelp: A New Interaction Design for Networked Home Appliances", *Personal and Ubiquitous Computing*, ISSN: 1617-4909, Vol. 10, Issue 2, pp. 187-190, January 2006 (Springer-Verlag, TR2004-114)

Duan, C.; Pekhteryev, C.; Fang, J.; Nakache, Y-P; Zhang, J.; Tajima, K.; Nishioka, Y.; Hirai, H., "Transmitting Multiple HD Video Streams over UWB Links", *IEEE Consumer Communications and Networking Conference (CCNC)*, Vol. 2, pp. 691-695, January 2006 (TR2006-008)

Xin, J.; Vetro, A.; Sekiguchi, S.; Sugimoto, K., "MPEG-2 to H.264 / AVC Transcoding for Efficient Storage of Broadcast Video Bitstreams", *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 417-418, January 2006 (TR2006-003)

Shipman, S.E.; Radhakrishan, R.; Divakaran, A., "Architecture for Video Summarization Services Over Home Networks and the Internet ", *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 201-202, January 2006 (TR2006-029)

Shipman, S.E.; Divakaran, A.; Flynn, M.; Batra, A., "Temporal-Context-Based Video Browsing Interface for PVR-Enabled High-Definition Television Systems", *IEEE International Conference on Consumer Electronics (ICCE)*, Digest of Technical Papers, pp. 353-354, January 2006 (TR2006-030)

Otsuka, I.; Radhakrishnan, R.; Siracusa, M.; Divakaran, A.; Mishima, H., "An Enhanced Video Summarization System Using Audio Features for a Personal Video Recorder", *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 297-298, January 2006 (TR2006-024)

Hancock, M.S.; Vernier, F.D.; Wigdor, D.; Carpendale, S.; Shen, C., "Rotation and Translation Mechanisms for Tabletop Interaction", *IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TableTop)*, pp. 79-88, January 2006 (TR2005-118)

Wu, M.; Shen, C.; Ryall, K.; Forlines, C.L.; Balakrishnan, R., "Gesture Registration, Relaxation, and Reuse for Multi-Point Direct-Touch Surfaces", *IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TableTop)*, pp. 185-192, January 2006 (TR2005-109)

Everitt, K.; Shen, C.; Ryall, K.; Forlines, C.L., "MultiSpace: Enabling Electronic Document Micro-mobility in Table-Centric, Multi-Device Environments", *IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TableTop)*, pp. 27-34, January 2006 (TR2005-110)

Shen, C., "Multi-User Interface and Interactions on Direct-Touch Horizontal Surfaces: Collaborative Tabletop Research at MERL", *IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TableTop)*, pp. 53-54, January 2006 (TR2005-119)

Ryall, K.; Ringel Morris, M.; Everitt, K.; Forlines, C.L.; Shen, C., "Experiences With and Observations of Direct-Touch Tables", *IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TableTop)*, pp. 89-96, January 2006 (TR2005-108)

Sidner, C.L.; Dzikovska, M., "A First Experiment In Engagement For Human-Robot Interaction In Hosting Activities", *Advances in Natural Multimodal Dialogue Systems*, N.O. Bernsen, L. Dybkjaer, J. van Kuppevelt (eds), December 2005 (TR2003-134)

Guvenc, I.; Sahinoglu, Z., "Threshold Selection for UWB TOA Estimation Based on Kurtosis Analysis", *IEEE Communication Letters*, ISSN: 1089-7798, Vol. 9, Issue 12, pp. 1025-1027,

December 2005 (TR2005-121)

Frankel, J.L.; Bromberg, D., "Combadge: A Voice Messaging Device for the Masses", *International Conference on Innovative Applications of the Information Technology for Developing World (Asian Applied Computing Conference (AACC))*, December 2005 (TR2005-161)

Esenther, A.W.; Wittenburg, K.B., "Multi-User Multi-Touch Games on DiamondTouch with the FTFlash Toolkit", *Intelligent Technologies for Interactive Entertainment (INTETAIN)*, November 2005 (TR2005-105)

Wu, J.; Mehta, N.B.; Zhang, J., "Flexible Lognormal Sum Approximation Method", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 6, pp.3413-3417, November 2005 (TR2005-099)

Wu, J.; Mehta, N.B.; Zhang, J., "Spectral Efficiency Analysis of Cellular Systems with Channel-Aware Schedulers", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 3, pp. 1383-1387, November 2005 (TR2005-104)

Mehta, N.B.; Molisch, A.F.; Greenstein, L.J., "Orthogonality Factor in WCDMA Downlinks in Urban Macrocellular Environments", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 6, pp. 3378-3382, November 2005 (TR2005-100)

Wang, Y.; Zhang, J.; Fossorier, M.; Yedidia, J.S., "Reduced Latency Iterative Decoding of LDPC Codes", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 3, pp. 1199-2204, November 2005 (TR2005-103)

Guvenc, I.; Sahinoglu, Z., "Multiscale Energy Products for TOA Estimation in IR-UWB Systems", *IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 209-213, November 2005 (TR2005-042)

Molisch, A.F.; Balakrishnan, K.; Cassioli, D.; Chong, C-C; Emami, S.; Fort, A.; Karedal, J.; Kunisch, J.; Schantz, H.; Siwiak, K., "A Comprehensive Model for Ultrawideband Propagation Channels", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 6, pp. 3648-3653, November 2005 (TR2005-142)

Fang, Y.; Gu, D.; McDonald, A.B.; Zhang, Z., "On the Performance Enhancement of Wireless LAN - A Multi-Polling Mechanism with Hidden Terminal Solution", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 1, pp. 571-575, November 2005 (TR2005-157)

Zhang, J.; Fossorier, M.; Gu, D.; Zhang, J., "Improved Min-Sum Decoding of LDPC Codes Using 2-Dimensional Normalization", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 3, pp. 1187-1192, November 2005 (TR2005-158)

Hu, R.; Raj, B., "A Robust Voice Activity Detector Using an Acoustic Doppler Radar", *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, pp. 171-176,

November 2005 (TR2005-159)

Raj, B.; Singh, R., "Reconstructing Spectral Vectors with Uncertain Spectrographic Masks for Robust Speech Recognition", *IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, pp. 27-32, November 2005 (TR2005-160)

Raskar, R.; Beardsley, P.A.; Dietz, P.H.; van Baar, J., "Tags are Coming: Exploiting Photosensing Wireless Tags for Assisting Geometric Procedures", *Computer Vision for Interactive and Intelligent Environment (CVIIE)*, pp. 145-150, November 2005 (TR2005-152)

Smolic, A.; Kimata, H.; Vetro, A., "Development of MPEG Standards for 3D and Free Viewpoint Video", *SPIE Conference Optics East 2005: Communications, Multimedia & Display Technologies*, Vol. 6014, pp. 262-273, November 2005 (TR2005-116)

* Zhang, X.; Molisch, A.F.; Kung, S-Y., "Variable-Phase-Shift-Based RF-Baseband Codesign for MIMO Antenna Selection", *IEEE Transactions on Signal Processing*, ISSN: 1053-587X, Vol. 53, Issue 11, pp. 4091-4103, November 2005 (TR2005-143)

Assis, F.; Yerazunis, W.S.; Siefkes, C.; Chhabra, S., "CRM114 versus Mr. X: CRM114 Notes for the TREC 2005 Spam Track", *NIST Text REtrieval Conference (TREC)*, November 2005 (TR2005-162)

Vannucci, G.; Leigh, D.L.; Katz, J., "Wide-Area Long-Range Unidirectional Sensor (WALRUS) Network", *IEEE International Conference on Mobile Ad-hoc and Sensor Systems*, pp. 396-404, November 2005 (TR2005-133)

Porikli, F.M.; Tuzel, O., "Bayesian Background Modeling for Foreground Detection", *ACM International Workshop on Video Surveillance and Sensor Networks (VSSN)*, ISBN: 1-59593-242-9, pp. 55-28, November 2005 (TR2005-101)

* Molisch, A.F., "Wireless Communications", *Wireless Communications*, ISBN 0470848871, November 2005 (John Wiley & Sons)

Hata, T.; Kuwahara, N.; Nozawa, T.; Schwenke, D.L.; Vetro, A., "Surveillance System with Object-Aware Video Transcoder", *International Workshop on Multimedia Signal Processing (MMSP)*, Poster Session 4, October 2005 (TR2005-115)

* Forlines, C.L.; Shen, C., "DTLens: Multi-user Tabletop Spatial Data Exploration", *ACM Symposium on User Interface Software and Technology (UIST)*, ISBN: 1-59593-271-2, pp. 119-122, October 2005 (TR2005-096)

* Dietz, P.H.; Harsham, B.A.; Forlines, C.L.; Leigh, D.L.; Yerazunis, W.S.; Shipman, S.E.; Schmidt-Nielsen, B.K.; Ryall, K., "DT Controls: Adding Identity to Physical Interfaces", *ACM Symposium on User Interface Software and Technology (UIST)*, ISBN: 1-59593-271-2, pp. 245-252, October 2005 (TR2005-069)

* Forlines, C.L.; Balakrishnan, R.; Beardsley, P.A.; van Baar, J.; Raskar, R., "Zoom-and-Pick: Facilitating Visual Zooming and Precision Pointing with Interactive Handheld Projectors",

ACM Symposium on User Interface Software and Technology (UIST), ISNB: 1-59593-271-2, pp. 73-82, October 2005 (TR2005-095)

* Lee, J.C.; Hudson, S.E.; Summet, J.W.; Dietz, P.H., "Moveable Interactive Projected Displays Using Projector Based Tracking", *ACM Symposium on User Interface Software and Technology (UIST)*, ISBN: 1-59593-271-2, pp. 63-72, October 2005 (TR2005-141)

Porikli, F.M.; Thornton, J.E., "Shadow Flow: A Recursive Method to Learn Moving Cast Shadows", *IEEE International Conference on Computer Vision (ICCV)*, ISSN: 1550-5499, Vol. 1, pp. 891-898, October 2005 (TR2005-058)

* Zhu, Q.; Avidan, S.; Cheng, K-T, "Learning a Sparse, Corner-Based Representation for Time-Varying Background Modeling", *IEEE International Conference on Computer Vision (ICCV)*, ISSN: 1550-5499, Vol. 1, pp. 678-685, October 2005 (TR2005-148)

* Agrawal, A.; Chellappa, R.; Raskar, R., "An Algebraic Approach to Surface Reconstruction from Gradient Fields", *IEEE International Conference on Computer Vision (ICCV)*, ISSN: 1550-5499, Vol. 1, pp. 174-181, October 2005 (TR2005-145)

* Feris, R.; Raskar, R.; Chen, L.; Tan, K-H; Turk, M., "Discontinuity Preserving Stereo with Small Baseline Multi-Flash Illumination", *IEEE International Conference on Computer Vision (ICCV)*, ISSN: 1550-5499, Vol. 1, pp. 412-419, October 2005 (TR2005-146)

* Lee, J.; Moghaddam, B.; Pfister, H.; Machiraju, R., "A Bilinear Illumination Model for Robust Face Recognition", *IEEE International Conference on Computer Vision (ICCV)*, ISSN: 1550-5499, Vol. 2, pp. 1177-1184, October 2005 (TR2005-093)

Raj, B.; Smaragdis, P., "Latent Variable Decomposition of Spectrograms for Single Channel Speaker Separation", *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pp. 17-20, October 2005 (TR2005-137)

Smaragdis, P.; Boufounos, P., "Learning Source Trajectories Using Wrapped-Phase Hidden Markov Models", *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pp. 114-117, October 2005 (TR2005-138)

Radhakrishnan, R.; Divakaran, A.; Smaragdis, P., "Audio Analysis for Surveillance Applications", *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pp. 158-161, October 2005 (TR2005-139)

Lee, K-C; Moghaddam, B., "A Practical Face Relighting Method for Directional Lighting Normalization", *International Workshop on Analysis & Modeling of Faces & Gestures (AMFG)*, October 2005 (TR2005-092)

Peker, K.A.; Divakaran, A.; Lanning, T., "Browsing News and Talk Video on a Consumer Electronics Platform Using Face Detection", *SPIE Conference Multimedia systems and Applications VIII*, Vol. 6015, pp. 430-435, October 2005 (TR2005-155)

* Ren, L.; Shakhnarovich, G.; Hodgins, J.K.; Pfister, H.; Viola, P., "Learning Silhouette

Features for Control of Human Motion", *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 24, Issue 4, pp.1303-1331, October 2005 (TR2004-068)

* Morency, L-P; Sidner, C.L.; Lee, C.H.; Darrell, T., "Contextual Recognition of Head Gestures", *International Conference in MultiModal Interaction (ICMI)*, ISBN: 1-59593-028-0, pp. 18-24, October 2005 (TR2005-107)

Guvenc, I.; Sahinoglu, Z.; Molisch, A.F.; Orlik, P.V., "Non-coherent TOA Estimation in IR-UWB Systems with Different Signal Waveforms", *IEEE International Conference on Broadband Networks*, pp. 245-251, October 2005 (TR2005-132)

Brand, M.E., "Nonrigid Embeddings for Dimensionality Reduction", *European Conference on Machine Learning (ECML)*, ISBN: 3-540-29243-8, Vol. 3720, October 2005 (TR2005-117)

* Lesh, N.B.; Marks, J.W.; McMahon, A.; Mitzenmacher, M., "New Heuristic and Interactive Approaches to 2D Rectangular Strip Packing", *Journal of Experimental Algorithmics (JEA)*, ISSN: 1084-6654, Vol. 10, September 2005 (TR2005-113)

* Martinian, E.; Yekhanin, S.; Yedidia, J.S., "Secure Biometrics Via Syndromes", *Allerton Conference on Communications, Control and Computing*, September 2005 (TR2005-112)

Ding, G.; Sahinoglu, Z.; Orlik, P.V.; Zhang, J.; Bhargava, B., "Reliable Broadcasting in ZigBee Networks", *IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON)*, pp. 510-520, September 2005 (TR2005-029)

Fang, Y.; Gu, D.; McDonald, A.B.; Zhang, J., "A Two-Level Carrier Sensing Mechanism for Overlapping BSS Problem in WLAN", *IEEE Workshop on Local and Metropolitan Area Networks (LANMAN)*, pp. 1-6, September 2005 (TR2005-156)

Zilani, F.; Velastin, S.; Porikli, F.M.; Marcenaro, L.; Kelliher, T.; Cavallaro, A.; Bruneaut, P., "Performance Evaluation of Event Detection Solutions: The Creds Experience", *IEEE Conference on Advanced Video and Signal Based Surveillance*, pp. 201-206, September 2005 (TR2006-018)

Raskar, R.; Beardsley, P.A.; Dietz, P.; van Baar, J., "Photosensing Wireless Tags for Geometric Procedures", *Communications of the ACM*, ISSN: 0001-0782, Vol. 48, Issue 9, pp. 46-51, September 2005 (TR2005-153)

* Molisch, A.F., "Ultrawideband Propagation Channels-Theory, Measurement and Modeling", *IEEE Transactions on Vehicular Technology*, ISSN: 0018-9545, Vol. 54, Issue 5, pp. 1528-1545, September 2005 (TR2005-037)

Nikovski, D.N.; Nishiuma, N.; Goto, Y.; Kumazawa, H., "Univariate Short-Term Prediction of Road Travel Times", *International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 1074-1079, September 2005 (TR2005-086)

Forlines, C.L.; Shen, C.; Vernier, F., "Under My Finger: Human Factors in Pushing and Rotating Documents Across the Table", *IFIP TC13 International Conference on Human-*

Computer Interaction (INTERACT), September 2005 (TR2005-070)

Ryall, K.; Li, Q.; Esenther, A.W., "Temporal Magic Lens: Combined Spatial and Temporal Query and Presentation", *IFIP TC13 International Conference on Human-Computer Interaction (INTERACT)*, September 2005 (TR2005-031)

Ryall, K.; Esenther, A.W.; Everitt, K.; Forlines, C.L.; Ringel Morris, M.; Shen, C.; Shipman, S.E.; Vernier, F., "iDwidgets: Parameterizing Widgets by User Identity", *IFIP TC13 International Conference on Human-Computer Interaction (INTERACT)*, September 2005 (TR2005-055)

Everitt, K.; Shen, C.; Ryall, K.; Forlines, C.L., "DocuBits and Containers: Providing e-Document Micro-Mobility in a Walk-up Interactive Tabletop Environment", *IFIP TC13 International Conference on Human-Computer Interaction (INTERACT)*, September 2005 (TR2005-053)

Forlines, C.L.; Schmidt-Nielsen, B.K.; Raj, B.; Wittenburg, K.B.; Wolf, P., "A Comparison between Spoken Queries and Menu-based Interfaces for In-Car Digital Music Selection", *IFIP TC13 International Conference on Human-Computer Interaction (INTERACT)*, September 2005 (TR2005-020)

Furuichi, M.; Mihori, Y.; Muraoka, F.; Estenther, A.W.; Ryall, K., "DTMap Demo: Interactive Tabletop Maps for ubiquitous Computing", *International Conference on Ubiquitous Computing (UbiComp)*, September 2005 (TR2005-091)

Guvenc, I.; Sahinoglu, Z., "TOA Estimation with Different IR-UWB Transceiver Types", *IEEE International Conference on Ultra-Wideband (ICU)*, pp. 426-431, September 2005 (TR2005-131)

Sahinoglu, Z.; Guvenc, I., "Threshold-Based TOA Estimation for Impulse Radio UWB Systems", *IEEE International Conference on Ultra-Wideband (ICU)*, pp. 420-425, September 2005 (TR2005-026)

Bansal, D.; Raj, B.; Smaragdis, P., "Bandwidth Expansion of Narrowband Speech Using non-Negative Matrix Factorization", *Eurospeech*, September 2005 (TR2005-135)

Raj, B.; Singh, R.; Smaragdis, P., "Recognizing Speech from Simultaneous Speakers", *Eurospeech*, September 2005 (TR2005-136)

Gezici, S.; Sahinoglu, Z.; Molisch, A.F.; Kobayashi, H.; Poor, H.V., "A Two-Step Time of Arrival Estimation Algorithm for Impulse Radio Ultra Wideband Systems", *European Signal Processing Conference (EUSIPCO)*, September 2005 (TR2005-028)

Zhang, J.; Wang, Y.; Fossorier, M.; Yedidia, J.S., "Replica Shuffled Iterative Decoding", *IEEE International Symposium on Information Theory*, pp. 454-458, September 2005 (TR2005-063)

Smaragdis, P., "From Learning Music to Learning to Separate", *Forum Acusticum*, August 2005 (TR2005-134)

Vetro, A.; Xin, J.; Sun, H., "Error Resilience Video Transcoding for Wireless Communications", *IEEE Wireless Communications*, ISSN: 1536-1284, Vol. 12, Issue 4, pp. 14-21, August 2005 (TR2005-102)

Gezici, S.; Tian, Z.; Giannakis, G.V.; Kobaysahi, H.; Molisch, A.F.; Poor, H.V.; Sahinoglu, Z., "Localization via Ultra-Wideband Radios: A Look at Positioning Aspects for Future Sensor Networks", *IEEE Signal Processing Magazine*, ISSN: 1053-5888, Vol. 22, Issue 4, pp. 70-84, July 2005 (TR2005-072)

* Weyrich, T.; Matusik, W.; Pfister, H.; Lee, J.; Ngan, A.; Jensen, H.W.; Gross, M., "Measuring Skin Reflectance and Subsurface Scattering", *ACM SIGGRAPH*, July 2005 (TR2005-046)

* Gezici, S.; Kobayashi, H.; Poor, H.V.; Molisch, A.F., "Performance Evaluation of Impulse Radio UWB Systems with Pulse-Based Polarity Randomization", *IEEE Transactions on Signal Processing*, ISSN: 1053-587X, Vol. 53, Issue 7, pp. 2537-2549, July 2005 (TR2004-017)

* McGuire, M.; Matusik, W.; Pfister, H.; Hughes, J.F.; Durand, F., "Defocus Video Matting", *ACM Transactions on Graphics (TOG)*, Vol. 24, Issue 3, pp. 567-576ISSN: 0730-0301, July 2005 (TR2005-049)

* Vlasic, D.; Brand, M.; Pfister, H.; Popovic, J., "Face Transfer with Multilinear Models", *ACM Transactions on Graphics (TOG)*, ISSN: 0730=0301, Vol. 24, Issue 3, pp. 426-433, July 2005 (TR2005-048)

* Matusik, W.; Zwicker, M.; Durand, F., "Texture Design Using a Simplicial Complex of Morphable Textures", *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 24, Issue 3, pp. 787-794, July 2005 (TR2005-094)

* Agrawal, A.; Raskar, R.; Nayar, S.K.; Li, Y., "Removing Photography Artifacts Using Gradient Projection and Flash-Exposure Sampling", *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 24, Issue 3, pp. 828-835, July 2005

Wu, J.; Horng, H.; Zhang, J.; Xiao, C.; Oliver, J.C., "Combining Orthogonal Space Time Block Codes with Adaptive Sub-group Antenna Encoding", *International Journal of Wireless Information Networks*, ISSN: 1068-9605, Vol. 12, Number 3, pp. 179-186, July 2005 (Springer Netherlands, TR2006-013)

* Molisch, A.F.; Win, M.Z.; Choi, Y-S; Winters, J.H., "Capacity of MIMO Systems with Antenna Selection", *IEEE Transactions on Wireless Communications*, ISSN: 1536-1276, Vol. 4, Issue 4, pp. 1759-1772, July 2005 (TR2005-144)

* Rich, C.; Sidner, C.L.; Lesh, N.; Garland, A.; Booth, S.; Chimani, M., "DiamondHelp: A Collaborative Task Guidance Framework for Complex Devices", *National Conference on*

Artificial Intelligence (AAAI), July 2005 (TR2004-114)

Porikli, F.M.; Tuzel, O., "Multi-Kernel Object Tracking", *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1234-1237, July 2005 (TR2005-122)

Xiong, Z.; Radhakrishnan, R.; Divakaran, A.; Husang, T.S., "Highlights Extraction from Sports Video Based on an Audio-Visual Marker Detection Framework", *IEEE International Conference on Multimedia and Expo (ICME)*, ISBN: 0-7803-9331-7, pp. 4, July 2005

Leigh, D.L.; Ryall, K.; Lanning, T.; Lesh, N.B.; Miyashita, H.; Hirata, K.; Hara, Y.; Sakura, T., "Sidelobe Minimization of Uniformly-Excited Sparse Linear Arrays using Exhaustive Search and Visual Browsing", *IEEE Antennas and Propagation Society International Symposium (AP-S International Symposium)*, Vol. 1B, pp. 763-766, July 2005 (TR2005-064)

* Yedidia, J.S.; Freeman, W.T.; Weiss, Y., "Constructing Free-Energy Approximations and Generalized Belief Propagation Algorithms", *IEEE Transactions on Information Theory*, ISSN; 0018-9448, Vol. 51, Issue 7, pp. 2282-2312, July 2005 (TR2004-040)

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

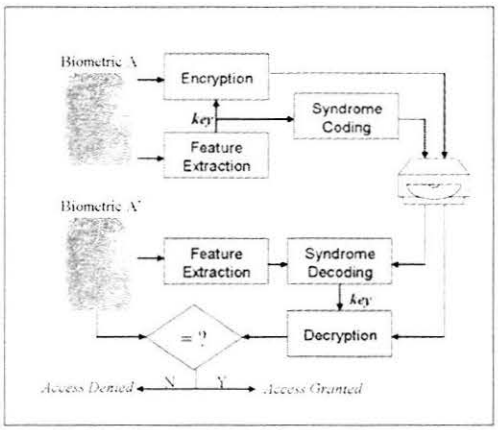
Computer Vision is the branch of computer science concerned with analyzing images to extract information about the world. This is the same function that the human visual system provides (although perhaps accomplished through different mechanisms). As sensors and computers drop in cost, these visual functions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

Much of the computer vision research at MERL is focused on security and 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. Another focus in the surveillance area is object tracking in video. Some of the work in tracking has used stereo cameras to track objects in 3-D. Other work has looked at the problem of tracking objects across different cameras in multi-camera systems.

The following project pages describe the many computer vision projects going on at MERL. They include work on biometric systems, tracking systems, event detection and classification, statistical classification, and advanced camera systems. Some of these projects are currently being applied in many MELCO's businesses like surveillance, security, and access control; and consumer products such as cell phones and DVD players.

Project Descriptions

Secure Biometrics



Current methods of using biometrics are often insecure since they store the biometric in the clear, compromising security and enabling identity theft. Our method obtains security by transforming the biometric into a syndrome (a compressed and scrambled bit stream that contains less information than the original biometric). Since only the syndrome and not the original biometric is stored, attackers who learn the syndrome cannot determine the original biometric and cannot impersonate the user.

Background and Objectives: Biometrics such as fingerprints, irises, and faces are becoming increasingly attractive for authentication, encryption and access control. Since biometrics are slightly different each time they are measured they cannot be stored in encrypted form as passwords are because the encrypted form of the original biometric and a later measurement of the same biometric would not much. Consequently, most systems currently store biometrics in the clear. For biometrics to be broadly accepted, we need a way to store biometrics in a secure form that cannot be used by an attacker to impersonate a valid user, while still being robust to the natural measurement variations of biometrics.

Technical Discussion: Our method obtains security by transforming the biometric into a binary vector which is then multiplied by the parity check matrix of a publicly known low density parity check code to yield a syndrome. The syndrome is a compressed and scrambled version of the original biometric with two essential features. First, the syndrome contains less information than the original. So if only the syndrome and not the original biometric is stored, attackers who learn the syndrome cannot recover the original biometric or impersonate the user. Second, when the syndrome is be combined with another measurement of the same biometric, it is possible to correct the measurement noise and exactly recover the original biometric through belief propagation decoding. The original biometric can then be used to decrypt, authenticate, or gain access using standard biometric techniques.

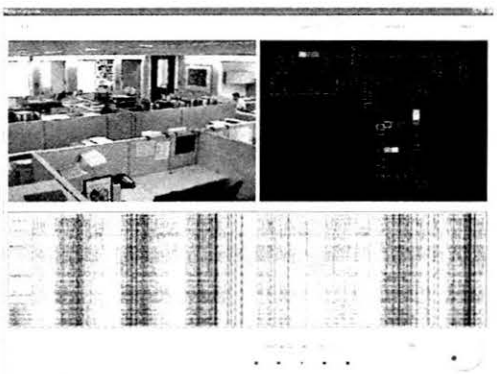
Collaboration: We are currently discussing applications and technologies with the Physical Security Project in Sentansoken and the Information Security Technology Department and Error Control Coding Team in Johosoken.

Future Direction: We plan to improve our secure iris prototype (which was the first published demonstration of this concept) and to develop a secure fingerprint prototype. In addition, we also plan to explore standardization opportunities for secure biometrics (e.g., through INCITS M1 in the USA and ISO/IEC JTC1/SC37 internationally).

Contact: Emin Martinian
http://www.merl.com/projects/secure-biometrics/

Lab: MERL Technology Lab
Project Type: Research

Integrated Event Recognition



See Color Figure 1

We are developing an approach to identification of human behaviors in a large spaces that are instrumented with video cameras and a variety of sensors. The approach is expected to make in impact in surveillance and monitoring applications.

Background and Objectives: Often a problem in surveillance and monitoring scenarios is the fact that human behaviors are extended in time and in space. Unlike in object detection and recognition, a single frame captured by a camera does not contain enough information to identify if an activity is taking place or not. Further, the spatial extent of the activity may move the subject outside the camera view which would result in the activity being only partially observed. Finally, geometric constraints of the monitored space may make using cameras prohibitively expensive, as more and more of them are required to provide reasonable uninterrupted coverage of the space.

To circumvent these problems we aim to develop a systematic approach for specifying and searching for activity events in a physical space instrumented with a set of heterogeneous sensors. In our initial implementation we are exploring a query system that allows us to specify search criteria in terms of the monitored space, rather than simply attempt to recognize activities in any individual camera view. We use over 150 motion sensors and 6 video cameras to monitor activities on an entire floor of an office building.

Technical Discussion: The main challenge of identifying human behaviors in large spaces is the fact that any individual sensor - a camera, a microphone or a motion sensor, only partially observe the behavior. In order to identify the behavior in its entirety we need a space-wide understanding of what these behaviors look like across sensors. Another challenge lies in actual specification of the search terms - the designer of the system can not predict what types of behaviors will be searched for. We opt for extracting high level features of video- and sensor streams and placing them in a database. Then the search could be quickly performed by combining these features within a query and using the database system to quickly locate instances in the video feeds where these high level features jointly satisfy a high-level search constraint, such as for example, a temporal order, co-occurrence, or appearance. These candidates dramatically reduce the search time in large scale video data sets to just a few candidates.

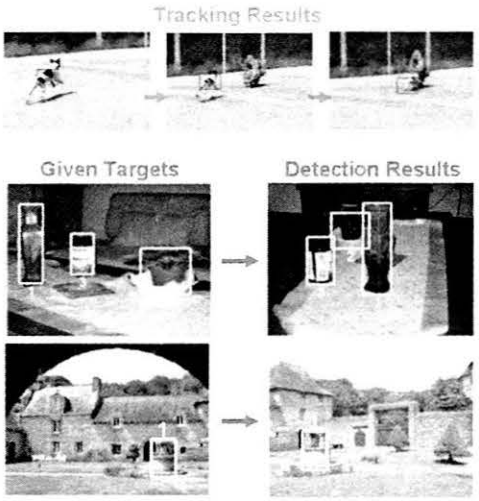
Collaboration: Integrated Civil Security SK.

Future Direction: In the future we will use these candidate hits as a first stage for a more refined search with more accurate but perhaps more computationally expensive classifiers to reduce the False Positive Rate.

Contact: Yuri Ivanov, Christopher R. Wren
http://www.merl.com/projects/integrateder/

Lab: MERL Technology Lab
Project Type: Research

Covariance Tracking



See Color Figure 2

The goal of this project is to develop robust, fast, and highly discriminative object representations and to integrate these representations within a state-of-science tracker. We also aim to develop more accurate object detectors by using this descriptor. Object tracking is one the most important and challenging tasks in computer vision. It has applications ranging from video surveillance to traffic management to to video summarization and compression. Covariance tracking is robust against severe illumination changes, noise, and erratic motion. It has been presented in Japan in January 2006 press release and it received significant media attention. Several companies already expressed their interest on this technology and investigating the possibility of joint developments.

Background and Objectives: Our objective is to improve the robustness and adaptability of surveillance tracking systems and strengthen MELCO's intellectual property in this area. Covariance based representation will increase the accuracy of these methods for uncontrolled lighting conditions and setups. Tracking has a wide spectrum of applications including video surveillance, traffic management, advanced vehicle control systems, robotics, and video summarization. Tracking is also a fundamental technology to extract regions of interest and video object layers as defined in JPEG-2000 and MPEG-4 standards.

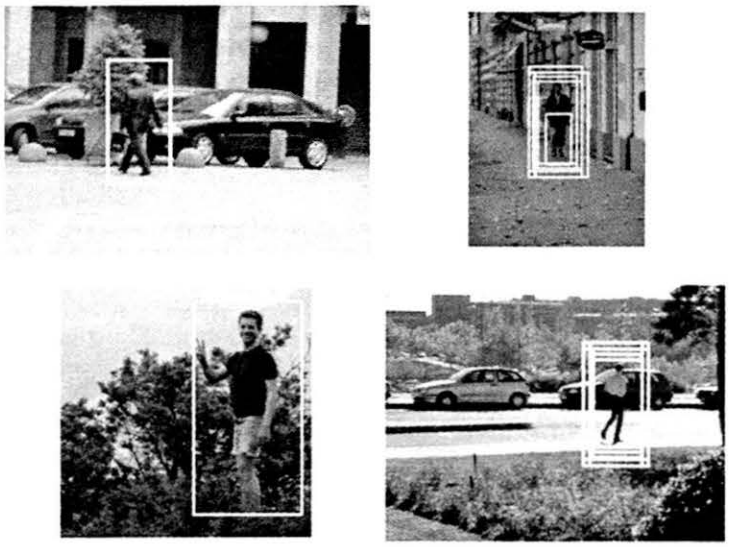
Technical Discussion: Covariance matrix representation embodies both spatial and statistical properties of objects, and provides an elegant solution to fusion of multiple features. Covariance is an essential measure of how much the deviation of two or more variables or processes match. In tracking, these variables correspond to point features such as coordinate, color, gradient, orientation, and filter responses. This representation has much lower dimensionality than histograms. It is robust against noise and lighting changes. To track objects using covariance descriptor, an eigenvector based distance metric is adapted to compare the matrices of object and candidate regions. Covariance tracker does not make any assumption on the motion. This means that it can keep track of objects even if their motion is erratic and fast. It can compare any regions without being restricted to a constant window size. In spite of these advantages, the computation of the covariance matrix distance for all candidate regions is slow and requires exponential time. An integral image based algorithm that requires constant time is proposed to improve the speed.

Collaboration: Joint work with the Physical Security Project at Sentansoken. Kyo-den, Sha-den, C-Den, Riki-den, Shin-den, Tsu-den, and Kama-den are also all either planning or considering covariance based methods for detection and tracking products.

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

Lab: MERL Technology Lab
Project Type: Research

Fast Human Detection



Human detection is a crucial technology for many surveillance applications. The problem is notoriously difficult because of the large variations in appearance due to lighting, pose and clothing. We developed a fast human detection algorithm that can detect humans in still images in real time. The detection accuracy is equal, or better, than state-of-the-art reported results, while running time is up to 70 faster than comparable methods.

See Color Figure 3

Background and Objectives: The problem of human detection in still images has been extensively investigated within the computer vision community. It is difficult because of the large variations in human appearance due to changes in illumination, camera position, clothing and body pose. We have developed a real-time system for human detection that combines a Histogram-of-Gradient representation, for high accuracy, with Integral Histogram representation, for real-time performance.

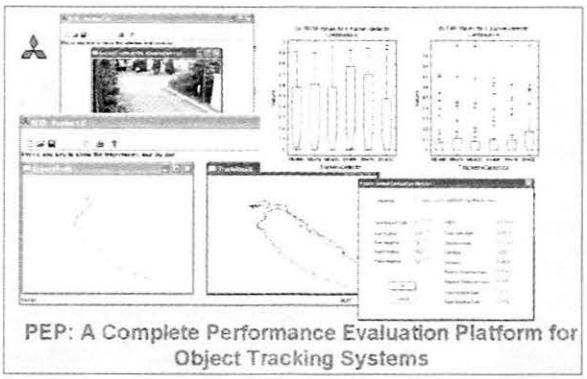
Technical Discussion: Our work combines two leading approaches to object detection. One that uses Histogram-of-Gradients (HoG) to represent objects in a robust manner and an Integral image representation that allows for fast and efficient implementation. We use an AdaBoost training algorithm to learn a cascade of rejecters that are based on HoG of windows of different size and position to quickly reject image patches that do not contain humans. The method compares favorably with other leading techniques in terms of accuracy, while running up to 70 times faster than comparable systems.

Future Direction: We plan to integrate our software into MERL's surveillance software and extend the method to use motion cues as well for improved accuracy.

Contact: Shai Avidan
http://www.merl.com/projects/fasthumandetection/

Lab: MERL Research Lab
Project Type: Research

PEP: Performance Evaluation Platform for Object Tracking Methods


PEP: A Complete Performance Evaluation Platform for Object Tracking Systems

See Color Figure 4

The goal of this project is to develop a complete performance evaluation platform for object detection tracking systems. We implemented all of the conventional metrics as well as MERL's own novel matrix based measures. We standardized the data representation and automated the evaluation of various configurations (different detector/tracker parameters) and tracking scenarios (multiple people/fast moving people/etc.).

Background and Objectives: The issue of evaluating the performance of video surveillance systems is becoming more important as more and more research effort is drawn into object detection and tracking. It's a natural question to ask whether there has been quantifiable progress in the form of robust, commercial- grade video surveillance systems as a result of past and ongoing research in this direction. We address the issue of comprehensive performance evaluation of automatic object detection and tracking systems. We designed several performance evaluation metrics for quantitative assessment of the performance of video surveillance systems. Initial efforts towards performance evaluation of video detection and tracking systems began with the workshops dedicated to the topic, namely VS (visual surveillance) and PETS (performance evaluation of tracking and surveillance) series of workshops.

Technical Discussion: Frame-based metrics are used to measure the performance of surveillance system on individual frames of a video sequence. This does not take into account the response of the system in preserving the identity of the object over its lifespan. The results from individual frame statistics are then averaged over the whole sequence. This represents a bottom-up approach. On the other hand, the object-based evaluation measures take the whole trajectory of each object into consideration. The various ways of finding the best correspondence (association) between individual ground truth tracks and tracker result tracks are analyzed. Finally, based on a particular association, success and error rates are computed and accumulated for all the objects. This represents a top-down approach. We propose metrics for both approaches.

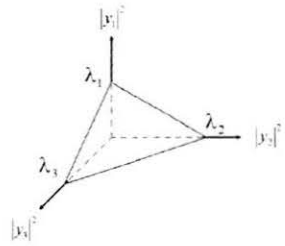
Collaboration: Joint work with the Physical Security Project at Sentansoken. There are also collaborators from academia including INRIA, EPFL, Queens Marry, etc.

Future Direction: We will make the PEP available through a MERL controlled website. This will encourage researchers and other companies to obtain an unbiased evaluation of their algorithms and see their overall rankings.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

Spectral Bounds for Sparse PCA and Sparse LDA



We develop exact methods for optimizing quadratic performance criteria (e.g., minimizing mean-squared-error or maximizing variance) subject to "sparsity" constraints (i.e., the solution vector must have few non-zero elements). Such *cardinality-constrained* optimization problems are encountered in a wide range of applied fields, from bioinformatics (gene selection) to computational finance (portfolio optimization). In specific, we have focused on sparse modeling techniques for computer vision and machine learning. We have developed an efficient and near-optimal greedy algorithm which yields better than state-of-the-art performance on "sparse PCA" benchmarks used in the statistics community. In addition, we derive a variational spectral method for finding optimal solutions using branch-and-bound search. Our algorithms have been fruitfully applied to various real-world computer vision and machine learning tasks.

Background and Objectives: Engineering and the applied sciences use numerical optimization techniques to maximize gain (or minimize loss) while satisfying various operational constraints. With modern-day (non)linear programming techniques, problems involving hundreds of (in)equality constraints are solved routinely by general purpose computers. But there is one (nonlinear) constraint that still remains a challenge: "sparseness", where the solution vector must have many of its elements be (exactly) zero. In this way, the solution will involve fewer measurements (variables), thus reducing associated costs, lowering computational complexity. Unfortunately, sparse constraints (mathematically equivalent to an L0-norm) are inherently non-convex and NP-hard.

Technical Discussion: Recently, in sparse signal representation important advances towards L0-norm optimization have been made. Under certain conditions an Lp-norm ($p<1$) on the solution vector of a set of linear equalities ($Ax=b$) is equivalent to an L0-norm. Hence, resulting in a sparse solution with minimized cardinality. Remarkably, this means that a formerly NP-hard combinatorial problem essentially becomes *convex* (under these conditions) thus leading to globally optimal solutions. However, with sparse PCA the objective function (a Rayleigh quotient) and the cardinality constraint are quadratic and therefore harder to optimize/satisfy. Recent advances in sparse PCA were achieved through continuous approximation techniques and convex relaxations of the hard cardinality constraint. In contrast, we considered an alternative discrete formulation based on variational eigenvalue bounds. We have devised an efficient near-optimal greedy algorithm that is several orders of magnitude faster than competing continuous methods. More importantly, our discrete approach leads to provably optimal solutions using branch-and-bound search. Furthermore, the discrete approach provides a simple renormalization step that improves approximate solutions obtained by any continuous method.

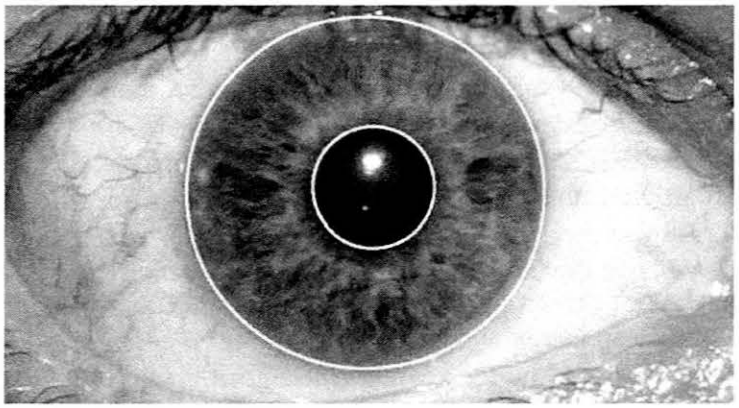
Collaboration: Yair Weiss (The Hebrew University of Jerusalem).

Future Direction: We are now investigating extensions of this spectral framework to supervised learning problems and feature selection for classification.

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

Lab: MERL Research Lab
Project Type: Research

Iris Recognition from 1-2 Meters



The goal of this project is to have a computer recognize a person from an image of his or her iris. The main application for iris recognition is in security, for example, access control. Iris recognition has been shown in independent studies to be the most accurate biometric. The drawback is that the iris is difficult to image. It is small and sometimes partly occluded by the eyelid and eyelashes. Current iris recognition systems require careful cooperation by the user and require the user's eye to be within about 10 inches of the camera. One goal of this project is to design hardware that can capture a good image of a person's iris from a distance of 1 to 2 meters with minimal user cooperation. This would overcome the inconvenience problem which is the main drawback of iris recognition technology. The major developments over last year are the completion of a novel segmentation algorithm and a novel comparison algorithm based on a new set of simple features.

Background and Objectives: The problem of iris recognition has been studied by a number of researchers. Our objective is to develop a state of the art system and give MELCO some intellectual property in this area. We are focusing on enabling iris recognition from further away than current products which will make it much easier to use. To do this we are investigating the hardware needed (cameras, lenses, pan-tilt units) as well as novel algorithms for comparing two iris images.

Technical Discussion: We intend to use MERL's state-of-the-art face detection and eye detection technology to enable a camera to be aimed at the iris from up to 2 meters away. With the right camera and lens, a high resolution image of the eye region can be captured at that distance. Once the eye region is captured, the iris must be segmented out. Given a segmented iris image, novel image features are computed to create a feature vector. Feature vectors can then be compared to yield a similarity value which is used to determine a match. We are investigating novel segmentation algorithms as well as novel features to use for comparison.

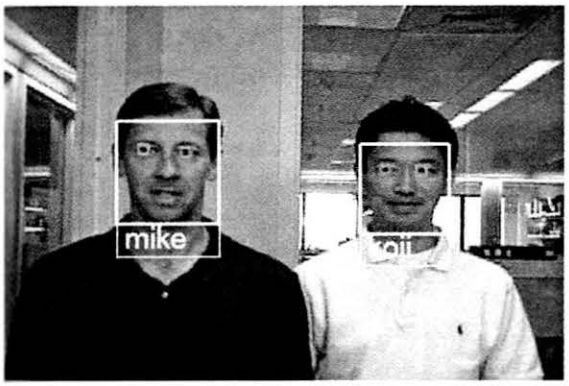
Collaboration: We are having discussions with the Physical Security Project group at Sentansoken and the cell phone group at Johosoken to decide how MELCO might use the technology we are developing.

Future Direction: In the future our goal is to put together the various pieces of the system to create a real-time live demo.

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

Lab: MERL Technology Lab
Project Type: Research

2-D Face Recognition



The goal of this project is to have a computer recognize a person from an image of his or her face. There are many applications for face recognition. Some examples are: access control, summarizing surveillance video, browsing image and video databases and user-interfaces. The face recognition problem is broken into two important steps. The first is aligning the face to a standard position, size and rotation. This is done by finding a number of facial feature points (such as the corners of the eyes, the tip of the nose, etc) and mapping those points to standard positions. The next step is to compare two aligned faces to get a similarity score. We are currently working closely with the PS-Pro group within Sentansoken to productize this technology for various business units. The major improvements over the last year include: 1) similarity score normalization which can dramatically improve accuracy, 2) a new face detector that works even on people wearing caps and helmets 3) a pose estimator that allows non-frontal faces to be discarded.

Background and Objectives: Our objective is to develop a state of the art face recognition system and give MELCO some intellectual property in this area. We are focused on an access control scenario in which the user must cooperate to gain access to a secure room or building. In this scenario lighting can be controlled to a large extent to insure good image quality.

Technical Discussion: The face detection and alignment step is done using the Viola-Jones detection framework developed at MERL. This framework yields very fast and accurate detectors for finding faces and facial feature points. Two techniques have been developed for the second step of computing the similarity score. The first is an extension of the Viola-Jones detection framework. The idea is to learn a set of simple local features that differentiates any two face images of the same person from any two face images of different people. The second technique, called Bayesian Matching, approximates the two distributions of intra-face pairs and inter-face pairs. These distributions are then used to compute the probability that a pair of face images comes from the intra-face class or the extra-face class.

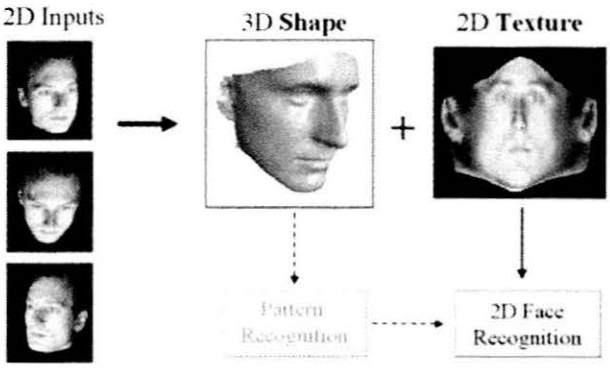
Collaboration: Joint work with the Physical Security Project at Sentansoken. Ina-den is planning to productize for access control in July 2006. Kyo-den, C-Den, Cho-den, Riki-den, MDIS, Shin-den, Tsu-den, and Kama-den are also all either planning or considering face recognition products.

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

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

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: A "Morphable Model" is an analysis-by-synthesis framework for capturing 3D models have 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.

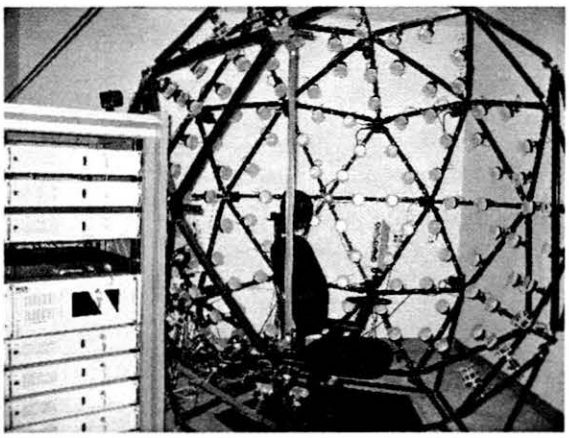
Technical Discussion: We use a linear combination of "eigenheads" obtained by Principal Component Analysis (PCA) of a training set of laser-scanned 3D human faces. The PCA coefficients are used as model parameters. We establish correspondence between faces with an efficient error metric (boundary weighted 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



The goal of this project is to build high-quality statistical models for human faces. The applications of such models include face recognition, digital face aging, user interfaces, and face synthesis. We have built a scanning system that is able to capture images of human faces from different viewpoints under varying illumination. We have acquired face data from a large cross-section of the population, including scans of the same individuals multiple times a day and over a long period of time. Using this data we are building statistical models that capture variations in 3D face geometry, expression, illumination, and small-scale mesostructure (e.g. wrinkles, bumps, pores, etc.).

Background and Objectives: Our research aims at building statistical models that capture how human faces vary between individuals and how they change over a period of time (e.g., within a day or a person's lifetime). There are several applications for such models, including: more robust face recognition, aging images of missing children/fugitives from a single photograph, digitally aging actors for entertainment purposes and re-animating faces of actors to match alternative sound tracks (e.g. foreign language dubs).

Technical Discussion: Our face scanner consists of a geodesic dome with 16 high-resolution color cameras and 150 computer-controlled white LED lights. The system captures the 3D geometry of the face using a commercial face scanning system. The data is used to compute a normal map and to estimate the diffuse reflectance at each surface point. We subtract the diffuse reflectance from the measured data and fit BRDF model to the remaining surface reflectance. We also measure the subsurface scattering of skin at few locations in the face using a special contact device and estimate skin translucency. We analyze the data for a large population of people of different gender, race, and age and under different external conditions (sweaty, cold, makeup, etc.).

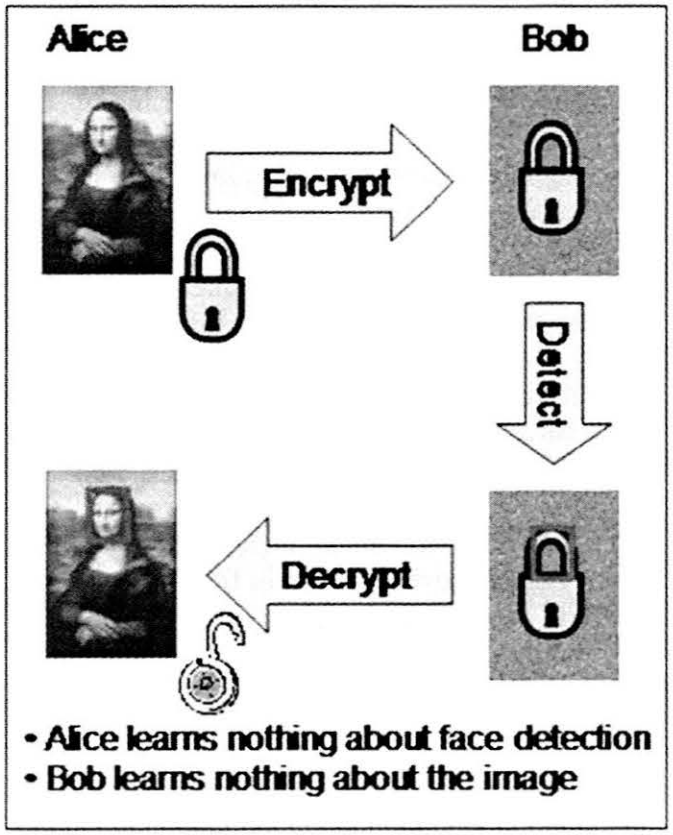
Collaboration: Tim Weyrich and Markus Gross, ETH Zurich; Jinho Lee and Raghu Machiraju, Ohio State University; Jovan Popovich and Daniel Vlasic, MIT; Henrik Wann Jensen, UC San Diego.

Future Direction: We continue to collect more face scans on a daily basis. We will use our models for more robust 2D face recognition, face recognition from high-resolution images, and digital face aging.

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

Lab: MERL Research Lab
Project Type: Research

Blind Vision



We have developed a general framework for secure image and video analysis that allows a client to have his data analyzed by a server, privately. For example, the client might submit his images to the server for face detection, without letting the server learn anything about the content of the images. Or, more generally, the client might use a query image to query an image database stored on the server, without revealing the content of the query image to the server. In the last year, we have implemented a secure face detector as a proof-of-concept, presented our work at a scientific conference and extended the method to work with different types of machine learning technologies.

Background and Objectives: The problem of image and video analysis has been extensively investigated within the computer vision community. However, privacy concerns were never taken into account, as the assumption was always that the data is available to the analyzing algorithm. We borrow techniques from the field of Secure Multi-Party Computations to derive secure image analysis algorithms.

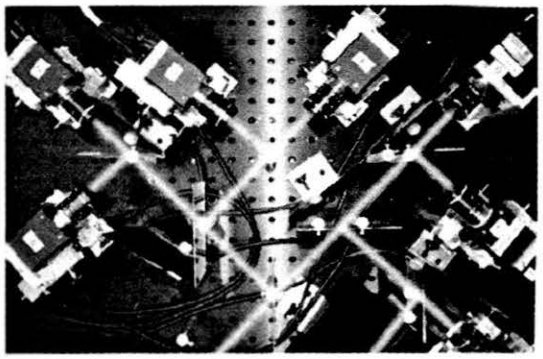
Technical Discussion: Our work combines methods from computer Vision and Secure Multi Party Computations. We use cryptographic primitives such as Oblivious Transfer to convert our non-secure face detector into a secure one. This is done by breaking the face detector into its components and devising a provably secure counter part for each component. In addition, we develop domain specific techniques that allows the user to trade security with speed, using standard information-theoretic tools.

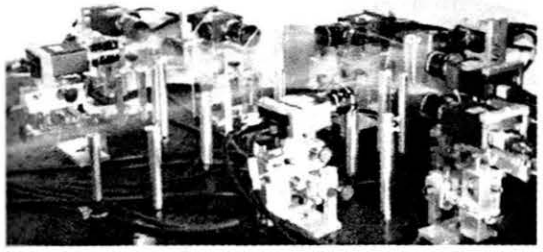
Future Direction: We plan to extend our analysis to general machine learning problems and to multi-player setup that involve more than two parties.

Contact: Shai Avidan
http://www.merl.com/projects/blindvision/

Lab: MERL Research Lab
Project Type: Research

Single-Axis Multi-Parameter (SAMP) Camera





The Single-Axis Multi-Parameter (SAMP) Camera is a research project at MERL that is attempting to rectify the problems with existing surveillance cameras. Taking an image with a standard camera requires the photographer or the camera itself to choose the resolution, the aperture, the exposure time, focus parameters, frequency and polarization parameters (usually done with lens filters), and frame rate (if a video camera). The SAMP camera takes multiple images with different settings of all these parameters with one exposure. Using these multi-parameter images, typical image-processing operations become much easier and result in imagery of markedly better quality.

Background and Objectives: In the context of computer vision, multiple images that are geometrically similar but radiometrically controlled are useful for many applications, such as high dynamic range imaging, focus/defocus analysis, multispectral imaging, high speed videography and high resolution imaging. We informally refer to this as "creating multiple copies of an image". Creating multiple copies of an image has been demonstrated in various contexts, but rarely more than three copies are created. This project addresses the question of how to make a larger number of copies within an optical system.

Technical Discussion: We have implemented a system using an optical splitting tree to generate up to eight copies of the scene. The internal nodes are optical elements like beam splitters, filters, and lenses, and the leaves are video sensors. Varying the optical elements allows us to capture at each virtual pixel multiple samples that vary not only in wavelength but also in other sampling parameters like focus, aperture, polarization, exposure, subpixel position, and frame time.

A wide variety of applications can be implemented using this system by simply reconfiguring its optical elements. We have implemented high dynamic range (HDR), multi-focus, high-speed, high-resolution video, and pulling a high-quality alpha matte and foreground from a video sequence.

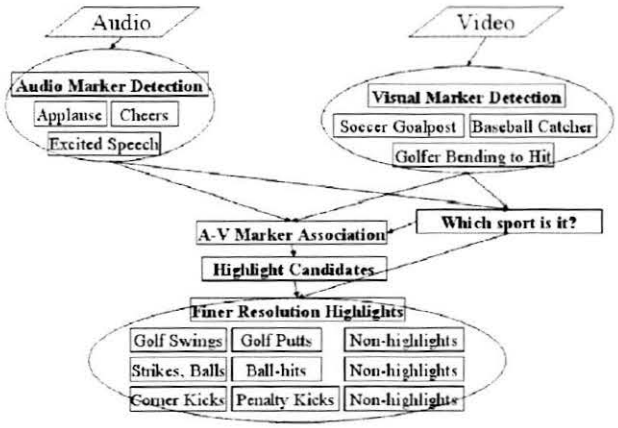
Collaboration: This is a collaboration with Morgan McGuire (Williams College), Shree Nayar (Columbia University), Fredo Durand (MIT), and John Hughes (Brown University).

Future Direction: In future research we would like to reduce the size of the SAMP camera to make it portable and to further enhance its functionality by adding additional cameras and image sensors.

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

Lab: MERL Research Lab
Project Type: Research

Audio-Visual Event Detection for Consumer and Surveillance Video



Audio-visual analysis for event detection in consumer and surveillance video enables us to summarize the content by only including "interesting" events in the summary. The focus of this project is to combine the audio-visual cues to detect these events in order to achieve higher accuracy as well as deeper understanding of audio-visual events.

Background and Objectives: In unscripted content such as sports and surveillance video, "interesting" events happen sparsely in a background of "uninteresting" events. Therefore, by detecting this set of "interesting" events one can summarize the content. 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. The events of interest have characteristic patterns in both video and audio features that may or may not be time aligned. By learning statistical models for these characteristic patterns from the audio-visual features, one can achieve higher accuracy for event detection than is possible when using only one modality (audio or video). In the past we used exclusively video or exclusively audio to detect events for both consumer and surveillance applications. We have used motion activity to detect patterns corresponding to sports highlights as well as changes in highway traffic density. We have also used the "Viola-Jones" real-time object detection framework to detect key video objects from sports video such as Goal Posts in Soccer, Baseball catchers etc. We have also used audio to detect sports highlights by finding long stretches of audience reaction in the form of cheering, applause and commentator's excited speech. While both modes provide powerful cues, we could achieve higher detection accuracy when we combine the results from audio and video analysis.

Technical Discussion: The challenge lies in identifying key audio-visual markers (objects) that are indicative of events of interest in a domain and then in developing statistical models to detect these key audio-visual markers. For instance, in sports video the audience reaction audio class gets us in the vicinity of an "interesting" event. However, to get to the beginning of the interesting event one needs to find a suitable key video object such as the goal post or the baseball catcher. Then, by associating a video marker with an audio marker one can capture the whole event. Another challenge is to deal with streaming as opposed to stored video.

Collaboration: Sentansoken

Future Direction: We plan to extend our techniques on the consumer video front to scripted content where the audio-visual markers will serve to expose the content's underlying structure.

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

Lab: MERL Technology Lab
Project Type: Research

Coded Exposure Photography: Motion Deblurring Using Fluttered Shutter

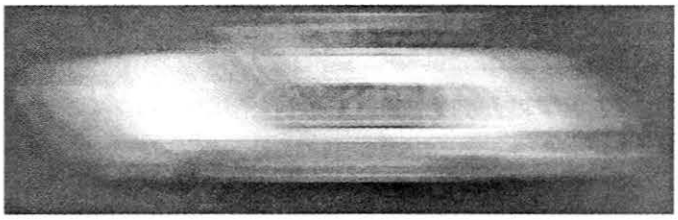


Figure 1 - Photo of a moving vehicle

Figure 2 - Sharp image obtained by deblurring

In a conventional single-exposure photograph, moving objects or moving cameras cause motion blue. The exposure time defines a temporal box filter that smears the moving object across the image by convolution. This box filter destroys important high-frequency spatial details so that deblurring via deconvolution becomes an ill-posed problem.

Rather than leaving the shutter open for the entire exposure duration, we "flutter" the camera's shutter open and closed during the chosen exposure time with a binary pseudo-random sequence. The flutter changes the box filter to a broad-band filter that preserves high-frequency spatial details in the blurred image and the corresponding deconvolution becomes a well-posed problem. We demonstrate that manually-specified point spread functions are sufficient for seval challenging cases of motion-blur removal including extremely large motions, textured backgrounds and partial occluders.

Background and Objectives: Motion deblurring is an ill-posed problem and current approaches try to get an appoximate or smooth solution from the deblurred image. By coding the exposure, high frequencies are preserved which makes the problem well-posed. This enables us to get exact solution with reduced noise and avoids rining artifacts.

Technical Discussion: The exposure of the camera is fluttered by driving an external ferro-electric shutter using a pseudo-random sequence. The sequence is chosen so that the resulting point spread function (PSF) has broadband response and high spatial frequencies are preserved in the blurred image. After manual estimation of the PSF, deblurring by solving a simple linear system results in a sharp image.

Collaboration: Prof. Jack Tumblin, Northwestern University

Future Direction: We are interested in building automatic moving object segmentation and deblurring system for applications in aerial imaging, surveillance, license plate identification and consumer photography.

Contact: Ramesh Raskar, Amit Agrawal
http://www.merl.com/projects/deblur/

Lab: MERL Research Lab
Project Type: Research

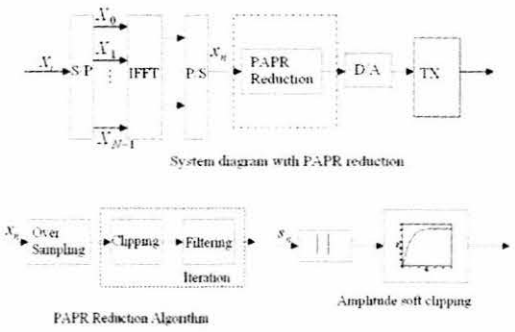
Digital Communications

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.

Project Descriptions

PAPR Reduction for WiMAX OFDM Systems



System diagram with PAPR reduction

PAPR Reduction Algorithm

Amplitude soft clipping

We have developed a new method to reduce Peak-to Average Power Ratio (PAPR) in an OFDM signals especially for WiMAX systems. The reduction is achieved by oversampling the OFDM signal, clipping the signal amplitude to predefined level and constraining the out-of-band spectral components by a bandpass filter. Instead of using a hard clipping, a soft clipping is performed which gradually clips the signal to achieve a better BER performance. The oversampling rate is relatively low and the out-of-band radiation is removed using filtering in the frequency domain.

Background and Objectives: Orthogonal frequency division multiplexing (OFDM) modulation is an effective modulation scheme for transmitting data at high rate. Because of its high spectral efficiency and robustness to the multipath, OFDM has been chosen for high speed wireless communication standards, such as 802.11, 802.16 and future super 3G. In OFDM scheme, the data are transmitted concurrently over number of equally spaced orthogonal subcarriers. OFDM uses an inverse fast Fourier transform (IFFT) to generate the signal waveforms. One of the major drawbacks of OFDM is amplitude varying in the transmitted signal. The latest WiMAX/802.16 standard includes the IFFT size of 2048. With the number of subcarriers increases, the PAPR will put more serious requirement on power amplifier. Therefore, we are interested in having an effective but simple PAPR reduction method.

Technical Discussion: A considerable amount of research has been dedicated in this topic. However, block coding needs exhaustive search for good codes. As number of subcarriers increases, this would be impossible. The complexity of the selective mapping increases as the number of candidate mapping increases and also side information is needed for the receiver. For the partial transmit sequences and interleaving methods, they also have the issue of complexity or side information. The hard clipping method is simple and can reduce the PAPR effectively, but BER performance is impacted. In this work, we perform a soft clipping on oversampled signal. The clipping function takes the form of a hyperbolic tangent function with two pre-defined parameters. After the soft clipping, a bandpass filter is required to remove the out-of-band radiation. This filtering process may result in partial peak regrowth. Therefore, a few iterations including clipping and filtering are suggested. Simulation results show that with two iterations, PAPR can be reduced to 4 dB while maintaining the BER performance for the case of 1024 subcarriers.

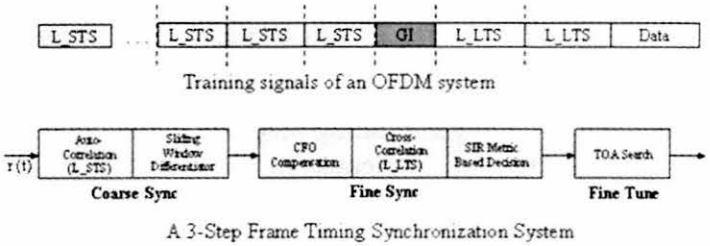
Collaboration: This research is in collaboration with Xiaozhou Huang from LSU.

Future Direction: We will continue to look into the PAPR issue as well as other technical challenges for 802.16/WiMAX systems.

Contact: Jinyun Zhang
http://www.merl.com/projects/papr/

Lab: MERL Technology Lab
Project Type: Research

Frame Timing Synchronization in OFDM Systems



A 3-Step Frame Timing Synchronization System

We are designing robust frame timing synchronization methods for Orthogonal Frequency Division Multiplexing (OFDM) based communication systems. For coarse synchronization, a sliding window differentiator is concatenated with a conventional auto-correlator so that the time ambiguity caused by the auto-correlation plateau can be avoided. For fine synchronization, a SIR metric is defined and used to determine the timing based on cross-correlation result. This method is more robust to multi-paths or pseudo multi-paths introduced by cyclic delay diversity (CDD) scheme. For Time-of-Arrival (TOA) detection, a local search operation can be applied to further improvement of the accuracy.

Background and Objectives: OFDM is a modulation technique for broadband wireless communications. It converts a high-rate data streams into a number of low-rate streams that are transmitted over parallel, narrowband channels. The OFDM technique can mitigate the effect of multi-paths and utilize spectrum effectively. Therefore, it has been widely adopted in various wideband systems, such as 802.11a/g, HiperLan2, 802.16 and latest 802.11n. However, one of its main disadvantages is that it is sensitive to synchronization errors. We are interested in improving the synchronization accuracy by avoiding some ambiguity.

Technical Discussion: Lots of research has been conducted in this area. Basically, proposed schemes in the literature can be divided into two categories: auto-correlation based and cross-correlation based. Auto-correlation exploits the repetition structure of the preamble symbols to find the starting point of each frame. However, due to the repetition structure, there is a well-known plateau problem. With a sliding window differentiator, we can easily remove the plateau and find the point where the repeating short traning symbols end. Cross-correlation is usually used for fine timing synchronization. It works well under AWGN channels. However, in multi-path channels, it may synchronize to the strongest path, but not the first path. Recently, the CDD scheme has been proposed for multiple antenna OFDM systems. The CDD can provide some spatial diversity, but introduces pseudo multi-paths and causes the cross-correlation failure. To remedy this problem, our detection is not based on the output of the cross-correlation directly. Instead, a SIR metric calculator is concatenated to computer the signal energy in the targeted window and the interference energy caused by multi-paths in the interference window. The timing is then determined by the peak of SIR.

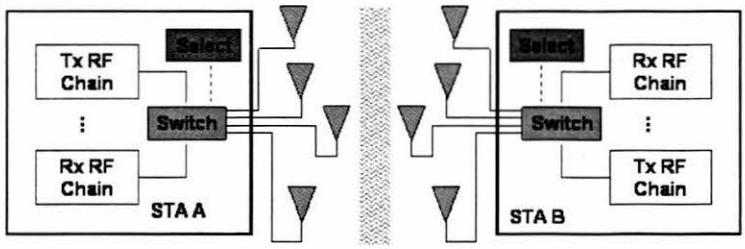
Collaboration: This work is in collaboration with Dr. Dong Wang, who was with University of Delaware and currently with Philips Research USA.

Future Direction: We plan to continue the research on OFDM receiver design techniques for upcoming high throughput 802.11n systems.

Contact: Jinyun Zhang
http://www.merl.com/projects/frametimesync/

Lab: MERL Technology Lab
Project Type: Research

Antenna Selection for IEEE802.11n



Diversity gains are achieved if antenna selection is based on current channel condition- a close-loop system. We developed a training method for selecting antennas in a MIMO-based WLAN. The method operates at MAC layer. The training method rapidly sends multiple training frames, designated as sounding packets. Each sounding packet conforms to a conventional PHY layer design and is for one subset of antennas, so that characteristics of the entire channel can be estimated by the receiver of the sounding packets. The estimated channel condition is used for either transmit or receive antenna selections, or both. In addition to training the MIMO channel, the sounding packets can also include data, which makes the method extremely efficient because training and data transfer are performed concurrently.

Background and Objectives: Multiple-input multiple-output (MIMO) systems provide significant gains in both system capacity and the reliability of transmission over wireless channels without requiring any additional bandwidth. The incoming high throughput WLAN standard - IEEE802.11n adopts MIMO technologies. However, the increased hardware and signal processing complexity has slowed its widespread adoption. Antenna selection reduces the complexity by using fewer RF chains. Channel state information (CSI) feedback from the receiver is often needed for the transmitter to know which antenna(s) to be used. While antenna selection achieves diversity gains, the channel training mechanism to estimate CSI may require some changes at both MAC and PHY layers, and introduce some overhead. So, efficient channel training schemes are needed for antenna selections in a MIMO system.

Technical Discussion: We developed a MAC-based training mechanism for antenna selection for IEEE802.11n standard. The training procedure is conducted by transmitting or receiving sounding packets over all disjointed antenna sets. To avoid channel distortions, these sounding packets are sent consecutively. The training information is exchanged by a control field residing in MAC header, and transparent to PHY. When both transmitter and receiver have antenna selection capabilities, training of transmit and receive antennas can be done one after another. The proposed training scheme minimizes training overhead at MAC and requires no changes at PHY.

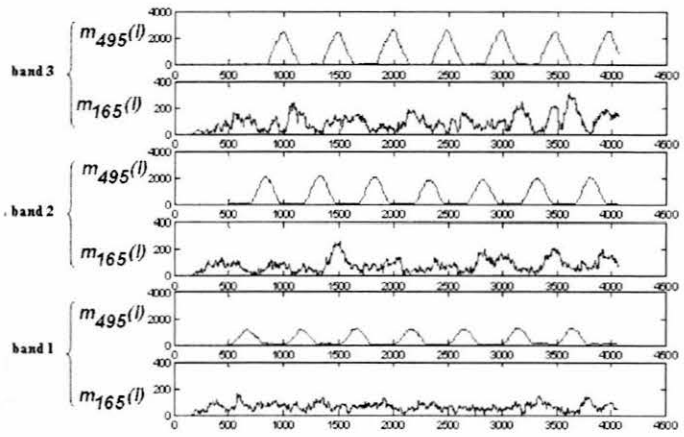
Collaboration: The project is collaborated with Johosoken.

Future Direction: The proposed training protocol for antenna selection has been accepted into IEEE802.11n draft standard. Participation of IEEE802.11n standardization and interoperability test are needed in future to maintain our proposal.

Contact: Daqing Gu
http://www.merl.com/projects/wlan/

Lab: MERL Technology Lab
Project Type: Research

Time Synchronization for MB-OFDM UWB



A problem in wireless communications is processing symbols received via a time dispersive channel causing intersymbol interferences. Orthogonal Frequency Division Multiplexing (OFDM) is an effective modulation technique for such channels. To further reduce OFDM symbol interferences, the OFDM symbols are sent through multiple frequency bands.

Proper decoding of the received OFDM symbols requires an accurate synchronization, i.e. a correct estimation of the arrival time at the receiver of the OFDM symbols. Detecting and synchronizing to the packets and estimating the related parameters becomes more difficult, however, when the symbols are transmitted over different frequency bands according to a frequency hopping pattern. Furthermore a new device attempting to join a network does not know a priori the sequence being used by the network.

MERL proposed synchronization techniques to detect frequency hopped packets and to synchronize, providing an optimal time for processing blocks of symbols during the preamble, in addition to an estimate of a carrier frequency offset.

Background and Objectives: Concurrently to its activities as contributor to WiMedia and co-author of the first UWB standard approved by ECMA International, MERL worked on some implementation issues of the Multiband OFDM receiver. One main achievement of this work is the invention of two patents on synchronization.

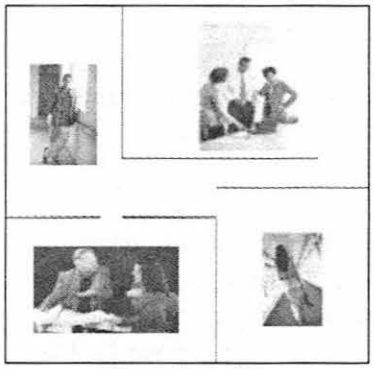
Technical Discussion: The first aspect of this work focused on synchronizing OFDM symbols provided in the preamble, considering these symbols do not have a cyclic prefix and including the effect of channels with different delay spread. The second aspect of this work looked at the challenge of a device joining a network without any prior knowledge of its frequency hopping pattern. To achieve a successful detection, the analysis of the signal received in the selected frequency sub-band relative to the periodicities associated with the respective predefined groups of partitioned frequency hopping patterns is performed.

Collaboration: This work was performed in collaboration with Renesas.

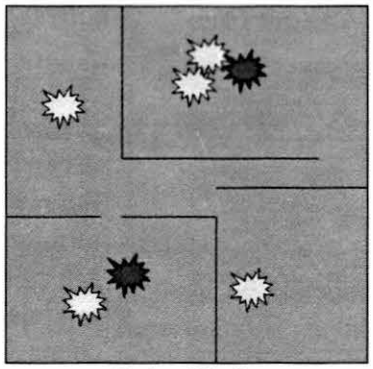
Contact: Yves-Paul Nakache, Jinyun Zhang
http://www.merl.com/projects/timesync/

Lab: MERL Technology Lab
Project Type: Research

Short Range Ultrawideband Radar



Floor of a building



System Display

See Color Figure 5

Although, radars have been primarily used for military applications, there is a growing interest to exploit radars for commercial use (e.g., games and toys), robotics, rescuing efforts and health care. The scope of this research is to investigate potential use of ultrawideband (UWB) signals as a short range radar and motion sensor for commercial and medical applications. If two radios are exchanging UWB signals, the highly structured preamble of the signal can be exploited for multi-purposes: ranging, synchronization and detection of change in multipath profile, while the signal payload is used to conduct information.

Background and Objectives: The high time-resolution of UWB signals enables separation of individual multipath components. Therefore, time-domain processing of the power delay profile can reveal information about changes and movements in an enviroment. Our objective is to develop a short-range cooperative ultrawideband radar network (CURN) that can detect, locate and track objects, and also a stand-alone radar device that estimates vital life parameters such as respiratory and heart rates.

Technical Discussion: The CURN utilizes UWB pulses in the preamble of a packet for SNR improvement. Multiple packet transmissions are used to identify changes in the multipath profiles and detect movements. While the receiver of the intended destination exploits the received packet preambles, the transmission source can also switch betweeen transmit and receive modes and acquire additional information about the charateristics of a target movement. The payload of each packet is used to share detection results with neighboring devices for data fusion. The first stand-alone use of the UWB radar we have studied is remote sensing of respiratory rate. We have derived Cramer-Rao bounds of respiratory rate estimation error over AWGN channels.

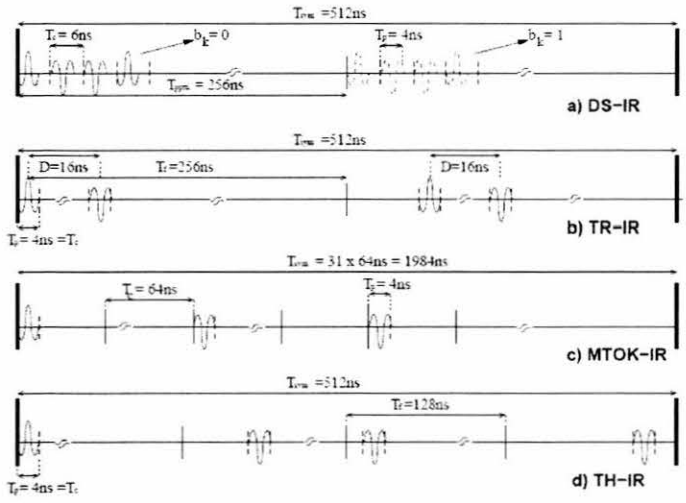
Collaboration: Our collaborations will start in FY07.

Future Direction: We are developing a framework for cooperative communication networks with radar and ranging capability.

Contact: Zafer Sahinoglu, Sinan Gezici
http://www.merl.com/projects/uwbradar/

Lab: MERL Technology Lab
Project Type: Advanced Development

Signal Waveform Design for UWB Ranging



MERL has extensively worked on ultrawideband signal design for high precision ranging, and provided contributions to the emerging IEEE 802.15.4a standard. Time-hopping impulse radio (TH-IR), direct sequence impulse radio (DS-IR), transmitted-reference impulse-radio (TR-IR) and M-ary Ternary Orthogonal Keying (MTOK) waveforms are the major ones that have been considered for use in packet preambles. We have identified the trade-off between these waveforms. We have been also evaluating alternative signal structures for IEEE 802.15.4a to increase the probability of detecting the start of frame delimiter (SFD) of IEEE 802.15.4a packets via both coherent and non-coherent receivers.

Background and Objectives: The waveform of a transmitted symbol has significant impacts on the performance of TOA estimation algorithms. By changing position and number of UWB pulses in a symbol and varying the symbol duration, unique symbol waveforms can be generated for use in ranging preambles. These parameters impact the level of inter-pulse-interference and inter-symbol interference, and necessitate development of post reception signal processing algorithms to achieve precision ranging.

Technical Discussion: The following constraints should be considered in waveform design. The maximum excess delay (MED) of the channel may vary from 30ns to 100ns or larger. Therefore, the waveform should have a large enough zero correlation zone prior to the leading edge to minimize self-interference. If multiple pulses per symbol are transmitted and the duration between pulses is smaller than the MED of the channel, aggregating the energies over multiple pulses may smooth the edges; and it may be better to operate on the non-aggregated samples for edge detection (e.g. for direct-sequence (DS)-IR).

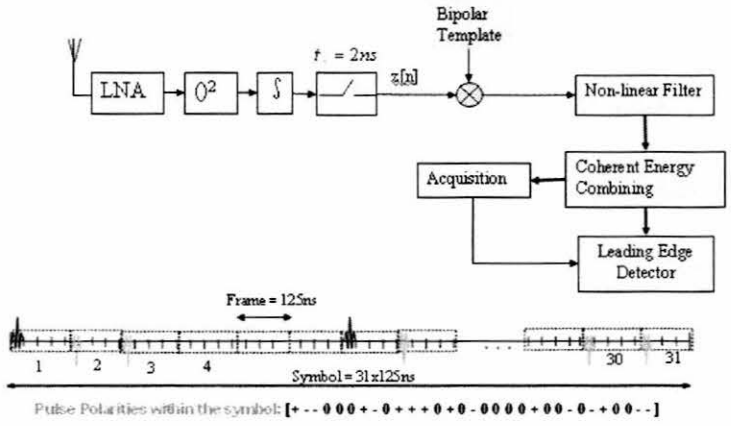
Collaboration: MERL and Musen-IP-Bu from ITC worked together on some SFD signal structures and design of non-coherent receivers for SFD detection.

Future Direction: Circuit implementation of the SFD waveforms and SFD detection algorithms; contribution of SFD waveforms to the IEEE 802.15.4a standard

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

Lab: MERL Technology Lab
Project Type: Advanced Development

Ultrawideband Ranging and Locating



The purpose in this project has been to develop high precision ranging algorithms for ultrawideband systems and provide technical contributions to the emerging IEEE 802.15.4a standard. During the course of this research, a concept of private ranging has been introduced; and it is now included in the 2nd draft specification of the IEEE 802.15.4a. Another technical contribution that is in the draft specification is scaling of the confidence interval for range estimation accuracy. Moreover, various receiver-side signal processing algorithms have been developed to improve time-of-arrival estimation.

Background and Objectives: Ultrawideband signals due to their wide bandwidth provide high time resolution, which helps to separate individual multipath components. In ranging, the arrival time of the first path is important, because it corresponds to the shortest distance between two devices. In obstructed line-of-sight environments, the first path might not be the strongest. Therefore, in ToA based ranging systems, post-reception signal processing algorithms are needed to locate the first arriving path. Another issue in ranging systems is that malicious devices should be prevented from learning ranges of other devices in a network. It has been also our objective to develop a defense mechanism against spoofing and impostor attacks in ranging networks.

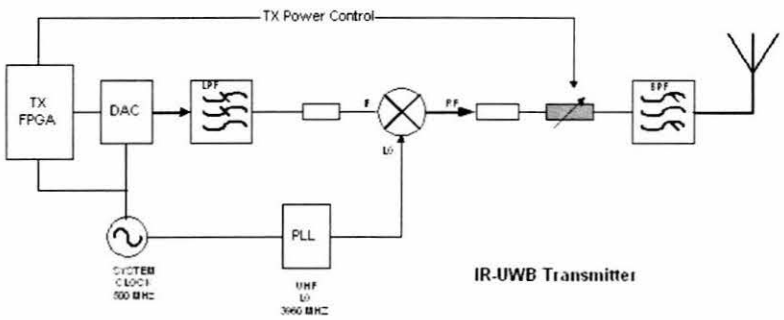
Technical Discussion: The private ranging scheme contributed to IEEE 802.15.4a relies on an initial handshaking protocol which uses encrypted messages between ranging parties. During handshaking, the two devices agree on waveforms that are to be used to form the preambles of the actual ranging packets. Hence, a malicious device is enforced to predict the selected waveforms to be able cause disturbance. If the preamble of a packet is corrupted with interference, ToA estimation might be quite challenging. Therefore, interference resilience also carries great importance for ranging. We have been also developing non-linear filtering techniques to mitigate multi-user interference.

Future Direction: Our next step is to build a prototype ultrawideband ranging platform and to develop and implement low-overhead positioning schemes for network wide position awareness.

Contact: Zafer Sahinoglu, Philip Orlik
http://www.merl.com/projects/uwbranging/

Lab: MERL Technology Lab
Project Type: Research

Impulse Radio UWB Design



Impulse radio Ultra-Wide Band systems communicate by using a sequence of short pulses. The standard for impulse radio UWB system has evolved in the last few years and is coming to maturity. Impulse radio UWB can potentially provide a solution for low cost, mid/high data rate communication link. It also has the advantage of being able to include ranging/location capability in each node. Due to the high bandwidth of pulses, it can potentially achieve very high accuracy ranging. All these features make impulse radio UWB system attractive for short range wireless network applications. We have investigated different implementation for impulse radio and designed the transmitter, two types of non-coherent receiver and a transmitted reference receiver.

Background and Objectives: MERL has been actively engaged in the impulse radio UWB standard (IEEE 802.15.4a) and has essential patents included in the standard. MERL proposed the hybrid system which allows coherent, non-coherent and transmitted reference (TR) nodes coexist in the same network. One of the challenges for impulse radio UWB is to implement such a system with low cost and high performance. The objective of this project is to study different architectures and implementations and to develop a system to prototype MERL's hybrid proposal. We also studied the possible hardware/software implementation of MERL's high precision ranging algorithm.

Technical Discussion: We investigated several different system architectures. We evaluated different impulse radio analog hardware implementation. We built the transmitter prototype using FPGA, high speed DAC and direct upconversion RF front end. Our transmitter is designed to transmit the hybrid system waveform. Our signal has greater than 500MHz bandwidth at 4 GHz. We studied different receiver designs and built two different types of non-coherent receiver and one TR receiver. We have identified issues associated with each designs and have proposed improvements.

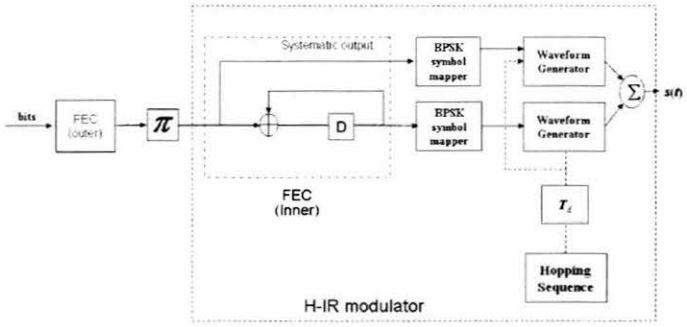
Collaboration: We have close collaboration with Johosoken on this project.

Future Direction: We are building a IR-UWB network based on the research from this project and looking into improvement in both cost and performance.

Contact: Chunjie Duan
http://www.merl.com/projects/iruwbdesign/

Lab: MERL Technology Lab
Project Type: Research

Hybrid Modulation for Impulse Radio



Impulse radio systems communicate by encoding each symbol as a sequence of short pulses. Due to the absence of frequency upconversion, low-cost transceiver structures are possible. Currently our project aims to make use of the low-cost transceiver capabilities of impulse radio systems by investigating modulation and coding techniques that are appropriate for embedded networking and sensor network applications.

Background and Objectives: Our previous work sought to solve cost and complexity issues associated with coherent UWB receivers by developing a modulation that combined features of BPSK, Transmitted Reference (TR) and Pulse Position Modulation (PPM). The result was modulation technique we termed Hybrid-Impulse Radio (H-IR). The key feature of H-IR modulation is that is allows simultaneous reception by variety of receivers types ranging from simple energy detectors to complex coherent rake receivers. After our initial discovery of the H-IR modulation is was observed that the technique may be viewed as a form of Trellis coded modulation (TCM) which led us to explore methods that could further improve the performance of both coherent and non coherent receivers.

Technical Discussion: A simple approach towards improved performance might be to include a forward error correcting code (FEC) prior to the H-IR modulation. However, the optimal use of FEC with H-IR is not the trivial application in which one simply codes the bits prior to H-IR modulation. This is due to the fact that the H-IR is already a coded system. In essence the addition of FEC led us to consider serially concatenated codes and a class of iterative decoders that can give performance results that approach those of turbo codes.

The H-IR modulation has been extended to consider the use of an outer FEC and interleaver along with a modified H-IR modulation that still enables the simultaneous reception by coherent, TR, and non-coherent receivers and also offers improved performance to all three types of receivers. The greatest improvement is seen in coherent receivers which can make use of advanced iterative decoding. In addition, the reduced complexity receivers (TR, and noncoherent) now obtain an error protection provided by the outer FEC encoder but do not require iterative decoders.

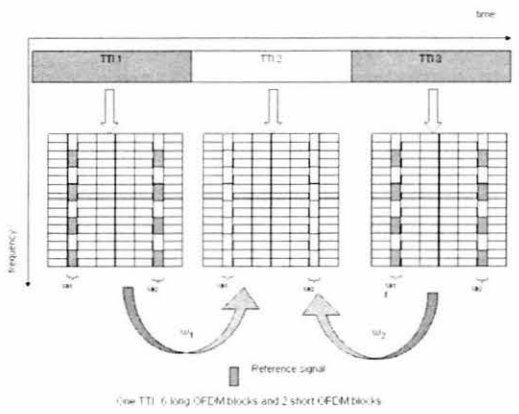
Collaboration: Johosoken

Future Direction: We will investigate further enhancements to the H-IR scheme such as the inclusion of higher order modulation for greater bit rates. Additionally, prototyping of the H-IR scheme and both coherent and noncoherent receivers is underway.

Contact: Philip Orlik
http://www.merl.com/projects/impulseradio/

Lab: MERL Technology Lab
Project Type: Research

Super 3G Uplink Reference Signal Design



One TTI: 6 long OFDM blocks and 2 short OFDM blocks

We present novel Transmission Time Interval (TTI) grouping techniques to reduce the significant uplink reference signal overhead in next generation wireless communication systems.

Background and Objectives: Reference training signals are inserted by handsets that transmit data to enable coherent demodulation of the signal received by the Node B (base station). More reference signals are needed when multiple antennas are used for transmission. In addition, the handsets, including those that are not currently transmitting data, also transmit additional reference signals to help the Node B estimate the instantaneous channel state and schedule users based on the channel estimates. Doing so enables the Node B to exploit multi-user diversity and significantly increases the spectral efficiency achievable by the system. The reference signal overhead in next generation systems is considerable. It consumes precious bandwidth and time resources that could otherwise be used by the handsets for transmitting data and control signals, or be used to accommodate channel quality estimates from more handsets. Reducing the overhead of reference signals used for coherent demodulation will improve the overall spectral efficiency of the system.

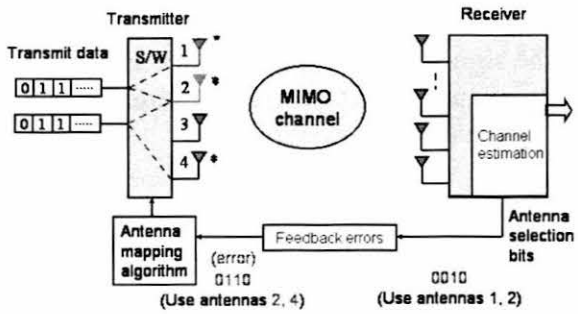
Technical Discussion: We present a reference signals design that is based on grouping multiple consecutive transmission time intervals (TTI). A TTI is the basic duration of transmission in third generation cellular systems, and is designed to be very short to meet tight latency requirements. In TTI grouping, the same reference signal pattern is not transmitted over all TTIs. Instead, the reference signal pattern is varied on a per-TTI group basis, and is tailored to reduce the total reference signal overhead. Shown in the figure is an example of a TTI grouping pattern consisting of 3 TTIs that reduces the reference signal overhead by 33%. At the receiver, channel estimation is performed over all the TTIs in the TTI group. Accurate channel estimation, and thus coherent demodulation, is possible so long as the channel does not vary significantly within the TTI group duration. Therefore, the method is suitable for lowering the reference signal overhead for slower moving handsets.

Future Direction: The method can be potentially be applied to CQI measurements and MIMO systems for reference signal design. It can also be adopted in OFDM-based wireless communication standards such as 3GPP Long term evolution (LTE), WiMAX (IEEE 802.16), and Wi-Fi (IEEE 802.11).

Contact: Shengjie Zhao
http://www.merl.com/projects/s3gref/

Lab: MERL Technology Lab
Project Type: Initial Investigation

Antenna Selection Signaling & Verification for Super 3G



Feedback channels in current wireless systems, which are used for signaling from the receiver (e.g., base station) to the transmitter (e.g., mobile handset), are often severely bandwidth-limited and are error-prone. We investigate novel designs for feedback signaling to enable transmit antenna selection in such error-prone channels. We also develop several 'verification' algorithms at the receiver for it to optimally cope with the errors.

Background and Objectives: Multiple-input multiple-output (MIMO) systems promise remarkable gains in both information rates and the reliability of transmission over wireless channels without requiring any additional bandwidth. However, the increased hardware and signal processing complexity has inhibited their widespread adoption.

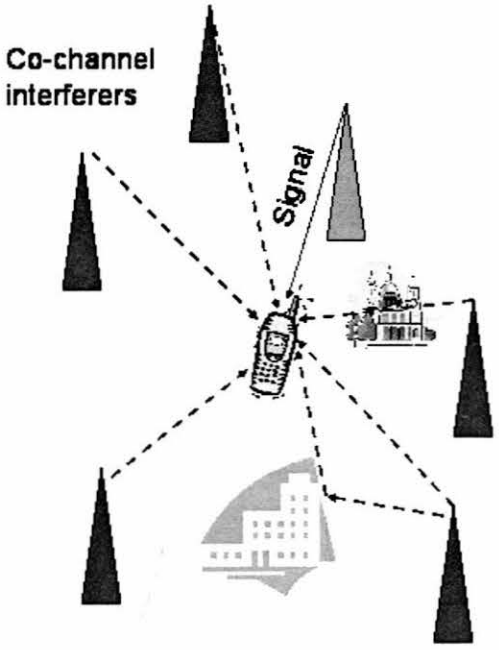
Transmit antenna selection reduces MIMO complexity by using fewer transmit receive chains. Feedback from the receiver is often needed for the transmitter to know which antenna(s) to use. While antenna selection achieves the highest possible diversity order under ideal assumptions such as instantaneous error-free feedback, its performance under non-ideal conditions such as with feedback errors has received relatively little attention. Not accounting for such errors can be catastrophic given the very low data error rates the systems should achieve.

Technical Discussion: Errors rates in the feedback channel can be as high as 4%. Subject to the feedback channel's bandwidth and latency constraints, we derive approximate performance metrics and efficiently find the optimal feedback signaling to minimize the average symbol error rate. Equally important is verification at the receiver, in which the receiver uses properties of the feedback channel while decoding data. We consider various verification algorithms such as single or multiple symbol-based verification, and verification with additional side information.

Future Direction: Our aim is to develop a common framework to design and analyze the feedback signaling and the corresponding receivers (with verification) for various MIMO schemes.

Contact: Neelesh Mehta
Lab: MERL Technology Lab
http://www.merl.com/projects/antenna_selection_signaling/
Project Type: Research

Downlink Co-channel Interference Modeling



We are developing a novel, simple, and parametrically flexible method that can provide an analytically tractable probability model for the downlink (base station to handset link) interference in cellular systems. Doing so has great relevance in cellular system analysis and design.

Background and Objectives: Co-channel interference, which arises from transmissions by neighboring base stations is a fundamental characteristic of cellular systems today. As the signals travel from the base stations to the handset, they undergo random fluctuations such as shadowing, which occurs due to buildings, and fading, which occurs due to multiple reflected copies of the signal arriving at the handset. These drive the behavior of the signal to interference and noise ratio (SINR), which determines how fast data can be transmitted reliably to the handset.

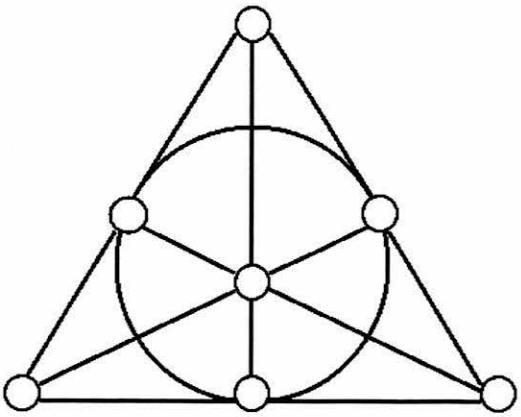
Technical Discussion: Attempts have been made over the past several years to characterize analytically the distribution of the sum of interferers given its direct impact on the SINR. While several methods such as the Fenton-Wilkinson method, Schwartz-Yeh method, etc., and their generalizations are available in the literature, each has its own disadvantages and none is unarguably better than the rest. We have proposed a moment generating function-based matching method to accurately model the probability distribution of the interference, which may arise from sum of lognormal or Rayleigh-Lognormal or Rice-lognormal interference components. Even the case in which the shadowing of different base stations is correlated is covered. The method is much more accurate than the conventional moment matching approaches in modeling smaller interference values, and matches the accuracy of the conventional approaches in modeling larger interference values.

Future Direction: Exciting opportunities exist to generalize the technique to cover a broad array of distributions that commonly arise in the literature.

Contact: Neelesh Mehta
http://www.merl.com/projects/lognormal/

Lab: MERL Technology Lab
Project Type: Research

High-Speed Decoding of Error-Control Codes



A simple finite geometry

We have developed high-speed decoders for error-control codes. Our "replica-shuffled" decoders are highly parallelized, and give excellent performance versus complexity trade-offs. We are currently building VLSI prototypes of these decoders for low-density parity check (LDPC) codes. We are also investigating codes and decoders that are suitable for very high-speed optical fiber channels.

Background and Objectives: Turbo-codes and LDPC codes are an attractive solution to the channel coding problem, capable in principle of performing very near the ultimate limits given by information theory. However, an important issue for decoders of such codes is the speed with which they can successfully decode--most standard iterative decoders for turbo-codes and LDPC codes require an excessively large number of iterations for good performance.

Technical Discussion: We have developed a class of high-speed decoders called "replica shuffled" decoders. In these decoders, different replicas decode different parts of an error-control code simultaneously, and exchange information in an intelligent way. Replica-shuffled versions of many different iterative decoders can be constructed, including belief-propagation decoders for low-density parity check (LDPC) codes, bit-flipping decoders for codes based on finite geometries, turbo-decoders, and turbo product code decoders. Replica shuffled decoders can typically successfully decode in far fewer iterations than standard decoders. We are constructing VLSI hardware prototypes of replica-shuffled LDPC decoders using FPGA's.

For the optical channel application, which has very challenging requirements, we have focused on bit-flipping decoders for LDPC codes based on finite geometries. Such codes have relatively good minimum distances so they do not tend to have an "error floor," which is an important feature for this application.

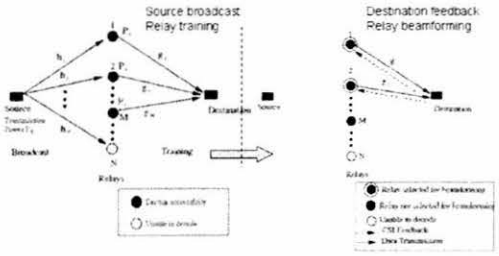
Collaboration: This work is in collaboration with Prof. Marc Fossorier from the University of Hawaii, his students Yige Wang and Juntan Zhang, and the error-control coding team at Johosoken.

Future Direction: We plan to build ASIC prototypes of these decoders. We are also collaborating with MERL's analog logic design team to develop analog versions of these decoding algorithms. Finally, we plan to continue to explore codes and decoding algorithms suitable for the optical channel application.

Contact: Jonathan Yedidia, Jeff Proctor
http://www.merl.com/projects/ecc_optical/

Lab: MERL Research Lab
Project Type: Research

Reliable Energy Efficient Cooperative Relaying



We are interested in the energy efficiency of a general class of cooperative transmission schemes that transfer information from the source to the destination node via intermediate cooperative relays over wireless channels. The wireless channels are a difficult environment to handle on account of their time-varying nature, and due to the presence of fast fading and shadowing attenuation. The basic question we ask and answer is whether the energy efficiency enabled by cooperative communication among many available relay nodes is significantly better than the overhead incurred in enabling this cooperation.

Background and Objectives: Cooperative communication networks, in which wireless nodes cooperate with each other in transmitting information, promise significant gains in overall throughput and energy efficiency. We are interested in energy-efficient cooperative relay networks than can deliver information from a source to destination with some guaranteed reliability. Energy efficiency can be improved by exploiting the broadcast nature of the wireless channel as multiple relays can simultaneously receive the signal transmitted by the source. Cooperative beamforming algorithms, in which the relays linearly weight their transmit signals to make them arrive coherently at the destination, provide an attractive energy efficient way to forward data to the destination. However, communicating the linear weights to the relays requires energy that increases as the number of relays increases. Minimizing the total energy consumption shows a distinct tradeoff between decreasing energy consumption for data transmission by using more relays and decreasing overhead by using fewer relays.

Technical Discussion: We consider a cooperative wireless network where a set of nodes cooperates to relay in parallel the information from a source to a destination. The relays forward information only after they have reliably decoded the message from their source. The source broadcasts the data to the relays, some or all of which cooperatively beamform to forward the data to the destination. We generalize the standard approaches for cooperative communications in two key respects: (i) we explicitly model and factor in the cost of acquiring channel state information (CSI), and (ii) we consider more general selection rules for the relays and compute the optimal one among them. Our simple relay selection and outage criteria exploit the inherent diversity of relay networks and satisfy a mandated outage constraint. These criteria include as special cases several relay selection criteria, such as best single relay selection, considered in the literature.

Collaboration: Ritesh Madan (Stanford University)

Future Direction: Our next step is to extend our insights to general energy-efficient reliable networks with multiple possible routes.

Contact: Andreas F. Molisch
http://www.merl.com/projects/cooperativerelaying/

Lab: MERL Technology Lab
Project Type: Research

Adaptive Tree Addressing Scheme for Wireless Ad Hoc and Sensor Networks

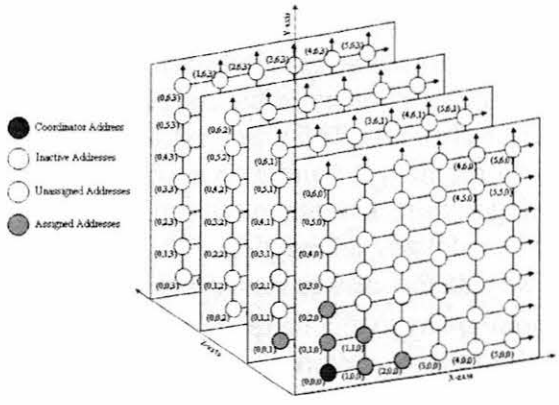


Figure 1

See Color Figure 6

Address assignment in wireless sensor networks, where nodes have limited resources, is a challenging issue. We introduce a novel concept of organizing address space into a multi-dimensional hypercube, as a 3-dimensional space shown by Figure 1, and then transforming that space into a tree structure. The tree dynamically and adaptively grows along any dimension as new nodes join. Our scheme is efficient and has low storage and communication costs. It avoids bottlenecks typically faced by central address assignment approaches and eliminates a need for address conflict detection and resolution mechanism as needed by many distributed addressing schemes. It also removes a need for routing tables by allowing nodes to route frames along underlying addressing tree.

Background and Objectives: Address space is a precious network resource that must be utilized judicially. An addressing scheme must be robust and flexible enough so that it adapts to geographical distribution of nodes to avoid formation of address starved pockets in networks. Finally, a distributed addressing scheme should be able to identify and reclaim addresses assigned to devices that ceased to function. Our scheme aims at addressing these.

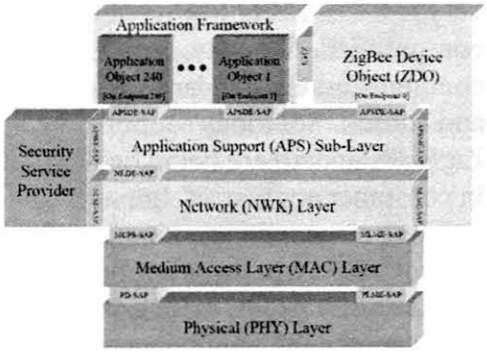
Technical Discussion: We organize address space into an N-dimensional hypercube, where N>1. Each address, thus, consists of N values called address components. There are constraints on the number of address values a node gets allocated, called its address share, which it can freely assign to its children. In general, nodes along nth dimension, 1=n=N, can have up to n children. That defines a mechanism for mapping the address space into a tree structure, which makes the basis of our approach. We partition the hypercube by cutting its size along all dimensions, for example, to half. Initially we use only this smaller cube, called active address space (AAS), for address assignment. If any node needs additional addresses, it can expand AAS along a suitable dimension by allocating an extra bit to the corresponding address component, thus, effectively doubling its size. A node can always expand AAS if at least one free bit is available. After each expansion, however, rest of the nodes must be informed about new composition of addresses. That is done by transmitting a network wide broadcast frame. Our scheme offers a trade off between adaptability and cost (due to broadcasts) making it well suited to a wide spectrum of applications.

Future Direction: We need to use simulation for performance characterization of this scheme under various conditions and deployment scenarios.

Contact: Ghulam Bhatti
http://www.merl.com/projects/adaptivetree/

Lab: MERL Technology Lab
Project Type: Research

Zigbee Application Interface



ZigBee is a wireless multi-hop mesh networking protocol that aims at a wide variety of applications including environmental and industrial monitoring, surveillance, home automation, tracking of goods and personnel, and automatic meter reading. In order to address the needs of all these diverse systems, ZigBee provides an interface with a common set of functions and services used by different profiles. A profile defines a standard set of attributes, parametric values, and functions that can be used by a given class of devices. The interface acts as a bridge between ZigBee network layer and the profiles (and application programs). MERL has developed software for ZigBee network layer, APS and ZDO sub-layers, and integrated this code with MAC and PHY on several platforms with different radios. ZigBee code developed at MERL has been used by Renesas for obtaining ZigBee platform certification.

Background and Objectives: ZigBee aims at facilitating low bit-rate, low cost and low power-consumption sensor networks with features such as security, automatic address assignment and link quality based routing. MERL has been at the forefront of ZigBee related activities including standardization of its specifications, development of software stack. The core ZigBee technology developed at MERL is being used by Renesas to produce a ZigBee chip set whereas MELCO will use this technology to develop value added application systems and services.

Technical Discussion: ZigBee follows a layered approach to offer a complete set of wireless networking solution. It uses IEEE 802.15.4 wireless standard as its MAC and PHY layers. The network layer, defined by ZigBee Alliance, facilitates multi-hop mesh networking with link quality based route discovery mechanism. The ZigBee interface consists of two sub-layers, called Application Support Sub-layer (APS) and ZigBee Device Object (ZDO). APS provides interface for data services such as transmitting and receiving data frames by using network layer. It also provides an indirect-addressing facility via its binding table and an acknowledge/retry service which can be used to improve reliability of data transfer. ZDO provides Device & Service Discovery, Node Management and Binding Management services to the applications.

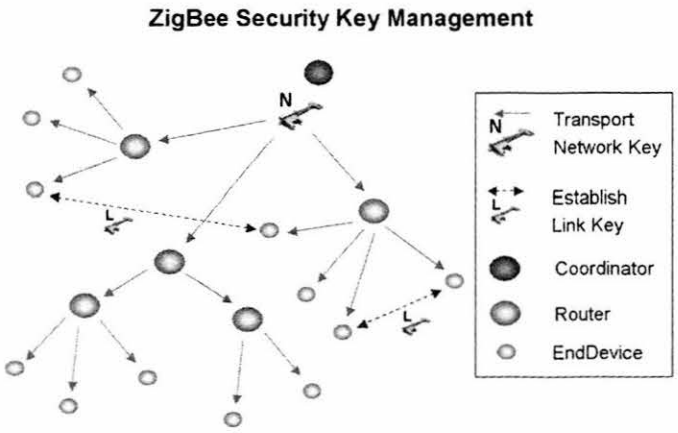
Collaboration: MERL has developed ZigBee stack on multiple platforms in close collaboration with Renesas, Johosoken, and Rikiden. We have been working closely with Sentansoken, and Jyukanken on ZigBee based application systems development.

Future Direction: MERL will continue development and research work related to ZigBee during next year. Such aspects as ZigBee security, device location estimation, and location based application systems will be focus of our effort. We expect to implement new features being included into the later versions of ZigBee specification.

Contact: Ghulam Bhatti, Frederick J. Igo, Jr.
http://www.merl.com/projects/zigbeeinterface/

Lab: MERL Technology Lab
Project Type: Advanced Development

ZigBee Security Software



See Color Figure 7

The ZigBee Alliance has designed a very low-cost, very low power consumption, two-way, wireless communications standard for sensing and control. Solutions adopting the ZigBee standard are being be embedded in consumer appliances, home and building automation and industrial controls. Many of these applications have security needs. MERL has already implemented the protocols required for a wireless, ad hoc network and is now developing the capability to make it secure for home, commercial and industrial applications.

Background and Objectives: Since 2002 MERL has actively participated in and contributed to the ZigBee Alliance. Presently, we are developing ZigBee Security services that 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 network.

Technical Discussion: Supporting the range of systems from lightweight home applications to robust industrial and commercial applications is required of ZigBee Security. By building upon the security of lower protocol layers (802.15.4 MAC/PHY and ZigBee NWK) MERL's implementation of the Application-Support Sub-Layer (APS) and ZigBee Device Object (ZDO) is able to complete the Security picture for ZigBee. The APS layer implementation provides secure transmission and reception of data frames and secure establishment and management of cryptographic keys. The ZigBee Device Object (ZDO) implementation manages the security policies and the security configuration by issuing primitives to the APS layer.

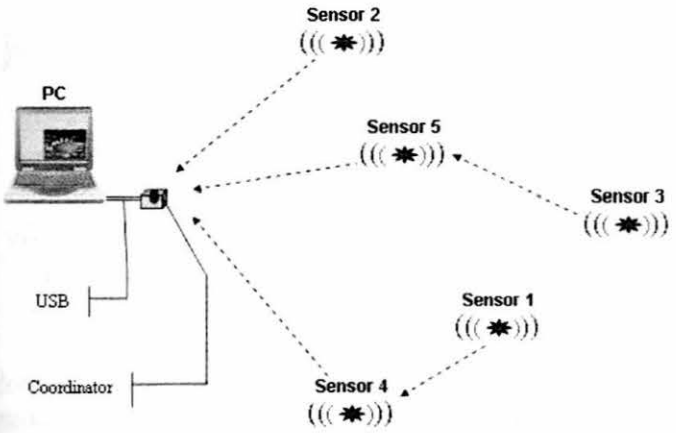
Collaboration: MERL has been working closely with both Johosoken and Renesas to deliver a fully functional ZigBee protocol stack including Security. Jyukan-ken, Shinden and Rikiden, are all either planning or considering embedding ZigBee into their system solutions.

Future Direction: ZigBee is still evolving as a standard and MERL will continue participating as an active contributor, specifically in the area of Security improvements. Distribution of group keys, which is a shared key within a select set of nodes, is being considered for future development. At the same time MERL will continue to help MELCO's business units leverage ZigBee's advantages and security by promoting and supporting inclusion in their system solutions.

Contact: Frederick J. Igo, Jr., Ghulam Bhatti
http://www.merl.com/projects/zigbeeaplsec/

Lab: MERL Technology Lab
Project Type: Advanced Development

Image Over ZigBee Networks



The ZigBee is a low rate and low power networking technology for short range communications, and it currently uses IEEE 802.15.4 MAC and PHY layers and does not have transport layer functionalities such as packet fragmentation and reassembly. While it is not a stated goal of the ZigBee Alliance to support the transfer of images over the network, it is clearly a desirable capability especially for surveillance systems.

Background and Objectives: Our objective is to develop a test bed system for JPEG image transmission over ZigBee network to explore adaptive source coding based on link quality and investigate various flow control mechanisms. The goal is to be able to reliably acquire 320x240 JPEG images from multiple sources in 2-3 seconds.

Technical Discussion: The prototype is implemented on a single M16C microprocessor. It includes an embedded processor, memory, a CMOS image sensor and JPEG encoder. M16C microprocessor firmware includes ZigBee networking (NWK),IEEE 802.15.4 MAC layer protocols, application and communication parts. Image Acquisition Board has CMOS image sensor and JPEG encoder. Fragmentation and reassembly are performed at the application layer. Source images captured and encoded to JPEG format in the real-time and transmitted to the destination either continuously or per request. One ZigBee device is connected to a PC via an USB connection and acts as a gateway into the wireless portions of the network. The gateway also allows the collection of various data from the network such as network topology, neighbor tables and route information. An application runs on the PC and controls parameters of the network such as routing and topology. It also displays received images for visual quality checking. The route that the packets traverse is also available to the gateway/PC data sink.

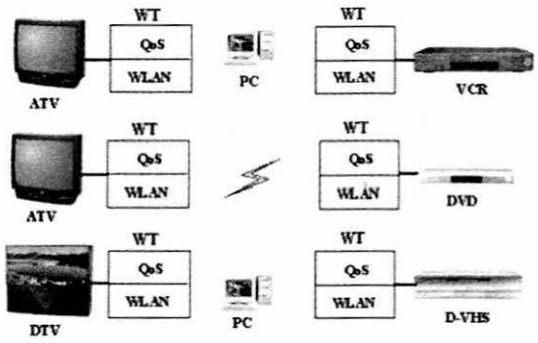
Collaboration: Sentansoken

Future Direction: Integration with APS and ZDO levels as defined by ZigBee, add sound encoding and transmission capability.

Contact: Georgiy Pekhteryev
http://www.merl.com/projects/zigbeeimage/

Lab: MERL Technology Lab
Project Type: Initial Investigation

QoS for Broadband Wireless Home Networking



MERL has developed QoS technologies for broadband wireless home networks to guarantee the required QoS for the higher priority traffics whenever networks are over loaded. MERL's QoS technologies dynamically monitor bandwidth usage of the wireless medium. If the bandwidth is in shortage the lower priority traffics are selected as victims based on a newly defined priority. The bandwidth allocation of victims shall be reduced to guarantee the QoS requirements of higher priority traffics. The bandwidth allocation is dynamically adjusted so that bandwidth is efficiently used. MERL's QoS technologies work on top of 802.11e standard at LLC layer, fully utilize existing 802.11e data as input, and do not need any new input from upper layers or users. MERL's QoS technologies have been implemented into QoS software delivered to Sentansoken and tested in a home networking system as shown in the figure. The results show that MERL's QoS technologies enhance and complement 802.11e QoS mechanisms to provide a better QoS solution. MERL has also completed a comprehensive QoS report to Johosoken, which includes the advanced QoS techniques developed MERL and others.

Background and Objectives: In FY04 and FY05, MERL has developed QoS software for Sentansoken's wireless dongle. MERL has identified that QoS technologies beyond MAC layer are needed to provide a complete QoS solution, and proposed dynamic QoS technologies to achieve efficient bandwidth management for broadband wireless home networks. The objective is to guarantee required QoS for higher priority traffics in the cost of lower priority traffics whenever networks are over loaded.

Technical Discussion: Due to the high bandwidth demanding and the variability of wireless links, QoS has been a challenge for AV stream transmission over wireless networks. IEEE 802.11e standard has introduced dedicated bandwidth in combination with priority techniques at MAC layer to enhance QoS capability of wireless networks. However, 802.11e's priority techniques can not guarantee the required QoS except in a very lightly loaded network. 802.11e's dedicated bandwidth technique also faces challenge when bandwidth is in shortage. MERL's QoS technologies aim at enhancing IEEE 802.11e QoS mechanisms to address QoS issue in the case of bandwidth shortage.

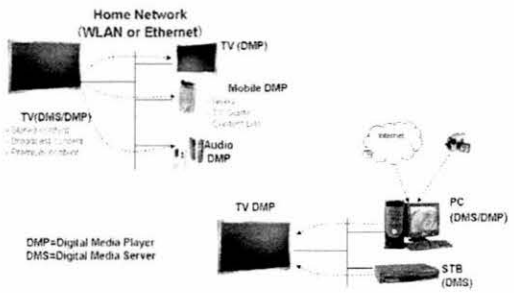
Collaboration: MERL has worked closely with Sentansoken and Johosoken on QoS for broadband wireless home networks and future multimedia application deployment.

Future Direction: MERL's current QoS technologies are designed for IEEE 802.11 networks. MERL will develop QoS technologies for UWB networks and other wireless networks.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

DLNA Digital Media Server (DMS) Design



The Digital Living Network Alliance (DLNA) is a cross-industry organization of consumer electronics, computing industry and mobile device companies. DLNA is focused on delivering interoperability guidelines based on open industry standards to complete the cross industry digital convergence. DLNA has published a common set of industry design guidelines that allow vendors to participate in a growing marketplace of networked entertainment and mobile devices, leading to more innovation, simplicity and value for consumers.

Background and Objectives: The DLNA Home Networked Device Interoperability Guidelines are use case driven and specify the interoperable building blocks that are available to build platforms and software infrastructure. They are focused on interoperability between the devices for personal media uses involving imaging, audio and video. In the DLNA digital home, it will be common for consumers to: Easily acquire, store and access digital music from almost anywhere in the home; Effortlessly manage, view, print and share digital photos; Take favorite content anywhere to share with family and friends; Enjoy distributed, multi-user content recording and playback

Our objective is to develop DLNA compliant Digital Media Server (DMS) prototype in collaboration with Johosoken M-Net team. Successful development and testing of our prototype will lead to quick integration in the future CE devices such as Digital TV and DVD. MERL has developed DLNA v1.5 compliant DMS. It has been certified by UPnP IC and DLNA D2G.

Technical Discussion: The basic function of a DLNA DMS is to implement the ContentDirectory and the ConnectionManager services as defined by the UPnP architecture. These services enable device to discover the presence and capabilities of other devices on the network, and control and collaborate with these devices in a uniform and consistent manner. In addition to UPnP services DLNA DMS provides media streaming transport via utilization "HTTP GET" method as specified in the DLNA Guidelines. In order to make network experience seamless to the end user we have implemented functionality recommended by IETF draft "Dynamic Configuration of IPv4 link-local addresses" also known as "Auto IP" feature. Our target platform is Linux OS.

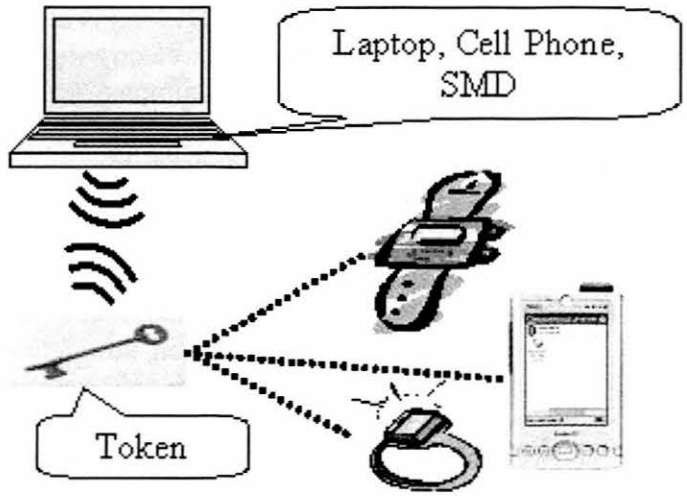
Collaboration: Johosoken

Future Direction: Enhancing current implementation by adding optional features such as time and byte range based seek/trick play, support more optional media formats, implementation of content Upload/Download, improve overall performance of DMS by employing more efficient algorithms for parsing and streaming content.

Contact: Georgiy Pekhteryev, Jinyun Zhang
http://www.merl.com/projects/dlna-dms/

Lab: MERL Technology Lab
Project Type: Advanced Development

Token-Enabled Authentication for Securing Mobile Devices



The increasing demand for sharing data, performing business transactions, and mobile computing has enabled a wide utilization of mobile devices, while increasing the vulnerability of exposing private/confidential data on mobile devices in the case of theft. Although the vulnerability can be closed with available mechanisms, such as passwords, biometrics, and asymmetric-key-based token authentication, they either place unpleasant burdens on the user, or degrade the system performance of the mobile device, such as power consumption and delay, significantly during the authentication process. In order to protect the private/confidential data on mobile devices from theft with minimal burden, we propose an approach to enable a symmetric key-based authentication between a token and a mobile device. The approach includes three processes: an initial authentication process, a key management process, and a subsequent authentication process with the key established in the previous operations. The patent focusses on the initial authentication process of the system, how do those secret keys get into the devices in the first place? A hardware prototype to prove the concept of the system has been demonstrated and integrated into Johosoken's development efforts.

Background and Objectives: MERL is working closely with the network security team in Johosoken for a secure memory device that will be used in both the cellular phone and the PC. The objective is to integrate this system into next generation cellular phone products.

Technical Discussion: The concentration here is on the initial deployment of the "binding" keys between the mobile device and the token that is worn by the user. Two methods have been proposed. The primary method involves the use of an online authentication server which identifies the mobile device and token to be bound. It then generates a binding key and deploys this key to both device using a secure protocol. The secondary method allows for the initial deployment of the binding keys to occur without the assistance of the authentication server. The secondary method pays the price of not being as secure as the first.

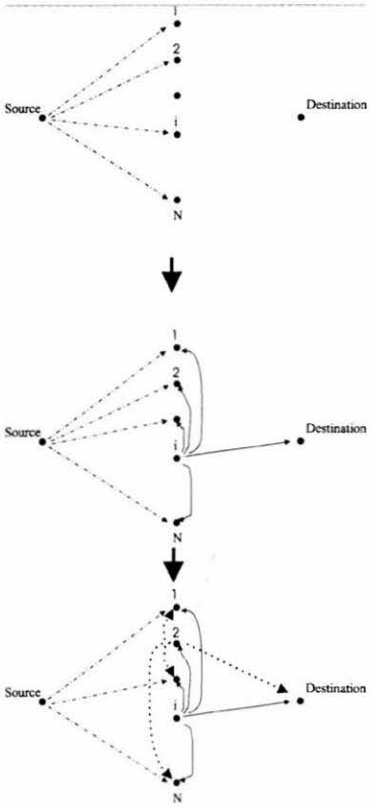
Collaboration: MERL is collaborating with the one of the Jouse teams in Johosoken headed by Mr. Munaka.

Future Direction: Deployment into product first within Japan as a result of this project. The improvement of seamless authentication in the field of network security and data security.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

Cooperative Communications Using Rateless Codes



We introduce a new method for transmitting information via a number of relay nodes. By using so-called rateless codes, we allow the receiver to accumulate information in a fast and energy-efficient way. We propose a quasi-synchronous, and an asynchronous protocol for the multiple-access by the different nodes. The quasi-synchronous protocol is more spectrally efficient, while the asynchronous protocol is more energy-efficient.

Background and Objectives: In standard wireless communications, information is transmitted directly from a transmitter to the receiver. The transmit power has to be large enough so that the information arriving at a possibly far-away receiver is sufficient. This leads to large energy consumptions, and also increases the interference to other devices. Cooperative communications circumvents this problem by using relay nodes. The information is sent over multiple hops, and also via parallel routes. At the receiver, the information from all the different relay nodes is added up. Instead of adding up the energy from the different relay nodes at the receiver, we introduce a method where bits from the different relay nodes are added up. As a consequence, the message can be transmitted faster, and in a more energy-efficient way.

Technical Discussion: The transmission from the source to the destination via several parallel relays uses rateless codes, also known as fountain codes. They encode and transmit the source information in an infinitely long codestream. The codes have the special property that a receiver can recover the original information from unordered subsets of the codestream, once the total obtained mutual information from multiple sources marginally exceeds the entropy of the source information. Thus, only the required energy is transmitted. In this project, we propose two protocols. In the quasi-synchronous protocol, after L relay nodes have decoded the information, the source stops its transmission, and all the L relay nodes then simultaneously transmit the information to the destination. If the relays use different fountain codes, the receiver accumulates the mutual information, while if all relays use the same fountain code, the destination accumulates the energy. In the asynchronous protocol, each relay node starts to transmit to the destination as soon as it has decoded the source data. This speeds up transmission as it provides useful information not only to the destination, but also to the relay nodes that have not finished the decoding process yet.

Future Direction: Identifying routing methods for accumulation of bits.

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

Lab: MERL Technology Lab
Project Type: Research

Digital Video

The field of Digital Video embraces techniques that span across several disciplines including traditional electrical engineering areas such as signal processing, information theory, communication and networking, as well as computer science areas, such as data analysis, content understanding, and database. 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: Work in this area has focused mainly on video encoder optimization, e.g., obtaining the best coding efficiency with limited complexity. A major initiative this year has been efficient storage of video using H.264/AVC. We are also engaged in the design of new coding techniques to compress multiview video.

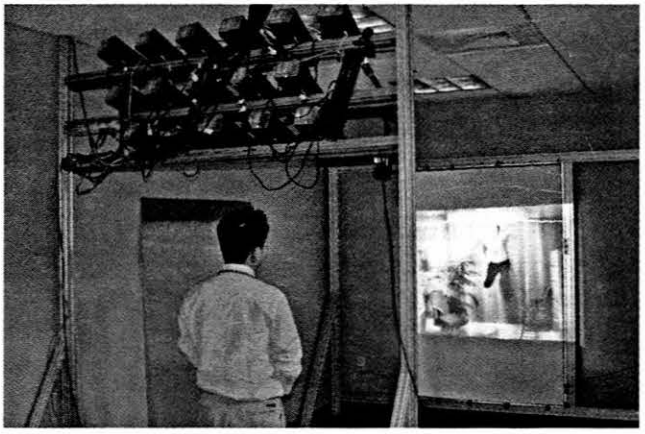
- Distribution: MERL has developed award-winning architectures and algorithms for transcoding video to specific bit-rates and spatio-temporal resolutions to satisfy terminal and network constraints. Active areas at the moment include scalable streaming of JPEG 2000 images sequences and IPTV.

- 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

3D TV



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: Matthias Zwicker (UC San Diego), Fredo Durand (MIT).

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

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

Lab: MERL Research Lab
Project Type: Research

Video Summarization for PVR's

See Color Figure 8

Personal Video Recorders have increasingly large storage capacity extending beyond 100 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. We are investigating the usability issues in a practical PVR application as well. We define tasks such as browsing a collection of programs, skimming through a program, watching the highlights, etc., and how summarization and video segmentation technologies apply.

Background and Objectives: In this project we emphasize the Personal Video Recorder application, which 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 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. Our sports highlights detection was featured in MELCO's DVR-HE50W DVD Recorder, the world's first DVD recorder with sports highlights playback.

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. An important constraint is feasibility on our target platforms. Our algorithms therefore have to be computationally simple and robust to the high variation in broadcast video. The second front is the usability of video summarization and browsing in PVR applications. From the user's point of view, there should be ease of use and flexibility. The technology should be seamlessly integrated with the typical tasks that a PVR user have, such as browsing through large number of programs, deciding what to watch, locating desired part of a program, or watching a summary of a program, etc. We plan to meet the flexibility requirement by developing scalable summarization algorithms that generate summaries of varying lengths.

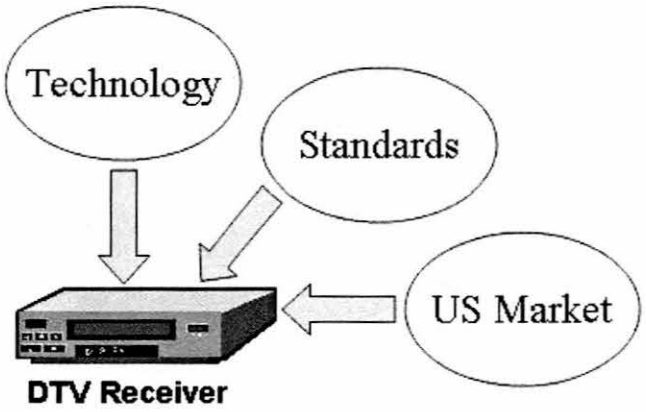
Collaboration: Sentansoken, Johosoken

Future Direction: We will extend our content analysis techniques to cover a wider range of genres using audio and video together. We will develop a holistic approach to application of video analysis, summarization and browsing technologies to PVRs that centers around user tasks.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

DTV Standards in the US



MERL is currently involved in activities related to the development of DTV receivers and in DTV related standardization and regulations in the US. We coordinate closely on these topics with our DTV related business units and other research labs within Mitsubishi Electric.

Background and Objectives: We participate and track standards and regulatory activities so that developed products will maintain compliance with the latest standards and US government mandates. Another important objective of our activity is to ensure information exchange on these activities throughout the company as part of the strategic planning process.

Technical Discussion: It is particularly important for Mitsubishi Electric to track related standardization and regulatory activities in the US that impact DTV receiver designs. We focus on standardization activities within CEA, ATSC, SCTE and SMPTE, as well as mandates issued by the FCC. Current topics of interest include electronic program guide (EPG), IPTV, advanced audio-visual codecs such as H.264/AVC and VC-1 for video, DTV tuner mandate, DTV closed caption, receiver performance guidelines, content protection systems, as well as digital cable-ready standards.

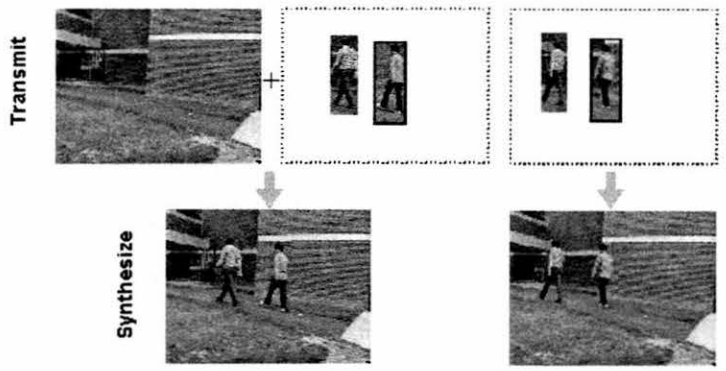
Collaboration: DTV related business units, as well as related departments in corporate research labs.

Future Direction: Continue to participate and track related standardization activities and developing technologies.

Contact: Anthony Vetro
http://www.merl.com/projects/dtv-standards/

Lab: MERL Technology Lab
Project Type: Advanced Development

Scalable 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 considering display and bandwidth limitations.

Background and Objectives: Recently, digital video recorder systems have come into wide use for efficient storage and transmission of surveillance contents. Many new systems employ JPEG 2000 for image compression and advanced image processing techniques to gain a better understanding of objects in a scene. To satisfy bandwidth constraints and display important parts of the images fast and precisely, the compressed images are efficiently transcoded in the compressed-domain based on the ROI information.

Technical Discussion: In our system, an image sequence is encoded and stored as a JPEG 2000 bitstream, and then the stored images are efficiently transcoded in the compressed-domain using a low-complexity adaptation technique that replaces data packets corresponding to higher quality layers with empty packets. In one particular streaming mode, the ROI's are transcoded with higher quality than the background to satisfy network constraints. We have also designed a dynamic rate control algorithm with frame skipping functionality that operates based on buffer occupancy, target rate and scene contents. This algorithm achieves high bandwidth utilization, hence better quality, and minimizes quality variations over time. Our transcoder also supports resolution reduction and cropping functions.

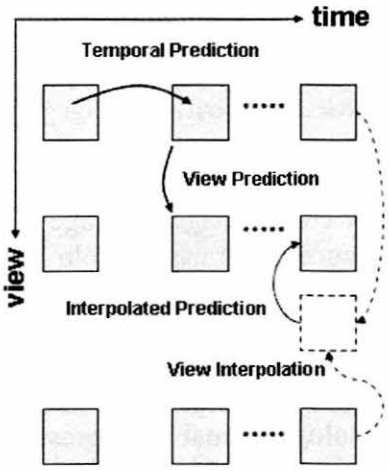
Collaboration: This project is done in collaboration with the Physical Security Group at Sentansoken.

Future Direction: We are considering novel error detection and recovery techniques for robust transmission.

Contact: Anthony Vetro, Derek Schwenke
http://www.merl.com/projects/roi-streaming/

Lab: MERL Technology Lab
Project Type: Research

Multiview Video Coding



We are working on developing advanced video compression algorithms for multiview video, i.e., video sequences recorded simultaneously from multiple cameras. We are also actively participating in the multiview video coding standardization activity in MPEG. Target applications for this work include 3D display and free viewpoint video.

Background and Objectives: The need for multiview video coding is driven by two recent technological developments: new 3D display technologies and the growing use of multi-camera arrays. A variety of companies are starting to produce 3D display technologies that do not require glasses and can be viewed by multiple people simultaneously. The immersive experience provided by these 3D displays are compelling and have the potential to create a growing market for 3D video and hence for multiview video compression. Furthermore, even with 2D displays, multi-camera arrays are increasingly being used to capture a scene from many angles. The resulting multiview data sets allow the viewer to observe a scene from any viewpoint and serve as another application of multiview video compression.

Technical Discussion: Our multiview video compression codec extends H.264/AVC to take advantage of correlations between different cameras. In standard video codecs, an important tool is motion compensated prediction where the encoder predicts the current frame from past or future frames in the same sequence. By coding only the resulting prediction error instead of the entire frame, significant savings are possible. In addition to standard temporal prediction, our multiview codec allows the encoder to predict the current frame from frames in other cameras or from virtual interpolated views. Specifically, we have modified the MPEG JSVM reference software to allow insertion of multiview frames into the Decoded Picture Buffer (DPB) and various reference lists. By decomposing the multi-camera sequence in various ways, we can obtain spatio-temporal prediction that is more efficient than pure temporal prediction. Furthermore, when camera parameters are available, we can interpolate a virtual view to use as a reference. For example, our codec can combine left and right views to interpolate a synthetic center view to use in predicting the center sequence. The interpolated views often perform better than temporal references.

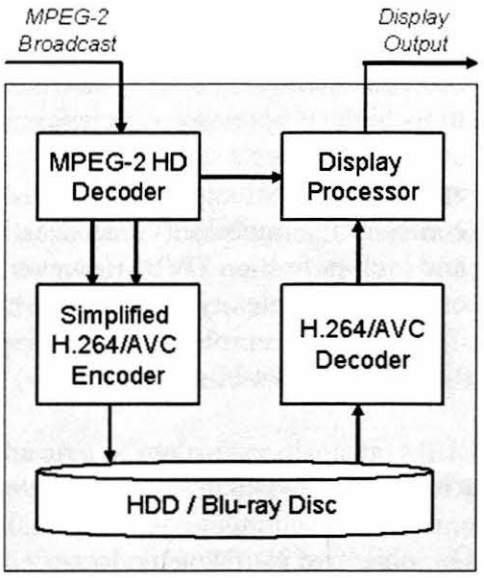
Collaboration: This project is done in collaboration with the Multimedia Information Coding & Transmission Technology Department at Johosoken.

Future Direction: We are working on improving the quality of our multiview codec and remain active in MPEG standardization activities including core experiments on buffer management, random access, and view synthesis.

Contact: Sehoon Yea
http://www.merl.com/projects/multiview-coding/

Lab: MERL Technology Lab
Project Type: Research

Efficient Storage of Broadcast Video



This work introduces a video recording system that employs MPEG-2 to H.264/AVC transcoding to achieve efficient storage of broadcast streams. Novel transcoding techniques have been developed for this purpose so that efficient conversion could be achieves with minimal increase in complexity.

Background and Objectives: MPEG-2 is the primary format for broadcast video, where the data rate for high-definition video is approximately 20Mbps. The latest video coding standard, referred to as H.264/AVC, is able to achieve the same quality as MPEG-2 with about half the data rate. Since the H.264/AVC format has been adopted into storage format standards, such as Blu-ray Disc, we expect H.264/AVC decoders to appear in consumer HDD systems soon. Certainly, as more high-definition content becomes available, long recording mode will be a key selling point for future HDD recorders. To satisfy this need, we are developing novel techniques that convert the MPEG-2 broadcast video to the more compact H.264/AVC format with low complexity.

Technical Discussion: In this project, we aim to transcode the incoming MPEG-2 bitstream to an H.264/AVC bitstream. The key to reducing the complexity is efficient motion re-estimation and mode decision, which would typically account for more than 80% of a full H.264/AVC encoder. Our current transcoder takes incoming motion vectors and coding modes from MPEG-2 bitstream and maps them in a simply way to the motion and modes of H.264/AVC. After mode and motion mapping, a small-range motion refinement around the mapped motion vector is performed. Extensive simulations show that the simple transcoder performed reasonably well with a fraction of the complexity of a full encoder.

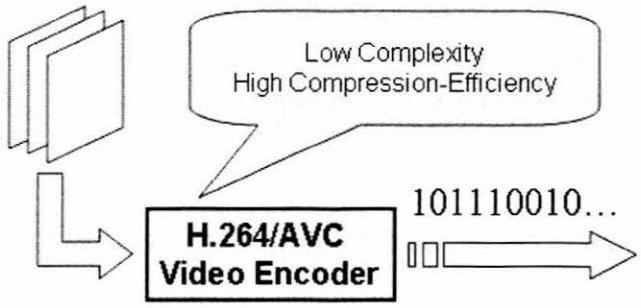
Collaboration: This project is done in collaboration with the Multimedia Information Coding & Transmission Technology Department at Johosoken.

Future Direction: Explore more advanced mode and motion mapping algorithms for the transcoder, and explore optimal combinations of H.264/AVC coding tools.

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

Lab: MERL Technology Lab
Project Type: Research

H.264/AVC Encoder Optimization



H.264/AVC is a recent major 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: We developed a hierarchical complexity control framework to efficiently manage the complexity of encoding process, with focus on motion estimation and mode decision algorithms. This framework may be used to develop a complexity scalable encoder. We also developed an efficient mode decision algorithm for intra-only H.264/AVC video coding. We exploited the correlation between optimal coding mode decisions of temporally adjacent pictures to reduce the computational efficiency of the encoding. Compared to conventional RD optimized mode decision algorithms, the proposed algorithm can significantly reduce the computational complexity with negligible loss of compression efficiency.

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

Future Direction: Explore further complexity reduction techniques, while maintaining high compression efficiency.

Contact: Jun Xin
http://www.merl.com/projects/avc-optimization/

Lab: MERL Technology Lab
Project Type: Research

Off-the-Destktop Interaction and Display

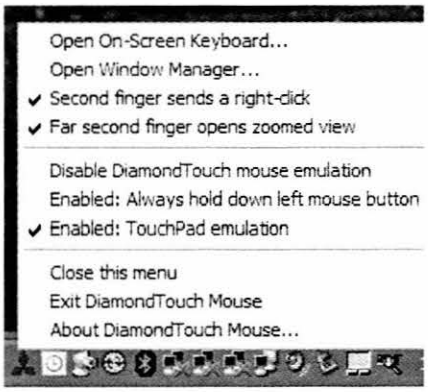
Computing devices are proliferating at all scales and sizes, from huge outdoor electronic signage and ever-larger digital televisions to navigation systems in automobiles and ever-smaller cell phones and projectors. Whereas desktop human-computer interfaces have become largely standardized, off-the-desktop devices have become the new frontier for human-computer interaction 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. This past year our research has focused on (1) multi-user touch technology and interactions for shared group displays and controls, (2) spoken-language interfaces for automotive and handheld devices, (3) task-based dialog technology for enabling human-robot communication as well as intelligent help systems, and (4) technologies for electronic displays.

In the area of multi-user displays, MERL plays a leading role in the international research community in its development of the DiamondTouch table and related software technologies. This year we extended the work into multi-surface group displays and also into shared controls. 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. Projects include SpokenQuery and its applications, Multilingual Dialog Tools, and ComBadge. Task and dialog representation underlie our work on human-robot interaction and DiamondHelp. Finally, projects in electronic displays include Saffron, a scalable font technology having worldwide impact, and Content-aware Media Browsing, incorporating concepts for unique presentation modes for high-definition TVs.

Project Descriptions

Mouse Emulation and Legacy Support for Multi-User/Multi-Touch Surfaces



DTMouse System Tray Context Menu

The DiamondTouch Mouse mouse emulation utility, one outcome of the Mouse Emulation and Legacy Application Support on Multi-User/Multi-Touch Surfaces effort (DTMouse), is provided with the DiamondTouch SDK to convert DiamondTouch table touch input to Windows system mouse input. DTMouse is a project to investigate issues related to providing backwards compatibility to (legacy) mouse-based software for users of a multi-user multi-touch DiamondTouch (DT) table. It is part of the broader DiamondTouch Applications project.

Background and Objectives: Contemporary operating systems fundamentally only allow the use of one mouse at a time. Only one window has mouse focus, only one text area can be selected for copying, etc. By providing support for legacy software, DTMouse enables arbitrary applications to run on DiamondTouch. It was originally developed as part of a remote collaboration system. The first standalone version of DTMouse offered arbitrary mapping of touch events to mouse events, and included a novel "2-Finger Resizing" capability for resizing arbitrary windows. Subsequent versions included support for laptop fingerpad emulation, multiple displays, zoomed views, and launching the onscreen keyboard. The current version supports fluidly conducting precision interactions with arbitrary applications, as well as a multi-user annotation tool for simultaneously scribbling on screenshots and then saving the annotations as image files.

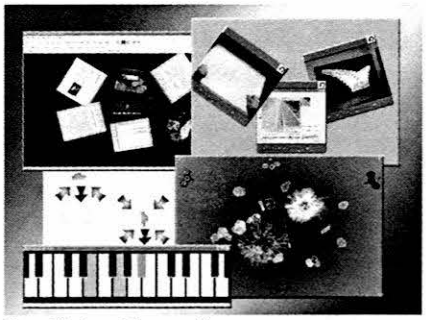
Technical Discussion: By default, the DiamondTouch Mouse utility lets any toucher act as the mouse. In the case of conflicts, the first to touch wins – it will ignore other touchers while the first toucher is interacting. Some of the challenges for realizing a mouse emulation system on a finger touch surface include: resolution, how to specify a particular pixel with a "fat" fingertip; obscured content, how to tell what is being touched if a finger is covering the content; movement vs. engagement, how to specify when the left mouse button is to be held down (just moving the mouse versus dragging with the mouse); timing and spacing, how two taps with a finger are not as temporally or spatially as close to each other as two clicks with a mouse; full mouse functionality, how to right-click, middle-click, and use the mouse wheel; integration, how to seamlessly incorporate this into existing operating systems user environments. This project also includes best techniques for automating existing software to support multi-touch (one user touching more than one place at a time), and multi-user (more than one user touching the surface at the same time) capabilities. These issues were addressed in the current software iteration through several mechanisms including support for touch modes, configuration settings, touch-friendly utilities, and O.S. integration. See the documentation for more details.

Future Direction: More study related to keyboard issues and advanced DT capabilities.

Contact: Alan Esenther, Kathy Ryall
http://www.merl.com/projects/dtmouse/

Lab: MERL Technology Lab
Project Type: Advanced Development

DTFlash: Rapid Prototyping for Multi-User Multi-Touch Applications



See Color Figure 9

DTFlash (DiamondTouch-Flash) is a project to develop next-generation rapid prototyping support that leverages an authoring environment rather than a more general-purpose programming environment. DTFlash will advance research into hardware and software to support multiple users interacting simultaneously with two hands on the same surface, and is designed to flush out multi-user interaction aspects more quickly and expose new areas of focus.

Background and Objectives: Our research into multi-user/multi-touch tabletop application development (DiamondTouch Applications) revealed significant shortcomings of traditional tools and development environments when developing for advanced input devices. The DiamondTouch SDK provides a low-level C API for accessing data about which touchers are touching a surface at which places. Early research into building on top of this SDK focused on providing an API based on a general purpose programming environment such as Java or .NET. DTFlash takes a different direction by leveraging the Macromedia Flash authoring environment to emphasize authoring over programming. For example, the standard Flash authoring tool can be used to create arbitrary shapes or objects which can then simply be marked as being draggable or rotatable. Literally no coding is needed, yet the new content is "multi-toucher-aware", allowing multiple people to interact with different shapes or objects at the same time.

Technical Discussion: Through our earlier work, we found that a rapid prototyping tool for multi-user/multi-touch applications requires fundamental low-level support for a variety of items: simultaneous users; multiple points of input from each user; an authoring environment for creating "multi-touch aware" content; multimedia support; the ability to simulate multiple touchers and touchpoints with a mouse and keyboard; debug-mode overlays for visualizing toucher information. DTFlash provides these capabilities, in part by defining primitive touch events, enhanced primitive events, and methods for semantic operations.

Also of note, DTFlash applications work as regular web pages, allowing for simple deployment and ushering in a new dimension of multi-user enabled web pages which eliminate the need to take turns with the mouse. Flash is also based on vector graphics and optimized for small downloads, so DTFlash applications have a small memory footprint. But it is the reliance on weak static typing and it's "expressiveness" which make Flash particularly well-suited for exploring drastic changes without breaking existing applications, and for facilitating the creation of complex and novel visual interfaces.

Collaboration: Indiana University, MDEA

Future Direction: Continue development of core code and explorations of interesting areas already flushed out. Exploring open source collaborations.

Contact: Alan Esenther, Kathy Ryall
http://www.merl.com/projects/dtflash/

Lab: MERL Technology Lab
Project Type: Initial Investigation

DiamondTouch

The MERL DiamondTouch table is a multi-user, debris-tolerant, touch-and-gesture-activated screen for supporting small group collaboration. The DiamondTouch table is available commercially as a developer's kit and includes: a selection of demonstration applications; a mouse emulator with onscreen keyboard to support common Windows applications; and a Software Developer's Kit allowing the development of new software applications that support gesture inputs and multiple simultaneous users.

Background and Objectives: Computers are good at facilitating collaboration between people in remote locations. However, when people are in the same place at the same time, computers tend to get in the way. DiamondTouch supports small group collaboration by providing a display interface that allows users to maintain eye contact while interacting with the display simultaneously (i.e., without having to take turns). DiamondTouch was first created in 2001 as an experimental multi-user interface device. Over the past few years, MERL has manufactured about 100 DiamondTouch tables, lending them to universities and research organizations around the world.

Technical Discussion: DiamondTouch is front-projected and uses an array of antennas embedded in the touch surface. Each antenna transmits a unique signal. Each user has a separate receiver, connected to the user capacitively, typically through the user's chair. When a user touches the surface, antennas near the touch point couple an extremely small amount of signal through the user's body and to the receiver. This unique touch technology supports multiple touches by a single user two handed touch gestures and distinguishes between simultaneous inputs from multiple users. DiamondTouch tables are available in two sizes (32" diagonal and 42" diagonal display), while custom sizes and shapes are available on spec.

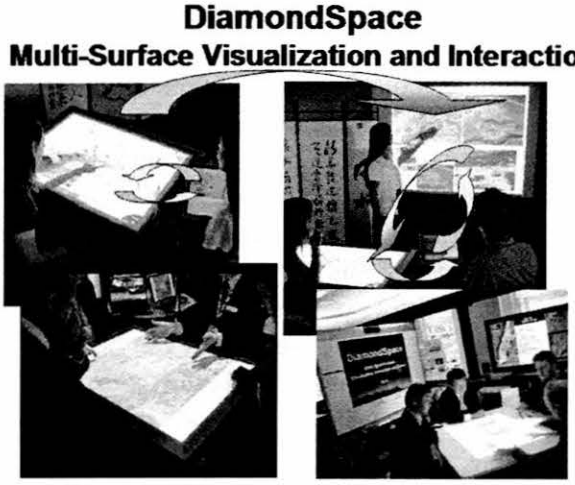
Collaboration: Within Mitsubishi Electric, MERL is working with Johosoken and Kamaden so that DiamondTouch may be included in contracts for the Electronic Systems business unit. Also, MERL is collaborating with many universities and corporations in researching further applications.

Future Direction: MERL will be conducting business incubation over the next few years, selling the DiamondTouch table as a developer's kit, along with several product options (projector, support packages, etc) with the intention of building product demand prior to transitioning the commercial business to Mitsubishi Electric.

Contact: Adam Bogue
http://www.merl.com/projects/DiamondTouch/

Lab: MERL Research Lab
Project Type: Advanced Development

DiamondSpace



See Color Figure 10

Fluid interaction and meaningful visualization are the key for multi-surface, multi-device, interactive spaces to become the users' true cognitive prosthesis. Such interactive environments, local or remotely connected, are suitable for a variety of application domains, including collaborative business and work spaces, emergency response and mission control centers, urban planning rooms, as well as geospatial visualization and analysis.

Background and Objectives: While incorporating many of our related research outcome and findings, including PDH, DiamondTouch, DiamondSpin, and UbiTable, DiamondSpace uniquely addresses the research challenge of how multiple interactive display surfaces, including multi-touch multi-user tabletops, electronic walls, and laptop/desktop displays, can be used in concert to create flexible visualization and workspaces in which people can explore, understand, utilize and manipulate information from many data sources and types.

Technical Discussion: Our research is situated within a three-pronged agenda:
(1) Multi-surface Visualization and Interaction (MVI): We are developing three distinct conceptual models of data sharing, visualization, manipulation and application interactions across multiple surfaces: independent (e.g., MultiSpace), reflective and coordinated/multi-view.
(2) Interaction Techniques: We have designed and developed a number of novel interaction techniques including CoR2Ds, a new type of draggable popup widgets for tools, menu and data, and Glimpse, a multi-level touch input model for real-time data exploration. Our ongoing research includes freehand touch gestures, as well as interaction techniques that span across multiple displays and interaction surfaces.
(3) Evaluation: We conduct evaluations, exploring the efficacy of new techniques, as well as the fundamental properties of touch surfaces. We have studied the nature and role of non-speech audio feedback in such settings, gradients of privacy and sharing, and overall impact on group interactions.

Contact: Chia Shen
http://www.merl.com/projects/dspace/

Lab: MERL Research Lab
Project Type: Research

DT Controls

See Color Figure 11

DT Controls is a new technique for instrumenting physical controls in multi-user systems. It can detect which controls each user is actuating at any given time. Therefore, the behavior of the system can depend on the identity of the user activating any control. Examples of per-user behavior differences include restricting access for some users, and control of different devices from a single physical control. In addition, the usage state information can be used to generate a very detailed audit journal.

Background and Objectives: DT Controls works by placing a uniquely identifiable signal near the surface of each control. Each user has a separate receiver on or near their person. When a user approaches a control, the signal is capacitively coupled through the user to that user's receiver. By examining the received signal, the system can determine which controls that user is currently near. DT Controls is very inexpensive and can be added to almost any kind of physical control, including push buttons, knobs, switches, dials, and touch screens.

Technical Discussion: Applications under consideration include:

(1) Automobile controls - For User Safety: By using DT Controls, Navigation system functions may be safely enabled for the passenger while the vehicle is in motion. For Cost reduction: DT Controls may be used to reduce the overall number of controls in a vehicle, and to provide per-user access control.

(2) Airplane cockpits and Train controls - DT Controls will add user information to black box journal. User-based access control can enforce use protocols. Touch duration data can be useful for black box and for control panel design and user training/evaluation.

(3) Control Rooms - User-based access control can assign different permissions to workers and supervisors, or to different members of a team based on their role. Per-user actuation and touch duration data can provide detailed use journal for safety as well as for control panel design and user training/evaluation.

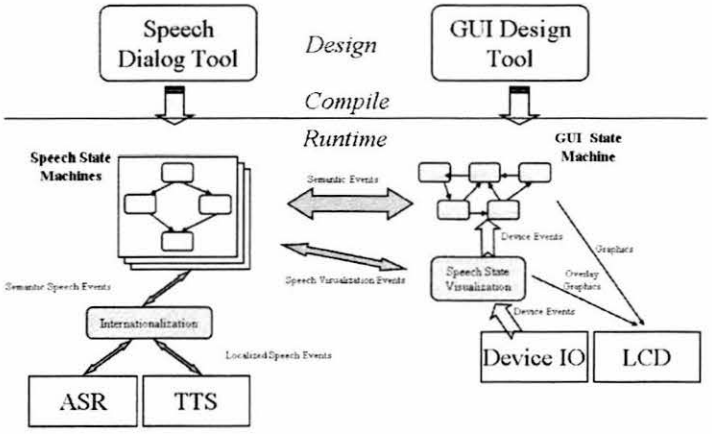
Collaboration: MERL is working with IDC to build novel multi-user interfaces using DT Controls, for a variety of business applications.

Future Direction: MERL will continue development of DT Controls by building several kinds of prototype control panels and studying the usability and functionality for a variety of application domains.

Contact: Paul Dietz, Bret Harsham
http://www.merl.com/projects/dtcontrols/

Lab: MERL Research Lab
Project Type: Advanced Development

Speech Dialog Tools for Automotive



The aim of this project is to develop a set of design and runtime tools that reduce the development and maintenance costs of building a speech dialog system for an automotive deployment.

Background and Objectives: Historically, the cost of building speech systems for cars has been very high. There are a number of reasons for this: 1) Multi-language deployments: car applications are frequently deployed in more than one language. The usual design goal is that the look and feel be substantially the same for each deployed language. 2) Distributed design: The specification and design responsibilities for these systems frequently reside in a separate entity from that responsible for the implementation. The ASR and TTS supplier is yet another entity. Late design or API modifications by any one entity can have very complicated effects on the implementation.

Technical Discussion: We will be building a set of modular tools which treat application specific speech dialogs as individual state machines. The speech state machines will be human language independent. Input and output language translation will be provided on a language specific basis. This will make the interface essentially the same regardless of the deployment locale.

Separate tool modes will allow separation of speech implementation and application implementation functions. This will lower development cost and time to market by reducing complexity. Furthermore, the tools will support multi-modal (speech and touch) input.

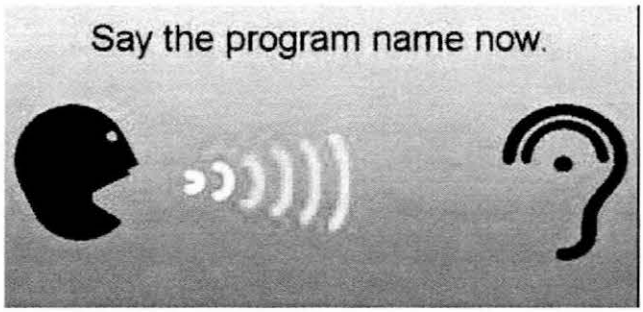
Collaboration: This is a joint project between MERL, Shakai-se, Johosken, and Sentansoken.

Future Direction: Future enhancements will include testing and verification tools, as well as additional design functionality.

Contact: Bret Harsham, Bent Schmidt-Nielsen
http://www.merl.com/projects/dialog4cars/

Lab: MERL Technology Lab
Project Type: Advanced Development

TV Content Search Using Voice



As the availability of television content continues to grow, a more effective method to search for programs is needed. We have developed a system that allows television viewers to search an electronic program guide (EPG) by voice. Our proposed solution is to add a microphone to remote controls to enable voice input for searching over variable collections of programs available through EPGs.

Background and Objectives: Consumers find television programs from a growing list of diverse sources including broadcast, cable, satellite, video-on-demand, Internet and disks. The need for effective search of content is growing as the number of choices for TV viewing and/or recording explodes. Existing search methods, such as text entry using conventional TV remote controls is awkward. This project aims to provide a more effective search using voice input.

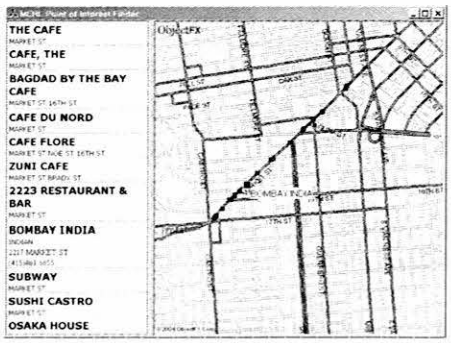
Technical Discussion: Our solution adds a microphone to remote controls that allows the user to quickly find programs or browse the nearest matches using voice. The system utilizes MERL's SpokenQuery technology in a new domain of EPG search. SpokenQuery allows search terms to be entered that are unrestricted by vocabulary or grammar and finds the best matches, even if some words deviate from the title. When searching EPG programs users may not know the full program name, or the recognizer or audio environment may lose a word. This allows us to take advantage of SpokenQuery's strength in combining probabilities over all the terms of a query and over all of the EPG entries to provide the user the best matches. To help the user understand this behavior we have combined SpokenQuery with result highlighting according to relevance.

Future Direction: We plan to refine the prototype and conduct further user testing. Advanced voice browsing techniques will also be investigated.

Contact: Anthony Vetro
http://www.merl.com/projects/epg-search/

Lab: MERL Technology Lab
Project Type: Research

SpokenQuery



This year the Speech group at MERL has delivered an R&D toolkit to MELCO clients and is nearing completion of an embeddable, productized C++ library and SpokenQuery server. We have also developed interactive speech-collecting applications that support prototyping and usability testing. With Point-of-Interest Finder we have collected a new database of unscripted, interactive speech queries.

Background and Objectives: Often end-users need access to information that they know is available on a device they're using, but their access is slowed or limited due to interface limitations. With SpokenQuery, users access such information simply by speaking. The SpokenQuery project supports the delivery of easy-to-use, speech-based search-and-selection capabilities in embedded application environments.

Technical Discussion: SpokenQuery is a patented information-retrieval technology that lets users perform searches using speech. SpokenQuery extends best-of-breed ASR (automatic speech recognition) engines with sophisticated post-processing, allowing natural, free-form spoken requests for information. Using SpokenQuery is like Googling for something you need to locate, but you say descriptive words in any order instead of by typing. Automotive-domain examples include iPod music selection and navigation destination entry.

Our work is focused on three main areas: 1) delivering software tools, libraries that allow MELCO BU's to develop embedded applications with SpokenQuery; 2) research and development on core ASR and SpokenQuery technology, including fundamental research on ASR usage, document retrieval using ASR results, and the collection and analysis of speech data from realistic interactive settings; 3) development of methodologies and applications for user studies that focus on improving interfaces and work-flow.

This year we delivered the Java-based SpokenQuery Toolkit 1.0 to MELCO clients to support their development of SpokenQuery-based applications. A embeddable C++ library and server instance was also developed. We have used the Point-of-Interest Finder desktop application to collect realistic speech data from users locating destinations. We also developed SpeakPod, an application for interactively selecting and playing music from an iPod.

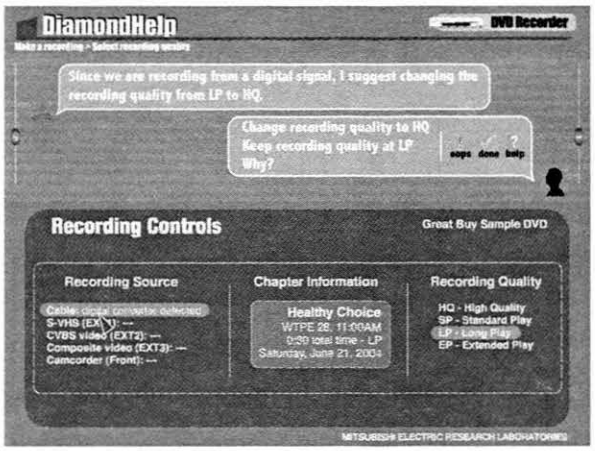
Collaboration: Shakaise/Taden, Johosoken, and Sentansoken.

Future Direction: The long-term goals of our research effort are improved task-completion and user-satisfaction rates and better tools and libraries for application developers. Our ongoing analysis of recently-collected speech data will improve fundamental SpokenQuery modeling and search technology. We are developing a dynamic domain-data update capability, as well as integrated support for mixing speech queries with speech commands.

Contact: Hugh Secker-Walker
http://www.merl.com/projects/SpokenQuery/

Lab: MERL Technology Lab
Project Type: Advanced Development

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 uses Collagen to track task context and to manage the conversational part of the interface.

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.

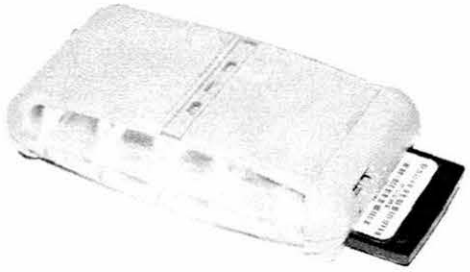
Collaboration: This research is contributing to the Display Conversation Technology (eActive) portion of the Interactive User Interface Platform project at Johosoken.

Future Direction: DiamondHelp is also contributing to the development of an industry standard for task model representation (CEA-2018) through Dr. Rich's co-chairmanship of the CEA Working Group (R7WG12) on Task-Based User Interface.

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

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.

Background and Objectives: This project is developing a communication device that can 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 in segments of the market where cell phone penetration is lowest, such as in the developing world.

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 – we currently support both English and Tamil (an Indian language) interfaces. 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, is used at all stages of the network and messages are delivered over the Internet.

Asynchronous messaging also has advantages for users. The device can be very small, since it has no display and 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 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.

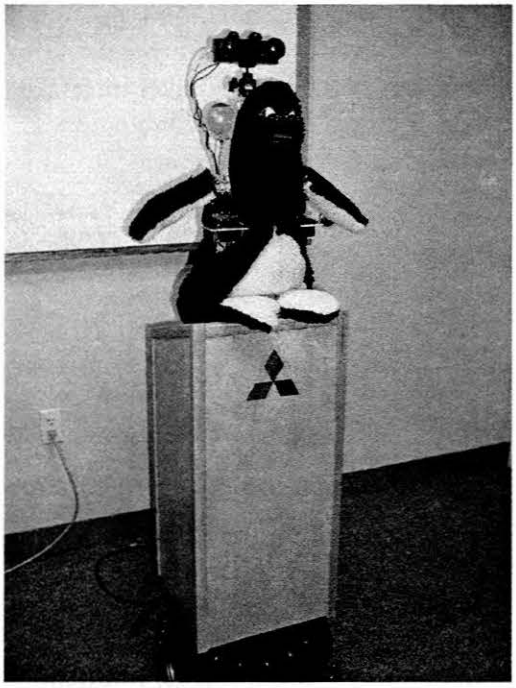
Collaboration: Through collaboration with Eric Brewer and the TIER project (Technology and Infrastructure for Emerging Regions) at UC Berkeley, we conducted a user study in Pudupalayam village in Tamil Nadu in India during August 2005. This work and its continuation is used to assess the technical, social, and business aspects of ComBadge.

Future Direction: We are investigating using the ComBadge hardware as a platform for other projects, such as sensor network nodes, radio implementations, and location aware systems. Also, we are considering using ComBadges in a mesh-connected environment in which no infrastructure is required -- intermediary ComBadges would form the network.

Contact: James L. Frankel
http://www.merl.com/projects/ComBadge/

Lab: MERL Research Lab
Project Type: Research

Human-Robot Interaction for Hosting Activities



We are developing a collaborative robot that can conversations with people and can perform and interpret physical gestures and movement during the interaction, thereby "engaging" the person.

Background and Objectives: We are investigating engagement in human-robot interaction by studying human-human engagement, developing rules to apply to human-robot interaction, and applying those rules to a functional robot. 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, it helps perform tasks, and it keeps track of the location of the people in the home. 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 tasks of demonstrating itself or the iGlassware system developed at MERL. People interacting with the robot do not need any training to interact. In its current form, the robot interprets head gestures to determine if the person is attending to itself and to objects of importance in the environment. It recognizes and responds to the normal head nods that people make during a conversation to indicate agreement or understanding of what is being said. The robot is fully mobile using a Pioneer II mobile base. Over 100 people have completed demonstrations with the robot. These studies show that (1) people nod more often if the robot returns their nodes, (2) they find the robot's gestures more natural than an unmoving conversational partner, and (3) they direct their attention more to the robot than the unmoving conversational partner.

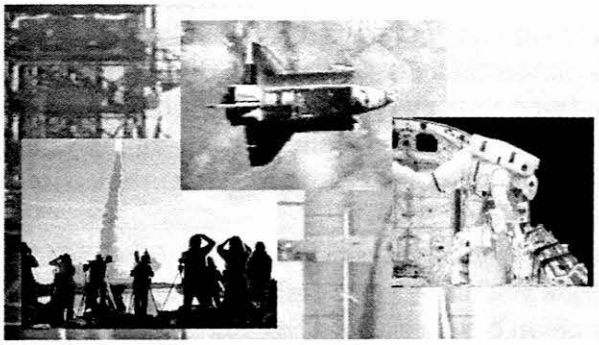
Collaboration: We have collaborated with members of the MIT Vision Group (http://www.ai.mit.edu/projects/vip/projects.htm) to use the Watson vision system and support vector machine learning algorithms to interpret head nods of people interacting with the robot.

Future Direction: The robot is being programmed to use vision recognition software to find a person in a group, and, using its mobile system, approach that person. Its decision to interact will depend on its own social and politeness parameters and what activities it perceives the person to be undertaking.

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

Lab: MERL Research Lab
Project Type: Research

Content-aware Media Browsing & Digital Video Playback Interface



The goal of this project is to improve the entertainment value of watching a recorded program by displaying that program in a novel, exciting way. We first decompose the structure of the video, and then present that video in a manner consistent with its structure. This manner of video playback is especially relevant to low bit-rate, low-quality video rendered on a high-resolution television.

Background and Objectives: While televisions, projectors, and computer monitors have become physically larger and include an increased number of pixels, the manner in which videos are displayed on these surfaces has remained the same. While newer content has a higher resolution, videos originally produced for smaller display surfaces or those streamed over a low-bandwidth network are simply scaled to fill the larger display surface. Compression artifacts and low-resolution result in poor image quality. Three assumptions that conventional video players make is that they should never display more than one frame from the same video at any one time, they should never display the same frame from a video in multiple locations on the screen, and that they should never move the presented content around the display. Our prototype digital video playback interface converts an input video to an output video with the aim of challenging the above assumptions about video playback.

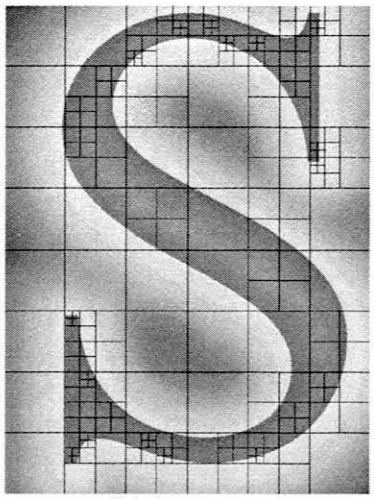
Technical Discussion: The input video is a series of frames that are normally displayed sequentially. The output video is the same series of frames that have been scaled, rotated, filtered, and displayed in parallel on different regions of the display in a manner that not only preserves the continuity of the story, but also supports the structure of the video. The manner in which the frames are selected, the length of the frames, and the treatment of previously displayed frames are based on the structure of the input video. We determine the structure by using a variety of known techniques in conjunction with a new method for scene detection to find the relationship between shots, the content of individual shots, and camera motion. By displaying the frames of a video in this manner, the context of the video is reflected in its presentation, and the viewing experience is arguably enhanced.

Future Direction: In addition to presenting videos in a manner consistent with their style and content, we are beginning to explore the presentation of low-resolution videos with very little analysis of the structure of the video. Without video analysis, our playback techniques may be more appropriate for the hardware constraints of consumer video devices.

Contact: Clifton Forlines
http://www.merl.com/projects/tricky/

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.

Saffron has been licensed to Macromedia (now Adobe) and is shipping in Flash Player 8. The Adobe Flash Player reaches close to 100% of all Internet-enabled desktops. Saffron has also been licensed to Monotype Imaging, a leading provider of fonts and font technologies, who will be integrating Saffron into their future products.

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, Kyoto Works; Sarah Frisken, Professor, Computer Science, Tufts University.

Future Direction: Continued collaboration with Jun Someya to develop a Saffron-based ASIC for MELCO products. A major focus of current development is to adapt Saffron for use in embedded applications and on mobile devices. Future development is focused on a new font representation using strokes.

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

Lab: MERL Research Lab
Project Type: Research

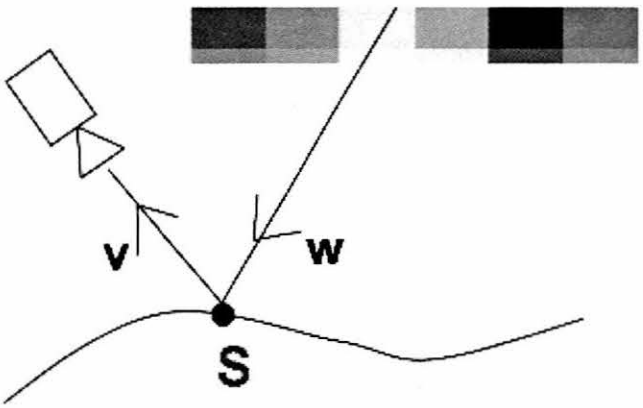
Sensor and Data Systems

Traditional computer applications have largely presented the user with content that has been processed according to fixed or human-generated algorithms. However, it is now becoming possible for a computer system to automatically collect data from the local environment, 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 create and support the technology itself to applications aimed at growing new businesses and enhancing current ones. Projects span the gamut from sensor hardware, which can autonomously and inexpensively collect data, to new systems for communicating that data, to algorithms for uncovering subtle underlying interrelations and trends in the data. Infrastructure cost is a limiting consideration for most applications of these technologies, and MERL's researchers have kept this in mind so that the new applications will be able to use commodity electronic and electronic and computing hardware, and be cost-effectively deployed.

Project Descriptions

3D Sensor for Specular Objects



See Color Figure 12

A new generation of robot cell systems is about to appear. Existing systems are inflexible, in particular the robot arm typically only retrieves parts that are presented at a preset location and orientation. The next generation of robot cells will have more human-like capabilities - to retrieve unsorted parts from bins, and to deal with both rigid and non-rigid parts. The two key challenges are making the robot cell quick and easy to deploy, and making it able to deal with all types of parts because any human interaction puts a bottleneck on the whole process. This project is on a new 3D sensor for specular objects.

Background and Objectives: The background for this project is the creation of robot cells that can compete economically with a human assembly cell. The overall objective is the creation of sensors to compute the pose of parts in bins around the workspace, so that a robot arm can subsequently retrieve the parts. The specific objective is to create a sensor that works on specular objects, such as mirror-surface objects or transparent reflective objects.

Technical Discussion: Technology for scanning specular objects is less well-advanced than for scanning objects with matte surfaces. We are building a system that uses controlled-reflection to compute 3D data for a specular surface. The sensor computes surface normals directly, and we are investigating ways to directly obtain pose from this information without going to full 3D shape. The advantages of our approach are the ability to handle varied shapes, and an expected ability to handle hybrid specular-diffuse surfaces for some degree of diffuseness.

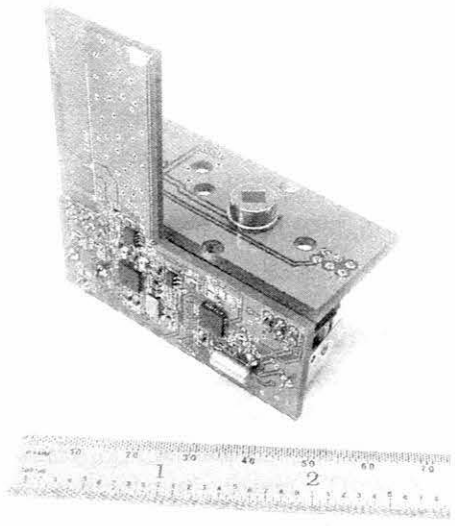
Collaboration: Dr Okuda, ATC.

Future Direction: ATC is working on a sensor for matte surface parts. MERL is working on a scanner for specular parts. In future, both groups will consider how to extend their work to a sensor for hybrid specular-diffuse parts.

Contact: Paul Beardsley
http://www.merl.com/projects/spec-sensor/

Lab: MERL Research Lab
Project Type: Initial Investigation

ROCkS: Reduced Operating Cost Sensors



ROCkS is an effort to create MERL-owned, standards-compliant, manufacturable, flexible, and low-power sensor network platform.

Background and Objectives: The primary goal is to demonstrate low-cost, ad hoc installation of sensor networks within buildings by utilizing standards-compliant wireless communication and multi-year battery life. With its modular architecture and easy to manufacture design it will readily serve as a flexible research platform for sensor network research at MERL.

Technical Discussion: The current implementation being used at MERL for building-wide sensor network research is built around technology licensed from MIT. ROCkS, the second-generation sensor box being designed at MERL, improves on the MIT design in terms of power consumption, radio compatibility, and manufacturability.

The ROCkS were designed specifically with power consumption in mind. An MSP430 microcontroller was chosen for its flexible clocking modes, which allow the software to enter various 'sleep' states and thus conserve power. This low power consumption profile means that the sensors can operate of small batteries for years at a time, and opens up to possibility of investigations into parasitic power models. The ability to install networks of sensors without the need for wires to deliver power represents a significant reduction in the cost of installation for the networks.

The new boxes employ an IEEE 802.15.4 radio; this is the physical layer typically used with Zigbee. These new radios are similar in concept to the radios used in the MIT boxes, but they are standards-compliant and therefore capable of interoperating with equipment designed by many different vendors.

The board layout is designed to be easy to manufacture and assemble, further reducing total cost of deployment. Modular design also means that experiments can leverage the work on the processor and communication systems for a multitude of other sensor modalities.

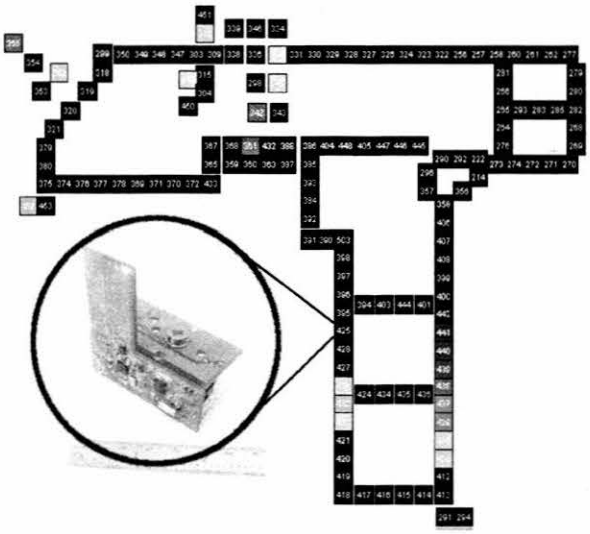
Collaboration: Shinsuke Azuma of Johoso-ken; Location-Based Services SK; Integrated Civil Security SK.

Future Direction: Currently we are working to deploy approximately 50 sensors at MERL as a first step beyond the prototype stage. Future directions may include deployment in the home environment, experiments with parasitic power, or different senor modalities.

Contact: Christopher R. Wren, Darren Leigh
http://www.merl.com/projects/rocks/

Lab: MERL Research Lab
Project Type: Research

Ambient Intelligence for Better Buildings



Sensor networks have the potential to allow the development of truly intelligent buildings that improve productivity, efficiency, safety, and security. To be practical, such networks must be efficient, scalable to very large spaces, and economical to manufacture, install and maintain. One answer is networks of passive infrared motion detectors. Sensors could be manufactured onto building infrastructure elements that Mitsubishi Electric manufactures, such as light fixtures. We are developing technology that enables cost-effective networks to recognize, predict, and index human activity in building-scale environments.

Background and Objectives: A key component of the cost for such systems is the installation, configuration, and operating costs. Sensor networks present complex perceptual challenges. It is imperative that the systems we develop can be installed by electricians, ventilation engineers, and system integrators. In "Functional Calibration for Pan-Tilt-Zoom Cameras in Hybrid Sensor Networks" we present a completely self-calibrating system that drives a pan-tilt-zoom camera from a simple sensor network. This reduces operational cost by eliminating the need for a human operator to be dedicated to the camera. It reduces installation cost by eliminating the need for a skilled technician to perform a complex calibration.

Technical Discussion: That is just one example of how a very low-cost sensor network can make many building systems more efficient. The sensor network is an economical way to capture context that is essential to the efficient operation of building services but was previously considered too costly to capture. We believe that there are significant implications for the energy efficiency, elevator operation, safety, and security in buildings. To test these claims we have installed a network of 155 sensors at MERL and are collecting a large, unique database from that space.

Collaboration: Shinsuke Azuma of Johoso-ken, Location-Based Services SK, Integrated Civil Security SK.

Future Direction: Currently we are building a large database that includes day-to-day, week-to-week, and month-to-month patterns of activity at MERL. This database includes uncommon events such as floor cleaning and fire alarms, as well as unique events such as high-level visits and office renovations. Future work involves developing and testing appropriate models for these macroscopic phenomena. We also plan to pursue work with Li-hon and Biru-hon on next-generation infrastructure.

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

Lab: MERL Research Lab
Project Type: Research

Scalable Activity Recognition for Sensor Networks

This project demonstrates that with an appropriate analysis methodology, motion-sensor networks are capable of providing useful contextual information to building services. The methodology we propose supports scalability by adopting a hierarchical framework that splits computation into localized, distributed tasks. We call this framework Scalable Activity Recognition for Sensor Networks (SARSEN).

Background and Objectives: There is locality in building context. The way hallways and intersections are used is the same in all buildings. The way those larger chunks fit together share commonalities, but also begin to have localized meaning. The overall structure of the building is often unique. If we are going to build robust systems that understand how people use building, then we need to account for these realities. The systems need to understand the building blocks of context. Those building blocks can be built and tested in the lab with some confidence that they will be portable. At the same time, systems need to be easily configured to the unique realities of each building.

As a secondary benefit, this locality can also be reflected in the computational structure. If it is advantageous to engineer the system as a mesh network with local computation, the SARSEN framework supports that engineering choice. Locality of context is reflected in the locality of computation. Local data can be aggregated in the nodes and summarized before it is passed along. This local analysis reduces communication overhead by summarizing data locally and only communicating the summaries. Also, the local models are computationally well suited to the simple processors likely to be embedded in a network.

Technical Discussion: Current empirical results include localization of meeting rooms, gathering points, and building resources such as printers. These results prove that the SARSEN framework is capable of correctly analyzing the complex behavior in our laboratory. Out analysis employs a database of over four million discrete events recorded over several months on the 8th floor of MERL from 155 sensors. Data collection is ongoing and data is currently being shared with Johosoken. The sensors are the product of the Reduced Operating Cost Sensors (ROCkS) projects.

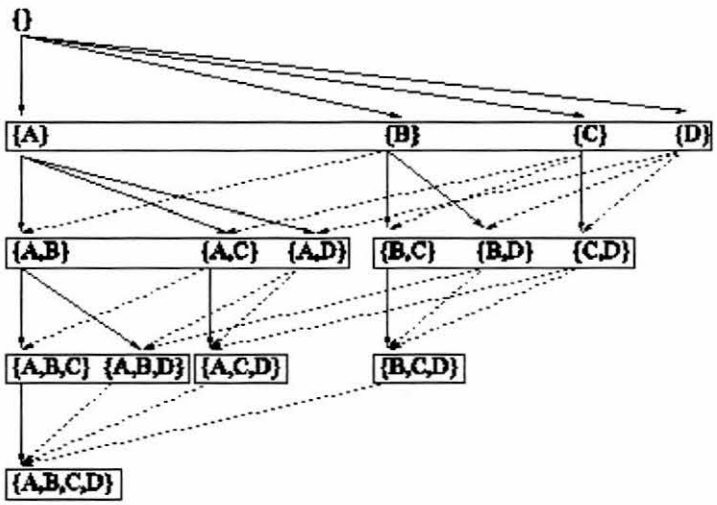
Collaboration: Shinsuke Azuma of Johoso-ken, Location-Based Services SK, Integrated Civil Security SK.

Future Direction: The current deployment of 155 sensors on the 8th floor of MERL allowed us to test the framework, and the initial results are very promising. We hope to be able to deploy sensors into other environments, such as different kinds of offices, homes, retail spaces, and industrial spaces to test the claims of portability to new environments.

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

Lab: MERL Research Lab
Project Type: Research

Induction of Compact Recommendation Trees



The availability of massive retail databases of past customer purchases has opened the possibility for personalized product recommendation tailored to the personal preferences of individual customers. We have found a very scalable solution for discovery of compact recommendation policies represented in the form of decision trees. The solution leverages an algorithm for efficient discovery of all frequent item-sets corresponding to groups of purchases, followed by a standard algorithm for induction of decision trees which uses the discovered frequent item-sets as training examples. The solution can work out-of-core and can handle purchase databases of practically unlimited size.

Background and Objectives: Previous work has shown that algorithms for personalized recommendation based on frequent item-set mining can be very effective when customer preferences are Boolean variables, for example when a database of past purchases is available. However, this approach to recommendation usually results in a huge number of association rules that are hard to maintain in memory and slow to search for each new customer. A more compact and easier to understand representation of the recommendation policy is needed.

Technical Discussion: The collection of frequent item-sets discovered by the rule mining algorithm defines implicitly a recommendation policy. Our solution is to explicitly enumerate this policy in the form of a mapping between a particular purchasing history and the optimal product to be recommended for this history. This turns the problem into a classification task, and the mapping can be used as training examples. We use standard algorithms for induction of decision trees to learn an explicit representation of the recommendation policy in the form of a binary decision tree. By combining two well known data mining algorithms - one for discovery of frequent item-sets, and the other for induction of decision trees, we achieve excellent scalability and complete out-of-core operation.

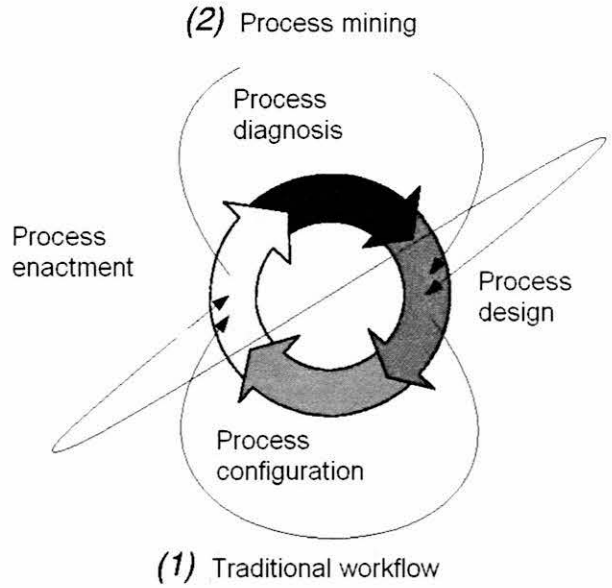
Collaboration: Johosoken/IS-hon

Future Direction: It remains to be seen whether the same level of compression can be achieved with sequential recommendation policies, when the order of purchases influences significantly purchase probabilities.

Contact: Daniel Nikovski
http://www.merl.com/projects/fimrec/

Lab: MERL Technology Lab
Project Type: Research

Business Process Management Software



Business Process Management (BPM) software is one of the fastest growing segments of the enterprise software market, due to the key role it plays in corporate re-engineering. Using BPM software tools, business people can document the workflow and processes in their enterprises, identify bottlenecks and other impediments to efficiency, and suggest alternative and improved business processes. MERL has been supported the development of Johosoken's BPM suite by implementing a discrete-event simulator for process analysis in FY05, and is working on business process mining software in FY06.

Background and Objectives: The purpose of BPM software is to support the documentation, analysis, monitoring, and re-design of the business processes in an enterprise. MERL has participated in the implementation of Johosoken's BPM suite with two modules: a discrete-event simulator for interactive analysis of business processes (FY05), and a process mining module that builds a model of a business process in standard notation from execution logs of that process (FY06).

Technical Discussion: To a large degree, the two modules are complementary: while the simulator produces synthetic execution logs from a given BP model, the process mining software produces a synthetic BP model given actual execution logs. In spite of this complementarity, however, the technology and technical challenges behind the two modules are completely different. The discrete event simulator uses efficient random number generation and management of event sequences in priority queues, while the main challenge in process mining is to select the most appropriate BP model among those that successfully explain the execution log. To this end, we use fundamental machine learning principles and methods such as Occam's razor, minimum-description length, and version spaces.

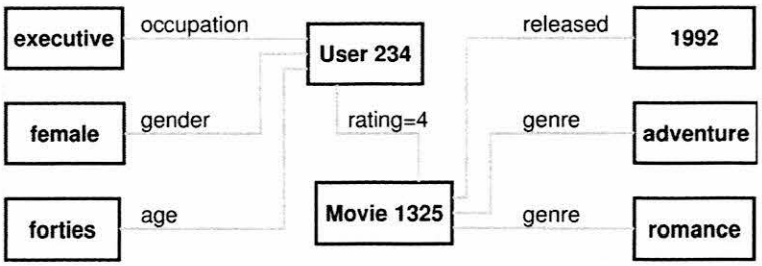
Collaboration: Johosoken (I-Shi-bu), Shakai-hon

Future Direction: We are exploring alternative representation of business process, such as Markov models, Petri nets, and their associated BP representation languages and search spaces. A more distant, and much more challenging goal is automated process re-engineering so as to maximize pre-specified key performance indicators such as cycle time, resource utilization, etc.

Contact: Daniel Nikovski
http://www.merl.com/projects/bpm/

Lab: MERL Technology Lab
Project Type: Research

Data Mining by Random Walks



Whereas search engines currently rank matches to a query by popularity, we show how to rank matches directly by relevance to the query, and even incorporate side information about the person making the query. Experimentally, this framework is shown to be very effective in recommending products to consumers, and can be used to optimize any combination of consumer satisfaction and vendor profit.

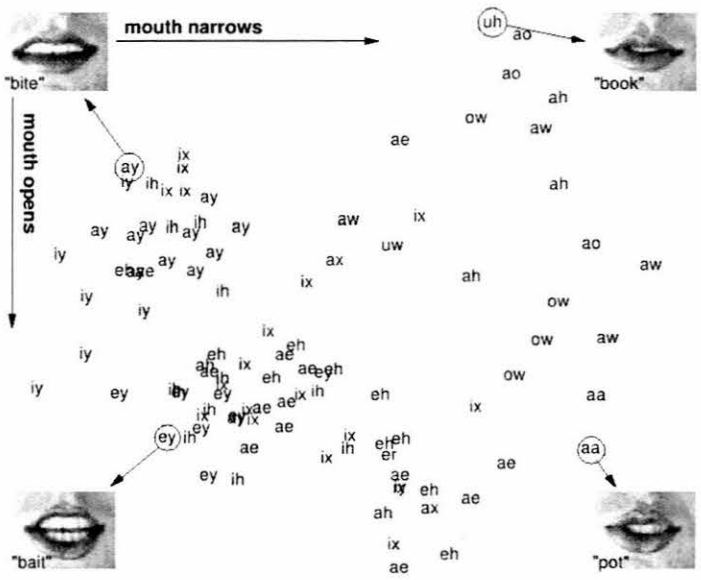
Background and Objectives: All search engines and collaborative filtering systems make recommendations on the basis of similarity and popularity. Popularity is not necessarily a good predictor of relevance. In the most sophisticated systems, the similarity metric is derived from random walks on the data. We adapt this framework to fuse information from transactional data (e.g., consumer-product purchasing records) and relational background data (e.g., consumer demographics and product attributes). We then seek to make the recommendation that is most relevant to a consumer's interests, regardless of popularity.

Technical Discussion: To find the most relevant items for a query, we view a relational database as a graph. The expected travel times of a random walks on this graph yield a natural distance metric related to the probability of association between any two items. We compute inner products between the query and all other items in a space consistent with that metric. There is a formal equivalence to viewing the database as a circuit and measuring conductance. From the inner product we can compute expected commute time, which is a measure of similarity, and random-walks cosine correlation, which is a measure of relevance that is almost invariant to sampling artifacts. We develop several recurrences that give all needed inner products in linear time. These recurrences also make it feasible to compute the expected discounted profit of recommendations when some profit margins are known for some items.

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

Lab: MERL Research Lab
Project Type: Research

Dimensionality Reduction



Nonlinear dimensionality reduction (NLDR) is a method of compressing high-dimensional data by mapping data points to locations on a low-dimensional manifold. We have developed fast methods for discovering the manifold and smooth compression/decompression algorithms.

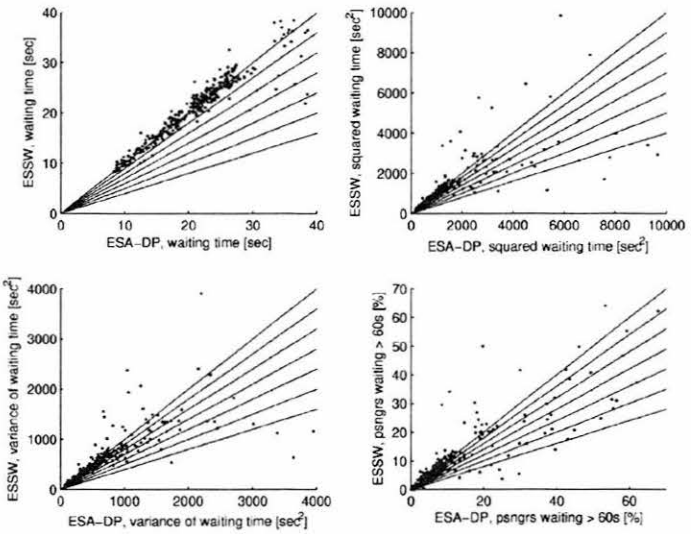
Background and Objectives: Images, video, and sound are examples of high-dimensional data streams that are observed to reside mainly on low-dimensional manifolds. Our goal is to replace linear dimensionality reduction methods such as the Karhunen-Loeve transform with nonlinear methods that are more efficient and isolating the intrinsic variation of the dataset.

Technical Discussion: Using our previous work in Geodesic Nullspace Analysis, we contruct a linear basis for a low-dimensional data embedding and its low-frequency vibrations. This can be done in linear time. We then construct and solve a semidefinite program (SDP) for an affine projection of the linear basis that preserves interpoint distances. This yields a locally isometric embedding of the data. The SDP has high complexity but has negligible size compared to the original problem. This entire framework can be projected into the space of possible radial-basis functions (RBFs) and used to solve for the RBF that gives the most isometric compression of the data. The result is the first NLDR algorithm that scales to millions of points, and works on datasets large enough to be useful for data-intensive applications like speech recognition. The above image shows embedding of acoustic feature vectors, yielding a representation that improves classification rates of spoken vowels.

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

Lab: MERL Research Lab
Project Type: Research

Group Elevator Control



We have developed decision-theoretic solutions for (western-style) re-assignment elevator scheduling and minimal-energy elevator scheduling.

Background and Objectives: Optimal assigning multiple elevator cars to pick up multiple waiting passengers is probably an NP-hard problem. In "Japanese-style" scheduling, all assignments are final. In "western-style" re-assignment scheduling, all assignments can be revised opportunistically. Although a much harder problem, advances in this area will be useful for increasing market share in the U.S. and E.U.

Technical Discussion: The optimal approach is to calculate the exact costs of a decision, marginalizing over all future scenarios compatible with unknown or uncertain variables. These variables include future passengers, unspecified destinations of current passengers, and the number of waiting passengers. The full calculation is intractable but in recent years we have found efficient algorithms to compute most of these marginalizations in the Japanese setting. This year we found a set of proxy costs that allow the "Japanese-style" solution to be used in the inner loop of a very efficient branch-and-bound solution for the "Western-style" re-assignment problem. We ported this algorithm to "ELEVATE 6.0", which is now the most commonly used elevator simulator and control testbed in the industry. MERL's western-style algorithm strongly outperforms all other controllers available on this platform, reducing waiting times by 10-40% and increasing the throughput of each shaft.

In an unrelated and mainly academic result, we have also developed an exact solution for the minimal-energy schedule, assuming symmetric up and down energy costs.

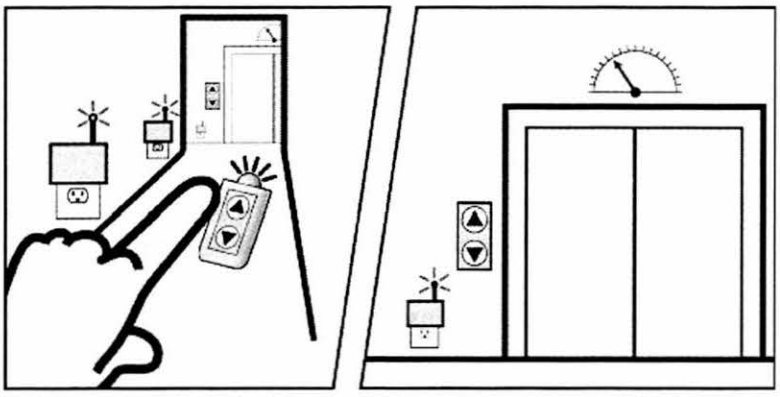
Collaboration: Dietmar Ebner, TU Wien; Jakub Czerny, Charles U (Prague)

Future Direction: We are seeking a partner for technology transfer within MELCO.

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

Lab: MERL Research Lab
Project Type: Research

Advance Elevator Call



This project is investigating the technology to allow elevator users to request an elevator to take them to another floor in advance of walking to the elevator lobby. This allows the elevator scheduler to better service the users and to more efficiently utilize the elevator cars.

Background and Objectives: At MERL, we have already developed elevator scheduling algorithms that are able to optimally schedule passengers to elevators at the time of the call to minimize their waiting time. This project will allow better scheduling by making the call information available earlier.

Technical Discussion: A prototype mains-powered wireless mesh network has been designed and deployed at MERL that receives requests from small battery-powered devices, or "fobs." When a button is pressed on a fob, our system determines the location of that fob using the signal strength received from it at a number of the mains-powered nodes. Each node in our wireless mesh network that receives a request from a fob forwards that request to a "primary location determination node" along with the ID of the originating fob, the ID of the receiving node, and the signal strength of the originating fob at the receiving node. Based on this information along with the known locations of the receiving nodes, the "primary location determination node" roughly locates the fob. The location together with the usual path, distance, and corresponding walking time is used to determine the approximate arrival time at the nearest elevator lobby. This information is forwarded to the elevator scheduler to improve elevator car arrival times and utilization. We are currently designing custom hardware to replace our current prototype system which will include keychain fobs (similar to remote car door unlocking fobs) with an "up" and a "down" button corresponding to the requested elevator direction.

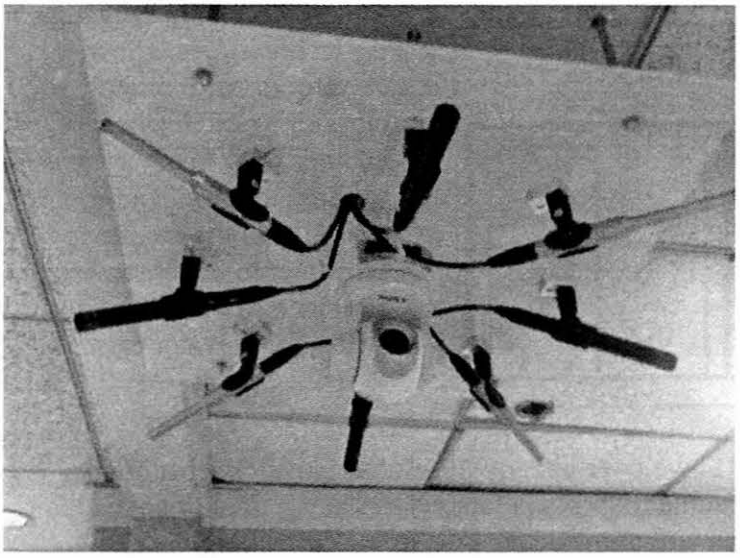
Collaboration: The information gained through this system will be made available to Matt Brand's best-of-breed elevator scheduler to allow even better scheduling.

Future Direction: We will also be investigating other location-aware applications, such as systems in which phone calls can follow the recipient.

Contact: James L. Frankel
http://www.merl.com/projects/elevatorcall/

Lab: MERL Research Lab
Project Type: Research

Audio-Assisted Cameras



The goal of this project is to enhance the functionality of cameras and overcome some of their fundamental limitations by performing analysis of their audio environment. This type if integration can have a significant impact on visual data mining with applications in surveillance, entertainment media, and human-computer interaction.

Background and Objectives: Cameras today are constrained by a limited field of vision, poor lighting conditions and bandwidth and storage limitations. The ability to point a camera at the right spot only at the right time is a desirable feature. To this end we can analyze the audio environment of a camera and determine whether certain target sounds have occurred and also determine their originating position. Armed with this information a camera can pan to the right position and record and illuminate only for the specified time frame thereby conserving power and bandwidth and also overcoming locality and temporal constraints. The same technology can be also applied to multi-camera systems where a human operator can be alerted to the proper monitor feed as directed by audio analysis of the relevant camera's environment.

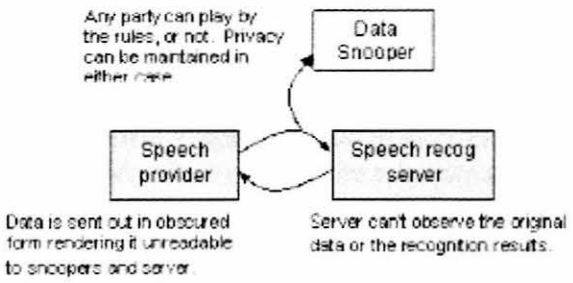
Technical Discussion: Sound recognition for camera assistance has already been used for previous projects in traffic monitoring and elevator security surveillance. By employing our sound recognition framework we can reliably detect a set of target sounds that require the attention of a camera. We have augmented this technology with custom algorithms that can learn when a sound originates from a set of target positions, and/or to arbitrarily localize sounds in a 3-dimensional space so that we can point a pan/tilt/zoom camera towards the desired activity. The system is fairly flexible so that it can operate on a very constrained environment where it is only interested in a small set of sounds and locations, or it can be more relaxed and curiously looking towards the most prominent sounds. Prototypes of this work have been built and are operating in real-time at MERL's lobby.

Future Direction: Future directions include integration with enhanced versions of the sound recognition framework that work with sound mixtures and a suite of applications that enhance the camera stream with audio derived information to assist multi-camera parsing by human operators.

Contact: Paris Smaragdis
http://www.merl.com/projects/aac/

Lab: MERL Research Lab
Project Type: Initial Investigation

Secret Audio



The goal of this project is to enable a client-server model for speech recognition while ensuring no private information leaking. Such a system can allow service providers to operate on private speech data without them directly observing the data or the recognition results. This can ensure that the party providing the speech data will keep the sent information and its transcription private.

Background and Objectives: In client-server processing models there is always the issue of trust. Privacy minded parties such as major corporations or government institutions often have vast amounts of speech data whose recognition they would like to outsource to data processing servers. However they cannot allow potentially sensitive information to be shared openly like that. This work addresses this issue and proves that it is possible to perform speech recognition as a server for another party without inferring any information from the speech data or seeing the results. Conversely we show that the data providing party will not obtain information on the recognizer models the server has used. Such a framework can enable a new business model for secure off-site speech processing and audio surveillance.

Technical Discussion: Using computational primitives from the field of secure multiparty computation we can construct arbitrarily complicated speech recognition and processing systems. We have constructed a process that can allow the training and execution of hidden Markov models in such a framework so that the nature of the data and the trained models are only known to their respective owners and are not compromised by the recognition process. This work is based on well-known cryptographic principles and is shown to be secure. The speech data is sent out on a garbled form that renders them unreadable but retains key mathematical properties. The server can operate on the data and return results that are similarly garbled as the received speech data. The speech provider is the only party that has the information to un-garble the results. The server and potential data snoopers cannot see the original speech data, or their transcription. Likewise the provider of the speech data cannot observe the speech recognition models of the server. Even in the case where one of the parties does not conform to the protocol and has malicious intent it is possible to still guarantee data privacy for both parties.

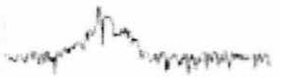
Future Direction: In the near future we are focused in finishing a practical implementation of the secret audio system. Further work will be on formulating secure methods for other speech and audio processing applications that can be outsourced and support such a client-server model.

Contact: Paris Smaragdis
http://www.merl.com/projects/secretaudio/

Lab: MERL Research Lab
Project Type: Research

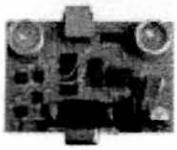
Acoustic Doppler Sensors for Surveillance

Doppler spectrum at peak stride



Doppler spectrum at mid stride

Complete spectrogram of a walking person



The MERL Doppler Radar

In this project we have developed an inexpensive Doppler sensor (costing less than $5 to manufacture in bulk), and technology that has been able to identify a person walking in a corridor towards or away from the sensor with over 90% accuracy in early tests. We are currently improving the technology to be able to identify a larger set of subjects, and to function in more open spaces.

Background and Objectives: Doppler acoustic sensors have long been used for measuring speeds of moving objects, as well as for motion sensing. Our goal is to utilize the sensing capabilities of a Doppler acoustic radar to develop a highly inexpensive, yet effective surveillance device that can be used both in isolation or to complement alternate surveillance devices such as cameras.

Technical Discussion: Doppler sensors can measure the velocities of moving objects. Various parts of articulated objects such as humans and animals have different velocities. These velocities change with time. A Doppler sensor can measure all these velocities and represent them as the spectrum of a reflected signal. The time-varying pattern of the velocities of the hands and legs of walking person, and thereby the temporal pattern of the short-time Fourier spectrum of a Doppler signal reflected off the person, is characteristic of the person. We use these spectra to identify the person. The Doppler based surveillance technology can be used by itself, or, more effectively, to augment other surveillance sensors such as cameras and microphones.

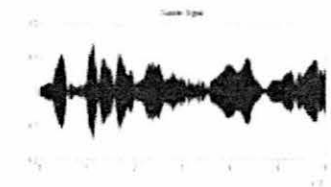
Collaboration: We are currently collaborating with MIT to develop portable devices that heterodyne high-frequency Doppler signals into audio-range frequencies, such that signals can be captured using conventional sound cards on a PC.

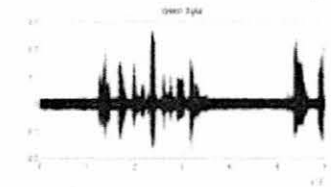
Future Direction: In the past year we have developed algorithms to extract the movement information from the Doppler signal as complex time-domain signals with precise, identifiable signatures. Future directions include the utilization of the extracted signature to enable identification of a larger set of subjects or motion classes than is currently possible. The current technology is suitable for constrained spaces such as corridors. We are also working on the incorporation of signatures from multiple sensors, in order to account for direction of approach and enable operation in open spaces.

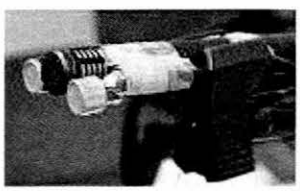
Contact: Bhiksha Raj
http://www.merl.com/projects/dopplersurveillance/

Lab: MERL Research Lab
Project Type: Research

Acoustic Doppler for Denoising Speech Signals







We have developed an acoustic Doppler based secondary sensing mechanism for improved denoising of noisy speech signals. The denoising mechanism can be utilized in diverse applications such as for denoising in cellphones or denoising of signals captured by hands-free farfield microphones, for either storage and transmission or for improved speech recognition.

Background and Objectives: Doppler acoustic radars are effective mechanisms to capture the motion of the mouth and facial tissue, when deployed at a short distance from the face. Such Doppler measurements are relatively unaffected by conventional acoustic noise and represent robust, if coarse, secondary measurements of the speech-production mechanism. Our objective is to harness these measurements for improved denoising of speech signals.

Technical Discussion: A Doppler-enhanced microphone such as the one in the figure incorporates an ultrasound transmitter, a high-frequency receiver and a conventional microphone. The transmitter emits an ultra-sound tone, and the receiver captures reflections of this tone from the speaker's face. The reflected signal undergoes a Doppler shift in frequency. The frequency shift and the energy of the reflected signal carry information about the motion of the speaker's face and mouth.
We have developed algorithms that can extract the signal signatures of the movements of the speaker's mouth from the complex time-varying Doppler spectra. We are able to detect speech activity accurately through the correlation of these signatures with recorded audio signals, even in very noisy conditions. The signatures are also correlated, albeit relatively weakly, with the actual sounds being uttered. They can hence be used to aid in the identification of the underlying sound to achieve improved denoising and speech recognition.

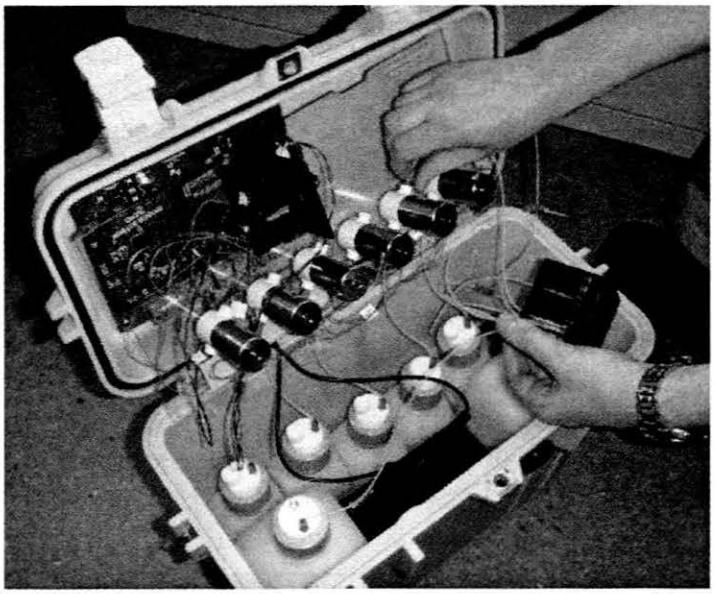
Collaboration: We are currently collaborating with MIT to develop portable devices that heterodyne high-frequency Doppler signals into audio-range frequencies, such that signals can be captured using conventional sound cards on a PC. MIT will deploy our technology in information kiosks at the Stata Center.

Future Direction: In the past year we have developed improved algorithms for extracting the motion signatures in the Doppler signal. We are currently able to achieve high-accuracy voice activity detection. Future work will be aimed at evaluating the voice activity detection under recording conditions that are normally considered extremely difficult, and the development of algorithms to improve VAD performance in such scenarios. We will also work on the utilization of the Doppler signatures for speaker-independent denoising of speech signals and speech recognition.

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

Lab: MERL Research Lab
Project Type: Research

LED-Based Sensors



By using a novel reverse-bias circuit, an LED can become a very sensitive and precise detector of light intensity. Alone, this can be used as a frost or liquid level detector. By adding chemical reagents, we can also detect chemicals or pollutants at the part-per-million level

Background and Objectives: LED-based detection of photons is now well-established in the literature, and we are moving forward with a fully autonomous, low-cost, low-power deployable pollution detection system based on LED sensors.

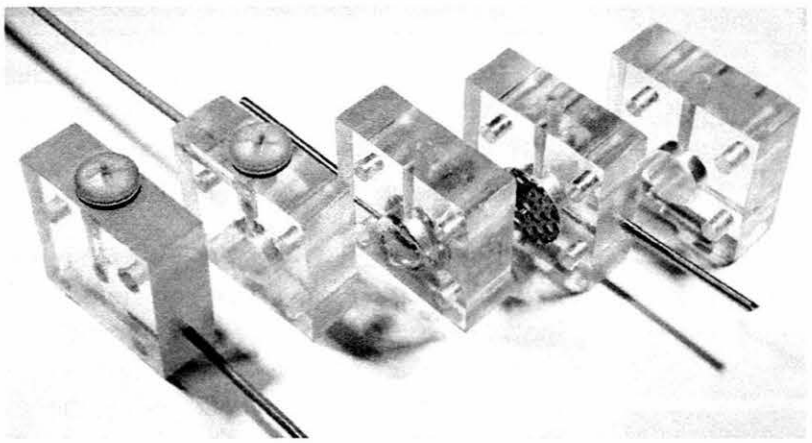
Technical Discussion: Our endurance-test LED system is now approaching two years of continuous operation (24 x 7 x 365 operation). The new test will use LED-based sensing to detect phosphate water pollution in rivers and streams. The goal is to build a phosphate water pollution sensor node that can give a phosphate reading every hour, for a full year, without requiring human attention, using solar power, LED sensing, microfluidic pumps, and GSM data modems to relay the data to a distant host.

Collaboration: In cooperation with Dublin City University and the National Centre for Sensor Research (Prof. Dermot Diamond, director).

Future Direction: We will do the test deployment of the phosphate sensor in Ireland's National Botanical Garden in the near future, and then go to a "deep" field test at the Ireland Environmental Protection Agency's sites.

Contact: William Yerazunis **Lab:** MERL Research Lab
http://www.merl.com/projects/LED_chemical_sensors/ **Project Type:** Advanced Development

Microfluidic Pumps using Fuel-Cell Actuation



By using the fuel-cell reaction in both forward and reverse, we can generate actuation pressures in excess of 5 atmospheres.

Background and Objectives: The goal of this work is to explore how fuel-cell reactions (water and electricity being converted back and forth to hydrogen and oxygen at pressure). We can use this pressure to operate any hydraulic or pneumatic device. Our first goal is to use a diaphragm pump, but any pressure-operated device (such as a valve, latch, or piston) can be operated.

Technical Discussion: In order to build small (or disposable) pollution sensors, we often need to pump very small amounts of sample and reagent liquids. To do this, one of the least expensive ways is to use the reversible fuel cell reaction. At very small pump sizes, fuel cell based pumps can be up to five times more efficient than a mechanical pump, and also a factor of ten less expensive. At larger scales, electromagnetic pumps become equally efficient; the largest pump we have built will drive approximately 100 cc/minute but is no more efficient than a commercial pump at that rate. Another benefit of a fuel-cell pump is that it can be operated "closed loop"; by using optical sensors to observe the hydrogen/water border we can control flow much more accurately than most mechanical pumps can, especially in situations of varying output pressure.

We are researching the realities of using these fuel-cell actuators for actual sensor systems, primarily as pumps, but secondarily as valves or mechanical actuators.

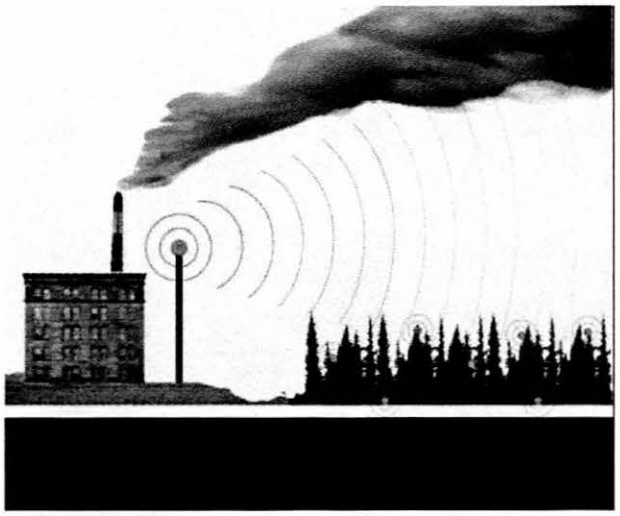
Collaboration: Dublin City University, National Centre for Sensor Research (Prof. Dermot Diamond).

Future Direction: Our next step is to build a fuel-cell pump that is sufficiently dependable to be integrated into a field-testable sensor system.

Contact: William Yerazunis, John Barnwell
http://www.merl.com/projects/micropump/

Lab: MERL Research Lab
Project Type: Research

Low-Power, Low-Bandwidth Communication for Sensor Networks



We are developing a communication system optimized for sensor networks that require extremely low power consumption and cost. The system will have a range of 100 meters, a data rate of 1 bit per second and a sensor node cost of about one dollar.

Background and Objectives: Most sensor networks that we have been envisioning have lax requirements for data rate and latency, but are severely constrained by power consumption and cost. While the sensor itself requires power and costs money, we can design the communication system to require less power and cost less money, so that the system will be limited by the sensors and not communication.

Technical Discussion: To limit the power consumption and cost of sensor nodes we use "backscatter" transmission to send data to a central hub. Backscatter is similar to using a signal mirror to send information: the mirror itself transmits no light, instead it modulates light from another source toward the receiver. Backscatter does this with radio, using a switched antenna instead of a movable mirror. This way information can be transmitted to the hub using a negligible amount of power -- the circuit will have a complexity and power consumption comparable to a digital watch. Our prototype operates in the 900 MHz ISM band. The hub transmits one watt of unmodulated RF power, which each sensor node reflects back toward it in a modulated fashion. The antenna switching happens at low modulation frequencies (~100 KHz) and not the much higher RF frequency (900 MHz), allowing the use of inexpensive electronics. The hub receives the modulated signal and decodes it into the data from each sensor node. Backscatter is very old technology used in radar and RFID systems. What distinguishes our system from others is achieving much greater range (up to 100 meters) with lowest-cost sensor communications circuitry. Some features we are developing to achieve these capabilities are: a very high dynamic range receiver, special cancellation circuitry (to prevent the hub's transmitter from jamming its co-located receiver) and novel modulation schemes.

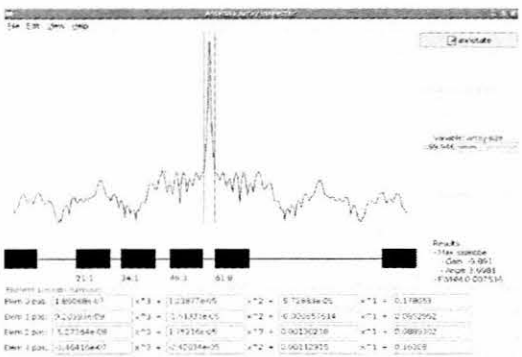
Collaboration: Some of the technology and circuitry from this project has been shown to the Microwave Department at Johosoken, who are interested in it for their work.

Future Direction: In the near future we will deploy a sensor network which uses the backscatter communication system. This will allow us to gauge its true effectiveness and cost in a realistic environment.

Contact: Darren Leigh, Joseph Katz
http://www.merl.com/projects/waesn/

Lab: MERL Research Lab
Project Type: Research

Human-Guided Antenna Design



Optimization-based approaches to antenna design have enjoyed limited success. The task is often computationally intractable 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: FY05 efforts built on our FY04 results, again addressing the issue of phased-antenna arrays with non-uniform spacing. The FY04 results provided a statistical background and framework for such design. In FY05 we used machine learning techniques to explore the space of antennas with the goal of rapid optimization. While in previous approaches we used exhaustive searches, which carry a large computational burden, we have now developed a new approach that results in significantly faster and "smarter" search, and also low computational complexity for fieldable applications. The trade-off for this new approach is the inability to guarantee that our results are truly optimal. However, using the statistics obtained from our FY04 project, we have shown that our new search results are only slightly sub-optimal. Such sub-optimal results are acceptable in physically realizable applications given the increased computational efficiency of the new methods.

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 extend our approach to handle two-dimensional, sparse phased-array antennas.

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

Lab: MERL Technology Lab
Project Type: Research

Color Figures

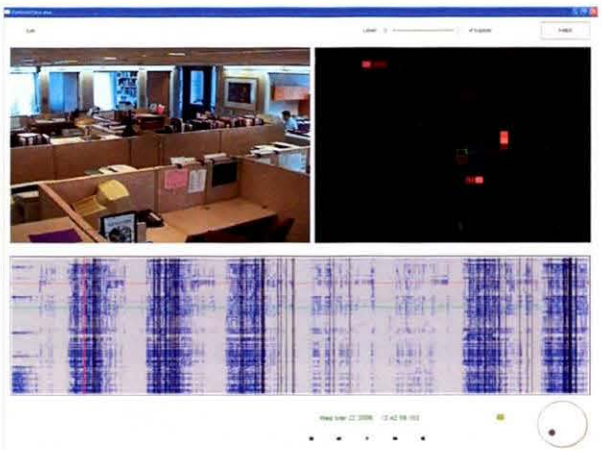


Figure 1 – see page 67

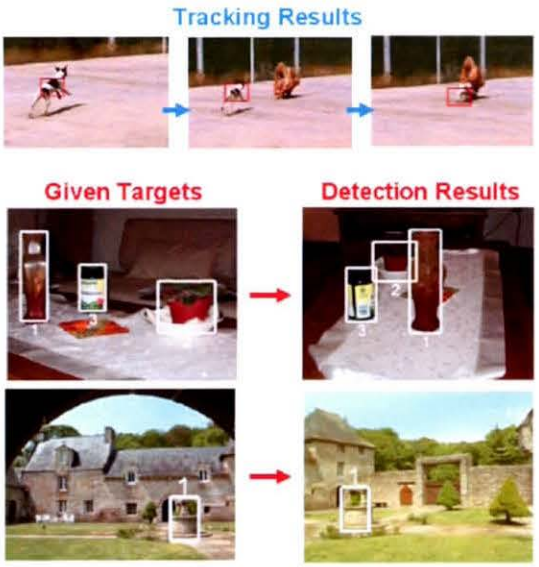


Figure 2 – see page 68

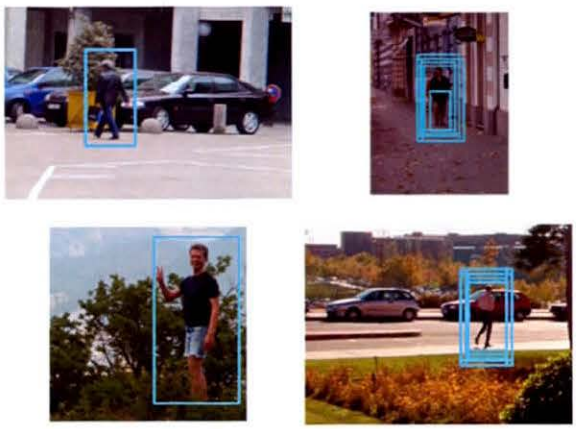


Figure 3 – see page 69

Figure 4 – see page 70

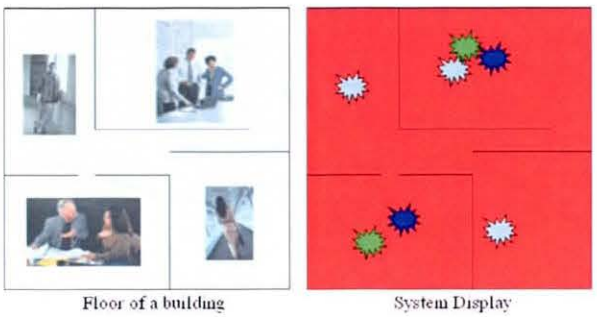


Figure 5 – see page 86

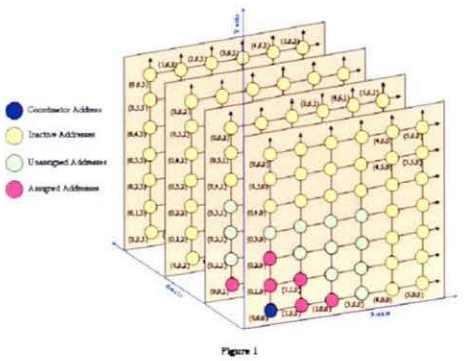


Figure 6 – see page 96

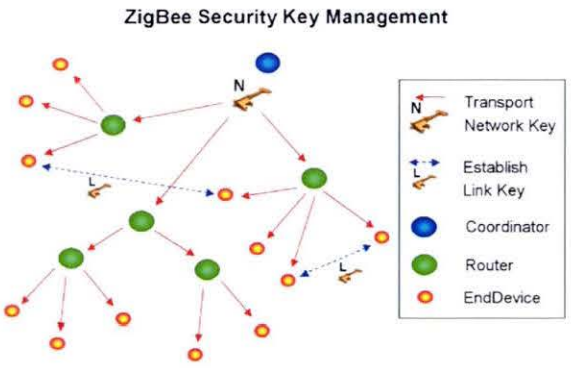


Figure 7 – see page 98

Figure 8 – see page 107

Figure 9 – see page 115

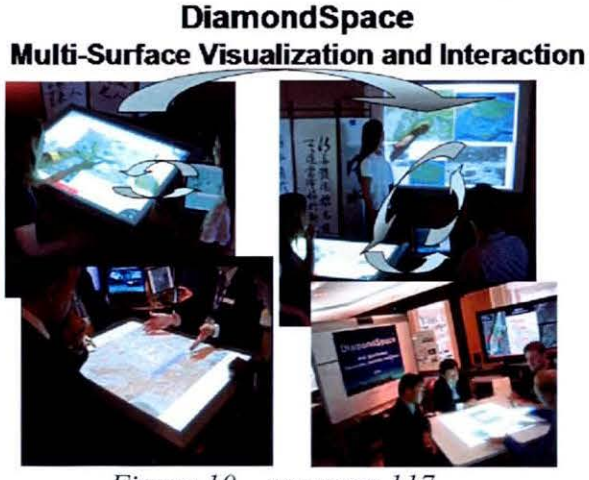


Figure 10 – see page 117

Figure 11 – see page 118

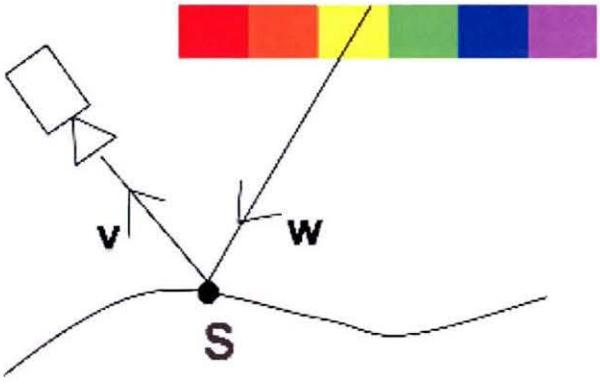


Figure 12 – see page 128