Mitsubishi Electric Research Laboratories (MERL)

Annual Report

July 2004 through June 2005

TR2005-00

(Published August 2005)

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 1

Awards and Commendations 9

Business Impacts 11

- Saffron Digital Typography 11
- Time Synchronization in IEEE 802.15.4b 13
- DiamondTouch in a Command and Control System 15
- Automatic Image Alignment for Electronic Whiteboards 16
- Down-Decoding in DTV Receiver Chip 18
- Post-Filtering for Professional HDTV Codec 20
- Multi-Projector Display System for Curved Surfaces 21

Technical Staff 23

Publications 37

Project Reports 49

- Computer Vision 51
- Digital Communications 73
- Digital Video 99
- Off-The-Desktop Interaction and Display 111
- Sensor and Data Systems 129

Color Figures 145

Mitsubishi Electric Research Laboratories

Mitsubishi Electric Research Laboratories (MERL) is the North American arm 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 networks and wireless transmission technology.
Digital Video – featuring encoding, decoding and analysis of video.
Off-the-Desktop Interaction and Display – featuring novel devices and interface concepts.
Sensor and Data Systems – featuring novel sensors, communication and system architectures.

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

We are strongly involved in the R&D community and standards activities, maintaining long-standing cooperative relationships with a number of research universities including MIT, CMU, 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
  Mr. Masatoshi Kameyama (EVP, CFO & CLO)
- Research Lab - *Basic & Applied Research, staff 21*
  Dr. Joe Marks (VP & Director)
  Dr. Joeseph Katz (VP & Deputy Director)
- Technology Lab - *Applied Research & Advanced Development, staff 43*
  Dr. Kent Wittenburg (VP & Director)
  Dr. Huifang Sun (VP, Research Fellow & Deputy Director)
- Marketing & Business Development
  Mr. Adam Bogue (VP)

Richard C. (Dick) Waters *Ph.D., MIT, 1978*
President, Chief Executive Officer & Research Fellow

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

Masatoshi Kameyama *B.S., Tokyo Institute of Technology, 1975*
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 & Research Fellow; Deputy Director Technology Lab

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.

MERL's Two Labs

MERL is organized as two labs because the style of management required for long-range research projects is quite different from the style of management required by shorter-range (advanced) development projects.

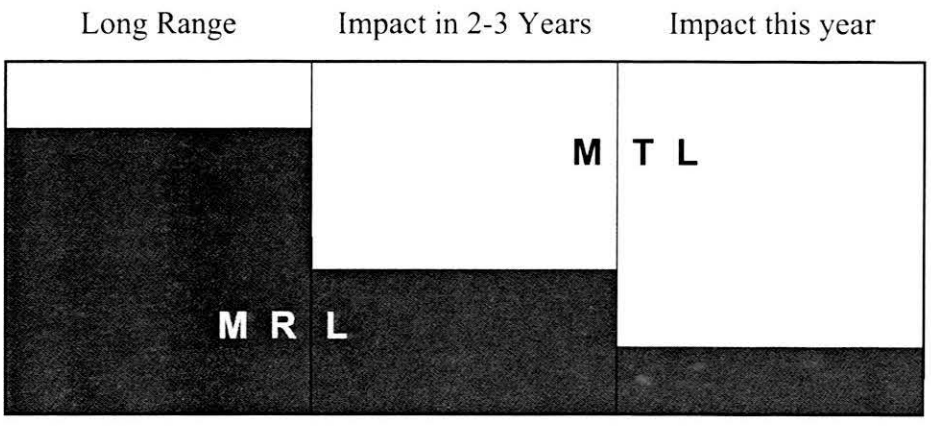
Long-range research projects are all about novelty. If it's research and it isn't novel, it doesn't matter whether it's ever done. In contrast, shorter-range projects are focused on meeting goals agreed with customers. If such a project doesn't meet its agreed goals, it doesn't matter how novel it is.

MERL's Research Lab (MRL) is primarily organized to foster novelty and generate new technology that can become part of future MELCO plans. The key elements of this are investigating as many potential ideas as possible and rapidly discarding those that seem least promising. This requires an emphasis on the individual pursuit of novel ideas. The best projects in this environment are ones with a potential for large benefit even if they have a high chance of failure and an indefinite time frame.

In contrast, MERL's Technology Lab (MTL) is primarily organized to deliver on projects that are part of larger MELCO plans. The key elements here are negotiating realistic project goals with MELCO customers and ensuring that these goals are met. This requires an emphasis on project management skills and meeting goals with high reliability even in the face of changing circumstances. The best projects in this environment are ones with a certainty of success.

The figure below symbolizes the way that MERL's activities are divided between MRL and MTL. Both labs participate in the full range of activities and time frames. It is vital for MRL to follow through on its research all the way to development and final delivery to MELCO. It is equally important for MRL's activities are longer-range research and the bulk of MTL's activities are shorter-range advanced development and development.

Together, the two labs are the yin and yang of MERL. They are different, but neither is more important than the other. They work very closely together with many projects spanning the two labs. Neither one would be very effective alone. Together, they allow MERL to be a vibrant research organization with a high level of impact on MELCO.



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 more than 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 estimated 2005 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. The largest of these (MDEA) is part of Lihon and has sales of approximately $800 million.

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 65 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. H.Ogata (Director), Mr. K.Kuroda (GM), 18 people (Tokyo) Managing MELCO's R&D	Hatsuhon
Advanced Technology R&D Center (ATC) Research & Advanced Development: Dr. K.Kyuma (GM), 926 people (Itami) Materials, Semiconductor Devices, Electrical & Mechanical Engineering	Sentansoken
Information Technology R&D Center (ITC) Advanced Development: Dr. H.Koezuka (GM), 847 People (Ofuna) Information Systems, Communications, Opto-Electronics	Johosoken
Industrial Design Center (IDC) Advanced Development: Mr. I.Arai (GM), 93 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), 57 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, 2004 through June 30, 2005.

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

Best Paper Awards

"Discovering Meaningful Multimedia Patterns with Audio-Visual Concepts and Associated Text", Lexing Xie, L. Kennedy, S.-F. Chang, A. Divakaran, H. Sun and C.-Y. Lin. IEEE International Conference on Image Processing, 2004. (One of three best student paper awards at ICIP 2004.)

Technology Awards From Outside Organizations

In September 2004, Ramesh Raskar was chosen by MIT Technology Review for inclusion in their list of the top 100 innovators under the age of 35 worldwide.

In September 2004, Wojciech Matusik was chosen by MIT Technology Review for inclusion in their list of the top 100 innovators under the age of 35 worldwide.

In the spring of 2005, DiamondHelp was selected as a finalist for the international design competition associated with the Third International Conference on Appliance Design (3AD) and a finalist for the "INDEX: Award 2005" international design competition.

In June 2005, MERL's Digital Merchandising technology as presented by Viaro (a Herman Miller company) won the "Best New Product 2005" award at the annual conference of the Society for Environmental Graphic Design.

Awards From MELCO

In Early 2005, MERL staff received a MELCO Valuable Invention Award for their patent "Route Discovery in Ad-Hoc Networks with Data Packets".

In June 2005, MERL staff received a MELCO Corporate R&D award for their work on "Visual-Related Technologies for Digital Home Appliance Business".

In June 2005, MERL staff received a MELCO Corporate R&D award for their work on "Visual Systems for Large Display Facilities".

In June 2005, MERL staff received an award from MELCO's Johosoken Research Laboratory for their work on the "Development of MPEG Decoder for Digital TV".

In June 2005, MERL staff received an award from MELCO's Sentansoken Research Laboratory for their work on "Personal Authentication Technology".

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.

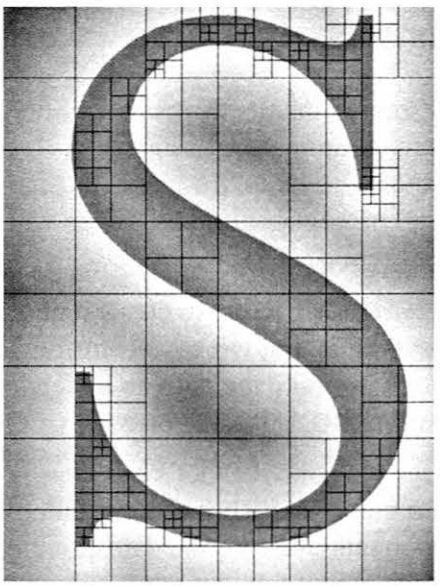
The following subsections detail what MERL's impact on MELCO has been in the 12 months spanned by this report. In addition, they summarize how this impact was achieved. It is worthy of note that there are several distinct models of how impact can be achieved ranging from work specifically requested by 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.

Saffron Digital Typography

In July 2004, Macromedia signed a license to include MERL's Saffron digital type rendering technology for the Macromedia Flash family of products. The Macromedia Flash Player is used by more than 300 million people world wide to view dynamic web content. Saffron displays scalable type esthetically and efficiently. Its use will greatly improve the look of type in the next version of the Flash Player, which will be released in the summer of 2005.

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 (see page 128).

Details: Saffron is the outgrowth of work that goes back to the late 1990s. At that time, 3D computer graphics was a prime area of research at MERL. One of many projects was a long-range fundamental research project to explore new ways of representing 3D graphics using implicit functions. This resulted in the creation of Adaptively sampled Distance Fields (ADFs).



The most common way of representating 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 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 computer 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.

In order 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, rendering the character 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 loosing 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.

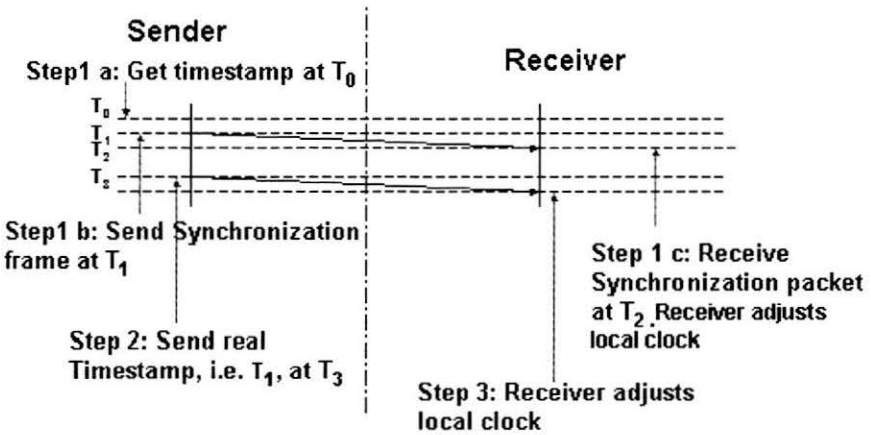
All in all, Saffron is better than traditional outline font rendering in every way but one—the memory required to represent a character as a distance map is considerably greatly than the memory required to represent a hinted outline. However, by using on-the-fly conversion from outlines to ADFs combined with caching, this memory problem can be dispensed within 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 wished to greatly improve the look of type in their Flash Player. 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 and Macromedia proceeded to integrate Saffron into Flash.

Time Synchronization in IEEE 802.15.4b

In September 2004, MERL's contribution "Time Synchronization Proposal" was accepted by the IEEE 802.15.4b Wireless Personal Area Network (WPAN) standard task group. This proposal is a fundamental contribution to the standard specifying how to synchronize different devices in an ad-hoc network. MERL's attribute macSyncSymbolOffset for detailed time adjustment between a sender and a receiver appears in draft standard IEEE P802.15.4-REVb/D1 of January 2005.



MTL is actively contributing to a wide range of wireless communications standards including High Data Rate Ultra wide Band (IEEE 802.15.3a) (see pages 82 & 85), Low Data Rate Ultra wide Band (IEEE 802.15.4a) (p. 83), High data rate Local Area Networks (IEEE 802.11n) (pp. 87 & 92), ZigBee (pp. 76-81), and future cell phone standards (3GPP) (pp. 93 & 94). MTL is also active in various efforts to make products relating to these standards.

Details: IEEE 802.15.4 is a low-level standard for low cost, low data rate, medium range WPANs. Potential WPAN applications include home automation and communication with sensors. IEEE 802.15.4 has been selected as the basis for the ZigBee standard for ad-hoc networks such as sensor networks. ZigBee is a high-level standard focusing on the network layer and above. Ad-hoc networks are ones where communicating devices have to self-organize into a network rather than being connected in a pre-planned fashion.

MELCO became a founding member of the ZigBee alliance in 2002. Since then, MTL has been actively involved in ZigBee standard development and ad-hoc networking research. Under the impetus of ZigBee standard developments, the IEEE 802.15.4b Task Group was formed to define enhancements and corrections to the original 802.15.4 standard. In early 2004, an MRL researcher working with MTL researchers identified time synchronization as an area where IEEE802.15.4 needed to be improved. They presented a proposal for improvement in July, 2004. After several months of intensive work and revision, the proposal was accepted in September 2004.

Time synchronization is required for fine-tuned coordination of wake/sleep duty cycles to reduce power consumption for low-power devices. It is also required to preserve event order for event-detection related applications. However, in the original 802.15.4 standard, there is no special treatment regarding this issue.

MERL's time synchronization proposal includes a basic scheme and a high accuracy scheme for achieving synchronization between sender and receiver. In the basic scheme, the sender sends the time read from its internal clock in a special synchronization frame to the receiving device. When a receiver gets a synchronization frame, it resets its local clock using the time in the frame.

In the high accuracy scheme, two fine tuning steps are added. One step compensates for the delay at the sender between the time T0 it fetches from its internal clock and the time T1 when the synchronization frame is eventually sent. This compensation is achieved by having the sender send a second synchronization frame containing T1. When the receiver gets T1, it increments its clock by T1-T0.

The second step compensates for the delay in the receiver between the reception of a synchronization frame and the time when the local clock is eventually set. This delay is vendor implementation dependent. It is specified by the attribute macSyncSymbolOffset, which is stored in the PAN Information Base and to correct the internal clock.

DiamondTouch in a Command and Control System

In December 2004, MERL's DiamondTouch hardware and software was included as part of Distributed Command and Control Experimental System Phase 2 (DCS2), which was developed by MELCO's Electronic Systems business unit for the Japanese Defense Agency. DiamondTouch is a touch surface that can distinguish multiple simultaneous touches by multiple users. Although only a small part of DCS2, DiamondTouch is a novel and unique technology that was a key differentiator between MELCO's DCS2 proposal and those of its competitors.

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 (see pages 114-116) that use DiamondTouch such as DiamondSpace (p. 112), 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. They have no way of identifying who is touching them, they merely record touches. Typically, they 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.

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. 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 connected to them.

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 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 the Electronic Systems business unit, began showing DiamondTouch to customers.

In 2003 the Electronic Systems business unit used DiamondTouch as a key part of its bid to build DCS2 and won the contract. In collaboration with Johosoken, MERL delivered DiamondTouch for use in DCS2. The Electronic Systems business unit is currently actively seeking additional customers for DiamondTouch.

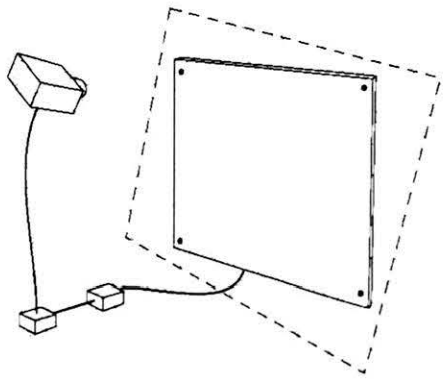
At the same time that opportunities were being pursued for DiamondTouch to be a subcomponent used within MELCO's traditional lines of business, MERL initiated an effort to determine whether DiamondTouch by itself could become a MELCO product. To generate interest, MERL has given prototype DiamondTouch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

As a final step in this commercialization effort, MERL will begin selling small quantities of DiamondTouch prototypes in the fall of 2005. We hope that this will lead to volume sales and that DiamondTouch will become a MELCO product in its own right.

Automatic Image Alignment for Electronic Whiteboards

In January 2005, MERL licensed automatic image alignment technology to a major manufacturer of electronic whiteboards. Electronic white boards support human computer interaction in two

ways. First, a person can write/draw on the white board and the writing/drawing is captured in computer form. Second, an image (e.g., a Power Point slide) can be projected onto the white board and the writing/drawing that is captured can be saved as annotation on the image. For this to work well, the exact position and shape of the projected image on the whiteboard must be determined and recorded so that the writing/drawing can be correctly positioned in the image. In the past, this alignment has been done by having the user touch a series of spots projected onto the white board. Using MERL's technology, alignment is done automatically.



Currently, the licensee is using MERL's technology merely to locate the position and shape of the projected image. MERL's technology is also capable of warping an image to achieve a desired position and shape. It is our intention to use the technology to the fullest in conjunction with our own work on touch surfaces (see the description of DiamondTouch in a Command and Control System above). MERL's work on aligning multiple projectors with each other (see the description of a Multi-Projector Display System for Curved Surfaces below) is closely related, but uses a different approach to image registration that requires a camera.

Details: Like many innovations, MERL's automatic image alignment technique arose from the interaction of people from quite different disciplines. In particular, a range of work at MRL and MTL on projector technology combined with work at both labs on Human/Computer Interaction (HCI) lead to a stream of projects going back to 2000 involving novel devices for interacting with projected images.

For example, in the DiamondTouch project, multiple users can simultaneously interact with an image projected on a touch table. In the Digital Merchandising project, interaction by touch with images projected on walls and products is used to great effect in retail settings. In both of these projects, a manual calibration step was used to align the projectors with the surfaces the images are projected on.

In early 2003, a MERL researcher began looking for automatic solutions to this alignment problem. Prior to the work described here, automatic alignment was, in general, done with cameras and computer vision algorithms that determine the positions of projected images and the objects they need to be aligned with. The MERL researcher sought to find a solution that did not use a camera. He did this both to save the cost of the camera and to achieve greater accuracy than can be achieved with typical cameras.

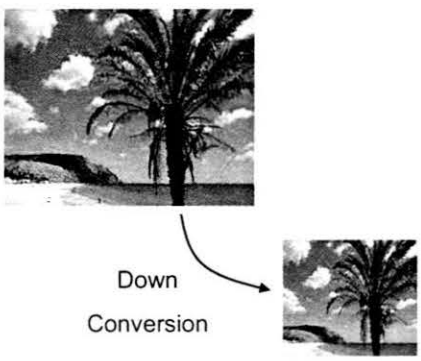
His fundamental insight was to turn the problem around---rather than using a camera to see what physical position a given part of an image is projected on, he uses a tiny light sensor (e.g., an optical fiber connecting to a photodiode) located at the physical position so that the physical position can "see" what is projected on it. Rather than capturing the whole image with low positional accuracy, only a few tiny spots of the image are captured, but very cheaply and with very high positional accuracy. A time-varying test pattern is projected that has a unique sequence of flashes at each pixel. This allows the system to tell which projected pixel is shining on each sensor.

An interesting subtlety in the system is that even assuming a sensor is smaller than a single projected pixel it is likely that it will see parts of at least two pixels P and Q. This causes problems.

For instance, suppose that the flash sequence corresponding to P is all ones and the sequence for Q is all zeros. In that case, an intermediate light level will be seen at each time step and every possible flash sequence will seem equally likely, yielding no positional information. What is needed is a pattern of flash sequences such that the codes for adjacent pixels never differ by more than one flash. Such a coding system was devised many decades ago by Gray. In addition to solving the problem above, using two-dimensional Gray codes makes the MERL system robust to pixel size variations (relative to the sensor) and focusing errors.

MERL's automatic alignment approach was first demonstrated at the UIST conference in 2004. This immediately generated significant interest from a number of electronic whiteboard manufacturers and licensing negotiation began in August 2004.

Down-Decoding in DTV Receiver Chip



In the first quarter of 2005, MELCO's associated semiconductor company Renesas began selling Digital TV (DTV) receiver chips (the R8A66950BG family) for the US market to a major television manufacturer. The chip is designed for low cost TVs that can receive High Definition TV (HDTV) signals but display their output using Standard Definition TV (SDTV) resolution display screens. An important part of the chip is an efficient "down-decoding" algorithm that allows the HDTV video input to be decoded and converted to SDTV video in a single step. In contrast to first decoding and then down-converting, this algorithm reduces the cost of the overall system by reducing the memory required for image storage, the memory bandwidth required for data transfer and the amount of computation that has to be performed. Key components of this algorithm were developed by MERL.

Digital video processing is one of the central areas of expertise at MERL. In addition to down-decoding, video processing work at MERL includes video compression (see pages 101-104) and post filtering to remove compression artifacts (see the description below of post-filtering for a professional HDTV codec).

Details: One of the predecessor labs that was merged to form the current MERL was founded in 1993 as part of a large project to create the first chip set that could decode signals conforming to the newly created US HDTV standard. This chip set was a landmark effort in the early days of the analog-to-digital transition in TV sets and was used in HDTVs produced by MELCO.

Under strong demands from the US government, the US TV industry was (and is) moving rapidly toward converting all TVs to handle HDTV input. However, only a relatively small percentage of people are ready to buy high-priced TV sets that can display images at full HDTV resolution. Therefore, there is a strong requirement for converting HDTV signals so that they can be viewed at SDTV resolution on low-priced TV sets.

The obvious approach to converting an HDTV input into an SDTV output is to fully decode the HDTV input creating HD images and then apply a filtering and sub-sampling process to reduce the resolution of each HD image to an SD image. However, this requires all the memory and processing that is needed for HDTV decoding, plus additional resources for the subsequent down-sampling. This is not appropriate for TV sets that are intended to be low-priced. In collaboration with MELCO video researchers in Japan, MTL researchers began to work on lower cost solutions in 1995.

The first step to a solution was the insight that theoretically, down-conversion can be applied early in the decoding process rather than after it is complete. When HDTV signals are encoded, each image is divided into blocks, which are individually compressed. Down-sampling can be directly applied to the compressed form of these blocks without decompressing them. This saves memory and computation in the overall process because there is subsequently less data to process when decoding. MERL developed a unique approach to early down-conversion that does a better job of preserving high-frequency information than other approaches. This yields sharper edges in the eventual down-converted images.

The insight above is powerful; however, it is not by itself a full solution. A key reason for this is the use of "motion compensation" in video compression and the tendency for "drift" to occur. In efficient video coding, most of the potential for compression stems from the fact that adjacent video frames are very similar. To take advantage of this, only a few frames are encoded individually using single image compression techniques. Most frames are represented in terms of how they differ from adjacent frames. A particularly important difference is that some parts of an image may move slightly from one frame to the next. Motion compensation represents these kinds of changes. Drift refers to the fact that errors can accumulate as a frame is computed from a prior frame and then is used as the basis for computing the next frame and so on.

In an HDTV signal, motion compensation information is represented with reference to the HD resolution blocks in the images. As a result, it cannot be directly applied to down-sampled versions of these blocks. MERL experimented with various ways to deal with this and other problems until a satisfactory solution was finally arrived at in 1997. In the crucial motion compensation step, the MERL algorithm up-samples the down-sampled blocks to determine the appropriate reference information for motion compensation. This use of up-sampling applied to down-sampled data allows the bandwidth to the decoder to be greatly reduced, but introduces reconstruction errors. To prevent this from causing drift, the up-sampling and down-sampling filters used in the MERL algorithm had to be carefully designed to minimize loss.

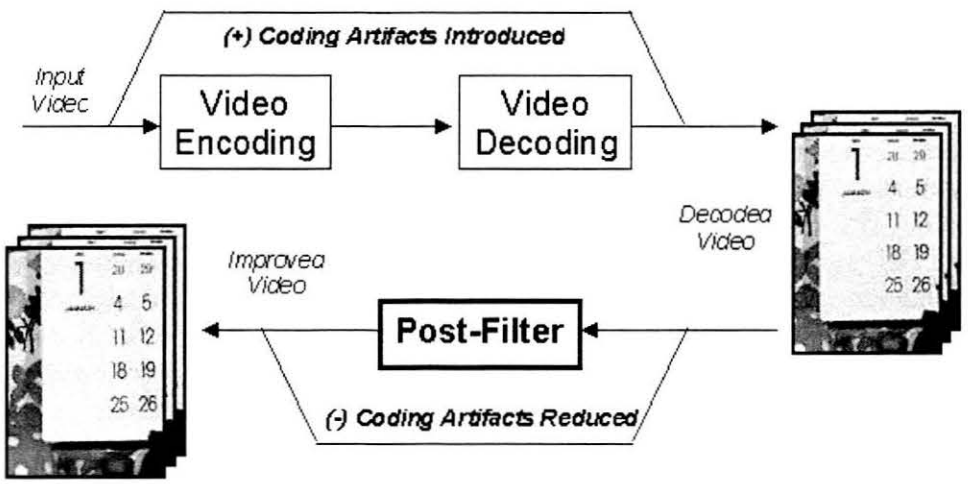
After the completion of the first HDTV chip set, MELCO's semiconductor business unit began a project to create a second-generation HDTV chip set. In 1998 MERL's down-decoding algorithm was selected to become part of this new chip set. In 1998 and 1999, MTL completed most of the work needed to put the down-decoding algorithm in the chip set. However, in 1999 production of the chip set was canceled and MERL shelved its work on down-decoding.

In 2003, most of the semiconductor operations of MELCO and Hitachi were merged to create a new company called Renesas. The new company, which included many people familiar with MERL's work on down-decoding, decided to make a low-cost DTV receiver chip for SDTV resolution TVs for the US market. MELCO's Johosoken laboratory was assigned responsibility for the video decoder design, with MTL providing support for the down-decoding components.

The hardware design was completed in 2004 and MTL was asked to support field testing of the chip in US cities. After two successful trials, sales of the chip began in early 2005.

Post-Filtering for Professional HDTV Codec

In the first quarter of 2005, MELCO's Koriyama Works began selling a professional HDTV coder/decoder (or codec) that includes post-filtering technology developed at MERL. In a TV broadcast studio, video is often compressed and decompressed several times before the video is finally broadcast to consumers. Post-filtering techniques are used after decompression to improve video quality by suppressing visual artifacts that can arise due to multiple compression/decompression cycles.



The professional HDTV codec is just one of many potential applications of MERL's post-filtering technology. For instance, we expect our post-filtering to be included in a DVD recorder chip that utilizes high compression coding of video for efficient storage.

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 the 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 quantization in 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.

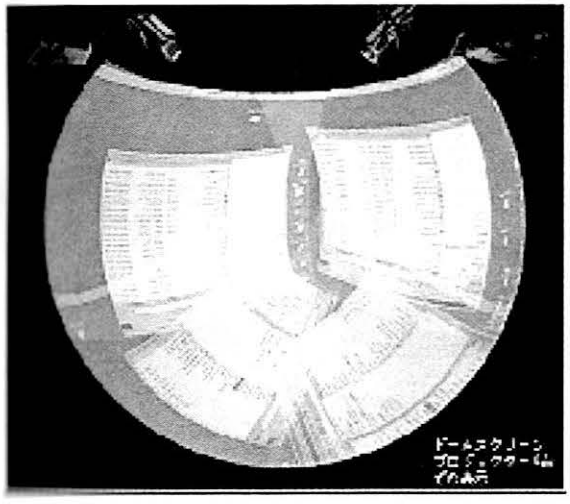
MERL's post-filtering algorithms are unique in attacking both blocking and ringing at the same time and doing so in computationally inexpensive ways. MERL developed a low-complexity block artifact detection method that does a good job of controlling adaptive low-pass filtering. In addition, MERL developed 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.

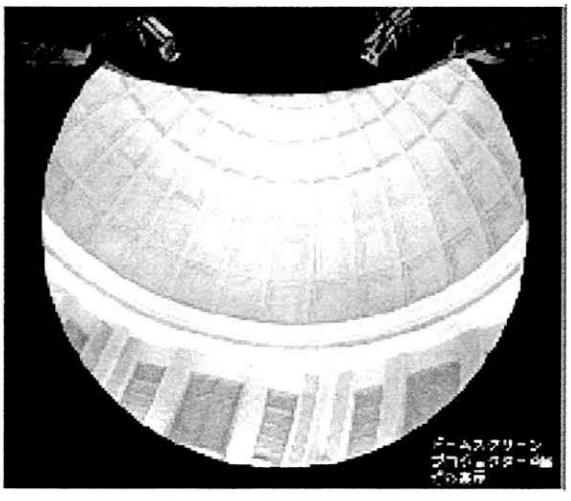
MERL developed its post-filtering algorithms during 2003 and early 2004. Extensive quality tests were conducted by MELCO and algorithmic refinements were made in close collaboration with Johosoken. The algorithm design was fixed in September of 2004. It appeared in Koriyama's professional HDTV codec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

Multi-Projector Display System for Curved Surfaces

In March 2005, MELCO's subsidiary Mitsubishi Precision (MPC) had their first sale of a product they call "Projection Cluster". This is a display system consisting of a two-meter hemispherical screen and four projectors whose images cover the screen. MERL provided the algorithms and software needed to combine the four projected images into a single seamless image.

In collaboration with MERL, a group at MELCO's Johosoken laboratory has extended this to a display system using a three-meter hemispherical screen and six projectors. Discussions are underway about making this a MELCO product as well. In addition, MERL is working on extending the calibration algorithms to support 360-degree wrap around screens. Beyond this, the basic concepts involved in this work have a range of applications including the automatic image registration work described above.





Details: MERL has been a key contributor in the field of projector-related research for a number of years. As part of his PhD thesis, an MRL researcher did basic work on combining multiple projected images into a single seamless image on a flat screen before coming to MERL in 2000. At MERL, he extended this work demonstrating fast, accurate, totally automatic and low-cost alignment for multiple projectors on planar surfaces based on feedback from a camera viewing projected test patterns.

Prior to this work, there were systems that used cameras for automatic alignment; however, they operated based on the highly precise determination of the positions specific projected points, which required highly precise (and expensive) cameras. The central feature of the MERL work stems from the fact that the distortion of an image that results from projecting it on an arbitrary plane surface can be completely described by a single matrix that in turn can be specified by a small number of independent parameters. These parameters can be determined with high accuracy without knowing the position of any specific projected point with high accuracy, if one can determine the approximate position of many projected points. This information can be captured using an imprecise, inexpensive camera.

Curved screen displays using multiple projectors have been in use for some time in application areas such as visualization, simulation and planetariums. However, these systems have typically required very cumbersome manual alignment. In the summer of 2002, the MRL researcher and a collaborator at MTL extended the fast calibration algorithms to work on curved screens. This was possible, because the distortion caused by projecting an image on a quadratic curved surface such as a cylinders or hemisphere can also be represented using only a small set of simple parameters. The figure shows a four-projector dome-projection system before (left) and after (right) automatic calibration to create a single seamless image.

MERL's multi-projector work was demonstrated to MELCO in Japan on several occasions. In mid-2003, MPC became interested in the technology and a business plan based on this technology was put in place. In early 2004, an MPC researcher visited MERL for three months to help mature the technology to product level. In mid-2004 the collaboration resulted in a product that was accepted by MPC and in late 2004 it was offered for sale.

Technical Staff

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

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

Avidan's research focus is computer vision with occasional detours into computer graphics and machine learning. He joined MERL in 2004 after three years at MobilEye, where he developed detection and tracking algorithms for vision-based driver assistance systems. 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.

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.

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

Cukier joined MERL in 1996. His initial focus was on digital systems for CATV, RF microwave transmitters & receivers, and front-ends for advanced TV receivers. His current interests are in advanced Digital Networking and Digital Signal Processing.

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

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

Paul H. Dietz *Ph.D., Carnegie Mellon University, 1995*
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

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).

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

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

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

Forlines' research interests include the design and evaluation of novel user interfaces. His current research projects span from three-dimensional presentation of and navigation through recorded digital video, to collaborative tabletop user interfaces, to using hand-held projectors for augmented reality. He is currently leading the user evaluation of three projects, MediaFinder, TimeTunnel, and DiamondSpin.

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

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

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

After joining MERL in 1999, Gu has been involved in many wireless communications and networking projects, and has many publications in the field. His current research interests include IEEE802.11 standardizations, QoS in wireless communications and networks, multimedia home networking and MIMO-OFDM technologies.

Jianlin Guo *Ph.D., Windsor University, 1995*
Principal Technical Staff MTL

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

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

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

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

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

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.

Mamoru Kato *B.S., Tokyo University, 1991*
Principal Technical Staff MTL

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

Keisuke Kojima *Ph.D., University of Tokyo, 1990*
Visiting 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.

Emin Martinian *Ph.D., Massachusetts Institute of Technology, 2004*
Member Technical Staff MTL

Martinian's doctoral thesis was in the area of dynamic information and constraints in source and channel coding at MERL. He has been working on problems of video processing, distribution, and compression. His broader research interests include digital communications, signal processing, information theory, belief propagation, and cryptography.

Wojciech Matusik *Ph.D., Massachusetts Institute of Technology, 2003*
Consulting Scientist, MRL

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

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

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

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

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

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

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

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

Molisch's current research interests are multiple-antenna systems, wireless channel measurement and modeling, ultra wideband systems, and OFDM. He is active in standardization (IEEE 802.15, 3GPP, COST273), and has authored or co-authored two books, five book chapters, some 50-journal papers, and numerous conference papers.

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

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

Barton Nicholls
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*
Senior Research Scientist MRL

Prior to joining MERL in 1998, Perry was a consulting engineer at DEC developing a three-dimensional rendering ASIC called Neon. Ron has consulted for many companies including Kodak, Atex, Adobe, Quark, and Apple over the last 20 years, developing software and hardware products in the areas of computer graphics, imaging, color, and desktop publishing. Ron's research interests are centered on key algorithms in computer graphics.

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

Pfister was the chief architect of VolumePro, Mitsubishi Electric's real-time volume rendering hardware for PCs. His research interests include computer graphics, scientific visualization, and computer architecture. His work spans a range of topics, including point-based graphics, 3D photography, volume graphics, and computer graphics hardware. Pfister was the general chair of the IEEE Visualization 2002 conference in Boston.

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).

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

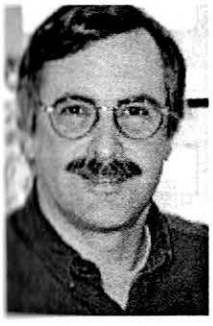
Pozerski's interests have followed the application of computers to a variety of manufacturing tasks including using PDP-11's to demonstrate control of multiple reactor chemical processes, using personal computers for production testing & manufacturing of chemicals and controlling multi-axis rotary assembly machines. Currently, Stan supports Windows and Linux clients, servers, networking, and the wide variety of PC applications used at MERL.

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

Raj works mainly on algorithmic aspects of speech recognition, with special emphasis on improving the robustness of speech recognition systems to environmental noise. His latest work is on the use of statistical information about speech for the automatic design of filter-and-sum microphone arrays. Prior to joining MERL, Raj worked at Compaq's Cambridge MA 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.

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

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

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

Ryall's research interests focus on human-computer interaction, user interfaces and improving human-computer collaboration. Her current research is on the design of interfaces and interaction techniques to support multi-user collaboration on shared surfaces. For 3 years before joining MERL, Kathy was an Assistant Professor of Computer Science at the University of Virginia.

Zafer Sahinoglu *Ph.D., New Jersey Institute of Technology, 2001*
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

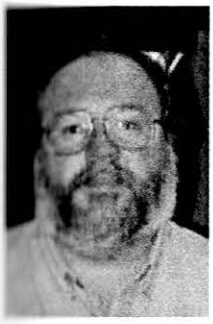
Schmidt-Nielsen spent 7 years at Dragon Systems applying speech recognition to useful products. At MERL he is paying a lot of attention to making speech interfaces robust and usable. He has very broad interests in science and technology. Among many other activities he has taught genetics at the University 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.

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

Shen's current research focuses on shared interactive surfaces. Previously, she led the MidART project, which MELCO has incorporated into several large 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.

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*
Member Technical Staff MTL

Van Baar's interests are in the fields of Computer Graphics, Scientific Visualization, Computer Vision and HCI. He first came to MERL as an intern in 1997. He joined MERL full-time in 1999. The projects he has been working on include points as rendering primitives, automatic keystone correction for projectors, and multi-projector displays on both planar and curved surfaces.

Anthony Vetro *Ph.D., Polytechnic University, 2001*
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*
Visiting Scientist MRL

Before joining MERL, he 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.

Joseph K. Woelfel *M.S., Rutgers University, 1992*
Principal Technical Staff MTL

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Zhang manages MTL's digital communication and networking group. Before joining MERL in 2001, Zhang worked for Nortel Networks for 10 years where she held engineering and management positions in the areas of VLSI design, Advanced Wireless Technology Development, Wireless Networks and Optical Networks. She has a broad technical background, specializing in system design and real-time embedded S/W for wireless communications.

Recent Major Publications

The following lists the 126 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 33 (26%) publications that were subject to highly stringent selection criteria where they were published. Some venues (such as major journals and certain key conferences) are very selective in what they publish and some (such as workshops and many conferences) are not. There are good reasons to publish something in a non-selective venue, the most important of which being that a given workshop or conference may be the best place at which to expose a particular piece of work to the scientific community. However, the mere appearance of a piece of work in a non-selective venue says little if anything about the quality of the work in the eyes of the scientific community.

As a basis for assessing the selectivity of various venues, the list below uses acceptance rates. For instance, certain key conferences such as CHI, ICCV and SIGGRAPH accept only 20% or less of the papers submitted to them, rejecting many papers that in fact describe fine work. In contrast, many workshops and regional conferences accept 80% or more of the papers submitted. 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.

* Pereira, F.; Smith, J.R.; Vetro, A., "Introduction to the Special Section on MPEG-21", *IEEE Transactions on Multimedia*, ISSN: 1520-9210, Vol. 7, Issue 3, pp. 397-399, June 2005

* Vetro, A.; Timmerer, C., "Digital Item Adaptation: Overview of Standardization and Research Activities", *IEEE Transactions on Multimedia*, ISSN: 1520-9210, Vol. 7, Issue 3, pp. 418-426, June 2005 (TR2005-056)

Lee, J.H.; Machiraju, R.; Pfister, H.; Moghaddam, B., "Estimation of 3D Faces and Illumination from Single Photographs Using a Bilinear Illumination Model", *Eurographics Symposium on Rendering (EGSR)*, June 2005 (TR2005-045)

Tuzel, O.; Porikli, F.M.; Meer, P., "A Bayesian Approach to Background Modeling", *IEEE Workshop on Machine Vision for Intelligent Vehicles (MVIV)*, June 2005 (TR2005-033)

* Brand, M.E., "A Direct Method for 3D Factorization of Nonrigid Motion Observed in 2D", *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2005 (TR2005-052)

* Avidan, S., "Ensemble Tracking", *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2005 (TR2005-065)

Tumblin, J.; Agrawal, A.; Raskar, R., "Why I Want a Gradient Camera", *IEEE International Conference on Computer vision and Pattern Recognition (CVPR)*, June 2005 (TR2005-083)

Wren, C.R.; Erdem, M.; Azarbayejani, A.J., "Automatic Pan-Tilt-Zoom Calibration in the Presence of Hybrid Sensor Networks", *IEEE International Conference on Computer vision and Pattern Recognition (CVPR)*, June 2005 (TR2005-084)

* Porikli, F.M., "Integral Histogram: A Fast Way to Extract Higtograms in Cartesian Spaces", *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2005 (TR2005-057)

Yuan, Y.; Gu, D.; Arbaugh, W.; Zhang, J., "Achieving Fair Scheduling Over Error-Prone Channels in Multirate WLANs", *International Conference on Wireless Networks, Communicaitons and Mobile Computing (WIRELESSCOM)*, June 2005 (TR2005-082)

Rich, C.; Sidner, C.L.; Lesh, N.; Garland, A.; Booth, S.; Chimani, M., "DiamondHelp: A Collaborative Interface Framework for Networked Home Appliances", *IEEE International Conference on Distributed Computing Systems Workshops (ICDCS)*, pp. 514-519, June 2005 (TR2004-114)

Almers, P.; Wyne, S.; Tufvesson, F.; Molisch, A.F., "Effect of Random Walk Phase Noise on MIMO Measurement", *IEEE Vehicular Technology Conference (VTC)*, May 2005 (TR2005-059)

Lau, B.K.; Ow, S.M.S.; Kristensson, G.; Molisch, A.F., "Capacity Analysis for Compact MIMO Systems", *IEEE Vehicular Technology Conference (VTC)*, May 2005 (TR2005-060)

Wyne, S.; Molisch, A.F.; Almers, P.; Eriksson, G.; Karedal, J.; Tufvesson, F., "Statistical Evaluation of Outdoor-to-Indoor Office MIMO Measurements at 5.2 GHz", *IEEE Vehicular Technology Conference (VTC)*, May 2005 (TR2005-062)

Johansson, A.; Karedal, J.; Tufvesson, F.; Molisch, A.F., "MIMO Channel Measurement for Personal Area Networks", *IEEE Vehicular Technology Conference (VTC)*, May 2005 (TR2005-061)

Pekhteryev, G.; Sahinoglu, Z.; Orlik, P.V.; Bhatti, G.M., "Image Transmission over IEEE 802.15.4 and ZigBee Networks", *IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2005 (TR2005-030)

Wu, M.; Vetro, A.; Yedidia, J.S.; Sun, H.; Chen, C.W., "A Study of Encoding and Decoding Techniques for Syndrome-Based Video Coding", *IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2005 (TR2005-038)

Su, Y.; Xin, J.; Vetro, A.; Sun, H., "Efficient MPEG-2 to H.264/AVC Intra Transcoding in Transform-domain", *IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2005 (TR2005-039)

Kong, H-S; Vetro, A.; Hata, T.; Kuwahara, N., "Fast Region-of-Interest Transcoding for JPEG 2000 Images", *IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2005 (TR2005-043)

Raskar, R.; van Baar, J., "Low-Cost Multi-Projector Curved Screen Displays", *International Symposium Scoiety for Information Display (SID)*, May 2005 (TR2005-079)

Porikli, F.M.; Shao, J.; Maehara, H., "Heli-Tele: Road Extraction from Helicopter Video", *IAPR Conference on Machine Vision Applicaitons (MVA)*, May 2005 (TR2005-041)

Du, J.; Gu, D.; Molisch, A.F.; Li, Y.; Zhang, J., "Layered Space-Time Structure with Statistical Rate Allocation", *IEEE International Conference on Communications (ICC)*, May 2005 (TR2005-081)

* Brand, M.E., "A Random Walks Perspective on Maximizing Satisfaction and Profit", *SIAM Conference on Optimization*, May 2005 (TR2005-050)

Hancock, M.S.; Shen, C.; Forlines, C.L.; Ryall, K., "Exploring Non-Speech Auditory Feedback at an Interactive Multi-User Tabletop", *Graphics Interface 2005*, May 2005 (TR2005-054)

Porikli, F.M., "Ambiguity Detection by Fusion and Conformity: A Spectral Clustering Approach", *IEEE International Conference Integration of Knowledge Intensive Multi-Agent Systems (KIMAS)*, pp. 366-372, April 2005 (TR2005-035)

Wren, C.R., "Large Networks of Ultra-low Resolution Sensors in Buildings", *International Conference on Integration of Knowledge Intensive Multi-Agent Systems (KIMAS)*, pp. 373-389, April 2005 (TR2005-003)

Porikli, F.M.; Wren, C.R., "Change Detection by Frequency Decomposition: Wave-Back", *International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*, April 2005 (TR2005-034)

Shen, C.; Hancock, M.S.; Forlines, C.L.; Vernier, F.D., "CoR2Ds: Context-Rooted Rotatable Draggables for Tabletop Interaction", *ACM Conference on Human Factors in Computing Systems (CHI) Short Paper*, April 2005 (TR2005-001)

Ryall, K.; Lesh, N.; Lanning, T.; Leigh, D.L.; Miyashita, H.; Makino, S., "QueryLines: Approximate Query for Visual Browsing", *ACM Conference on Human Factors in Computing Systems (CHI) Short Paper*, April 2005 (TR2005-015)

Everitt, K.; Shen, C.; Ryall, K.; Forlines, C.L., "Modal Spaces: Spatial Multiplexing to Mediate Direct-Touch Input on Large Displays", *ACM Conference on Human Factors in Computing Systems (CHI) Short Paper*, April 2005 (TR2005-002)

Forlines, C.L.; Shen, C.; Buxton, B., "Glimpse: A Novel Input Model for Multi-Level Devices", *ACM Conference on Human Factors in Computing Systems (CHI) Short Paper*,

April 2005 (TR2005-024)

Sidner, C.L.; Lee, C.H., "Robots as Laboratory Hosts", *Interactions*, ISSN: 1072-5520, Vol. 12, Issue 2, pp. 24-16, March 2005

* Nie, Y.; Kong, H-S, Vetro, A.; Barner, K., "Fast Adaptive Fuzzy Post-Filtering for Coding Artifacts Removal in Interlaced Video", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-6149, Vol. 2, pp. 993-996, March 2005 (TR2005-018)

* Guinness, J.; Raj, B.; Schmidt-Nielsen, B.K.; Turicchia, L.; Sarpeshkar, R., "A Companding Front End for Noise-Robust Automatic Speech Recognition", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-6149, Vol. 1, pp. 249-252, March 2005 (TR2005-023)

* Wang, B.; Wang, Y.; Selesnick, I., Vetro, A., "Video Coding Using 3-D Dual-Tree Discrete Wavelet Transforms", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-6149, Vol. 2, pp. 61-64, March 2005 (TR2005-019)

* Xie, L.; Kennedy, L.; Chang, S-F; Divakaaran, A.; Sun, H.; Lin, C-Y, "Layered Dynamic Mixture Model for Pattern Discovery in Asynchronous Multi-Modal Streams", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-6149, Vol. 2, pp. 1053-1056, March 2005 (TR2005-078)

Zhang, J.; Wang, Y.; Fossorier, M.; Yedidia, J.S., "Replica Shuffled Belief Propagation Decoding of LDPC Codes", *Johns Hopkins University Conference on Information Scienses and Systems (CISS)*, March 2005 (TR2005-005)

Porikli, F.M.; Tuzel, O., "Object Tracking in Low-Frame-Rate Video", *SPIE Image and Video Communications and Processing*, Vol. 5685, pp. 72-79, March 2005 (TR2005-013)

Radhakrishnan, R.; Divakaran, A., "Systematic Acquisition of Audio Classes for Elevator Surveillance", *SPIE Image and Video Communications and Processing*, Vol. 5685, pp. 64-71, March 2005 (TR2005-076)

Porikli, F.M., "Computationally Efficient Histogram Extraction for Rectangular Image Regions", *SPIE Real-Time Imaging IX*, Vol. 5671, pp. 36-43, February 2005 (TR2005-014)

Chimani, M.; Lesh, N.; Mitzenmacher, M.; Sidner, C.L.; Tanaka, H., "A Case Study in Large-Scale Interactive Optimization", *IASTED International Conference on Artificial Intelligence and Applications AIA 2005*, ISSN: 1027-2666, Track - Planning and Scheduling, February 2005 (TR2004-113)

* Xin, J.; Lin, C-W; Sun, M-T, "Digital Video Transcoding", *Proceedings of the IEEE*, Vol. 93, Issue 1, pp. 84-97, January 2005 (TR2005-006)

* Chang, S-F; Vetro, A., "Video Adaptation: Concepts, Technologies and Open Issues", *Proceedings of the IEEE*, ISSN: 018-9219, Vol. 93, Issue 1, pp. 148-158, January 2005 (TR2005-010)

* Weyrich, T.; Pfister, H.; Gross, M., "Rendering Deformable Surface Relectance Fields ", *IEEE Transactions on Visualization and Computer Graphics*, ISSN: 1077-2626, Vol. 11, Issue 1, pp. 48-58, January 2005 (TR2005-008)

Yerazunis, W.S.; Chhabra, S.; Siefkes, C.; Assis, F.; Gunopulos, D., "A Unified Model of Spam Filtration ", *MIT Spam Conference,* January 2005 (TR2005-085)

Divakaran, A.; Forlines, C.L.; Lanning, T.; Shipman, S.E.; Wittenburg, K.B., "Augmenting Fast-Forward and Rewind for Personal Digital Video Recorders", *IEEE International Conference on Consumer Electronics (ICCE)*, Digest of Technical Papers, pp. 43-44, January 2005 (TR2004-136)

* Beardsley, P.A.; van Baar, J.; Raskar, R.; Forlines, C.L., "Interaction Using a Handheld Projector", *IEEE Computer Graphics and Applications*, ISSN: 0272-1716, Vol. 25, Issue 1, pp. 39-43, January 2005 (TR2004-107)

* Tan, K-H; Feris, R.S.; Turk, M.; Kobler, J.; Yu, J.; Raskar, R., "Harnessing Real-World Depth Edges with Multiflash Imaging", *IEEE Computer Graphics and Applications*, ISSN: 0272-1716, Vol. 25, Issue 1, pp. 32-38, January 2005 (TR2005-067)

Radhakrishnan, R.; Otsuka, I.; Xiong, Z.; Divakaran, A., "Modeling Sports Highlights Using a Time-Series Clustering Framework and Model Interpretation", *SPIE Storage and Retrieval Methods and Applications for Multimedia*, Vol. 5682, pp. 269-276, January 2005 (TR2005-077)

Nie, Y.; Hong, H-S; Vetro, A.; Shimada, T.; Minegishi, N., "Implementation of Edge Map Guided Fuzzy Filtering for Artifact Reduction in Highly Compressed Video", *IEEE International Conference on Consumer Electronics (ICCE)*, Digest of Technical Papers, pp. 325-326, January 2005 (TR2005-007)

Otsuka, I.; Nakane, K.; Divakaran, A.; Hatanaka, K.; Ogawa, M., "A Highlight Scene Detection and Video Summarization System Using 'Audio' Feature for a Personal Video Recorder", *IEEE International Conference on Consumer Electronics (ICCE)*, Digest of Technical Papers, pp. 223-224, January 2005 (TR2005-011)

Rich, C.; Sidner, C.L.; Lesh, N.; Garland, A.; Booth, S., "Collaborative Help for Networked Home Products", *IEEE International Conference on Consumer Electronics (ICCE)*, Digest of Technical Paper, pp. 209-210, January 2005 (TR2004-083)

* Otsuka, I.; Nakane, K.; Divakaran, A.; Hatanaka, K.; Ogawa, M., "A Highlight Scene Detection and Video Summarization System Using Audio Feature for a Personal Video Recorder", *IEEE Transactions on Consumer Electronics*, ISSN: 0098-3063, Vol. 51, pp. 112-116, Janaury 2005 (TR2005-011)

Wren, C.R.; Porikli, F.M., "Waviz: Spectral Similarity for Object Detection", *IEEE International Workshop on Performance Evaluation of Tracking & Surveillance (PETS)*, January 2005 (TR2005-004)

* Lesh, N.; Marks, J.W.; Rich, C.; Sidner, C.L., "Man-Computer Symbiosis Revisited: Achieving Natural Communication and Collaboration with Computers", *IEICE Transactions on Electronics (IEICE)*, December 2004 (TR2004-053)

* Shakhnarovich, G.; Moghaddam, B., "Face Recognition in Subspaces", *Handbook of Face Recognition*, December 2004 (TR2004-041)

Xin, J.; Vetro, A.; Sun, H., "Efficient Macroblock Coding-mode Decision for H.264/AVC Video Coding", *Picture Coding Symposium (PCS)*, December 2004 (TR2004-079)

Vetro, A.; Matusik, W.; Pfister, H.; Xin, J., "Coding Approaches for End-to-End 3D TV Systems", *Picture Coding Symposium (PCS)*, December 2004 (TR2004-137)

Wang, B.; Wang, Y.; Selesnick, I.; Vetro, A., "Video Coding Without Motion Compensation Using a 3-D Dual-tree Wavelet Transform", *Picture Coding Symposium (PCS)*, December 2004 (TR2004-138)

* Avidan, S.; Butman, M., "The Power of Feature Clustering: An Application to Object Detection", *Neural Information Processing Systems (NIPS)*, December 2004 (TR2004-141)

Wren, C.R.; Minnen, D.C., "Activity Mining in Sensor Networks", *Advances in Neural Information Processing Systems (NIPS) Workshop on Activity Recognition and Discovery*, December 2004 (TR2004-135)

Divakaran, A.; Otsuka, I.; Radhakrishnan, R.; Nakane, K.; Ogawa, M., "Audio-Assisted Video Browsing for DVD Recorders", *IEEE Pacific-Rim Conference on Multimedia (PCM)*, November 2004 (TR2004-139)

Xin, J.; Vetro, A.; Sun, H., "Converting DCT Coefficients to H.264/AVC Transform Coefficients", *IEEE Pacific-Rim Conference on Multimedia (PCM)*, Lecure Notes in Computer Science, ISSN: 0302-9743, Vol. 3332/2004 pp. 939, November 2004 (TR2004-058)

* Zhou, S.K.; Chellappa, R.; Moghaddam, B., "Visual Tracking and Recognition Using Appearance-Adaptive Models in Particle Filters", *IEEE Transactions on Image Processing*, ISSN: 1057-7149, Vol. 13, Issue 11, pp. 1491-1506, November 2004 (TR2004-028)

Sudarshan, P.; Mehta, N.B.; Molisch, A.F.; Zhang, J., "Spatial Multiplexing and Channel Statistics-Based RF Pre-Processing for Anetenna Selection", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 6, pp. 3947-3951, November 2004 (TR2004-112)

Gezici, S.; Tufvesson, F.; Molisch, A.F., "On the Performance of Transmitted-Reference Impulse Radio", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 5, pp. 2874-2879, November 2004 (TR2004-143)

Karedal, J.; Wyne, S.; Almers, P.; Tufvesson, F.; Molisch, A.F., "UWB Channel Measurements in an Industrial Environment", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 6, pp. 3511-3516, November 2004 (TR2004-142)

Wu, J.; Horng, H.; Zhang, J.; Olivier, J.C.; Xiao, C., "Combining Orthogonal Space Time Block Codes with Adaptive Sub-Group Antenna Encoding", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 1, pp. 346-350, November 2004 (TR2004-144)

Yu, W.; Sahinoglu, Z.; Vetro, A., "Energy Efficient JPEG 2000 Image Transmission Over Wireless Sensor Networks", *IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 5, pp. 2738-2743, November 2004 (TR2003-111)

Sidner, C.L.; Lee, C.H.; Kidd, C.; Lesh, N., "Explorations in Engagement for Humans and Robots", *IEEE RAS/RSJ International Conference on Humanoid Robots (HUMANOIDS)*, November 2004 (TR2004-048)

* Ryall, K.; Forlines, C.L.; Shen, C.; Ringel-Morris, M., "Exploring the Effects of Group Size and Table Size on Interactions with Tabletop Shared-Display Groupware", *ACM Conference on Computer Supported Cooperative Work (CSCW)*, ISBN: 1-58113-810-5, pp. 284-293, November 2004 (TR2004-124)

* Ringel-Morris, M.; Ryall, K.; Shen, C.; Forlines, C.L.; Vernier, F., "Beyond Social Protocols: Multi-User Coordination Policies for Co-located Groupware", *ACM Conference on Computer Supported Cooperative Work (CSCW)*, ISBN: 1-58113-810-5, pp. 262-265, November 2004 (TR2004-123)

Catovic, A.; Sahinoglu, Z., "The Cramer-Rao Bounds of Hybrid TOA/RSS and TDOA/RSS Location Estimation Schemes", *IEEE Communication Letters*, ISSN: 1089-7798, Vol. 8, Issue 10, pp. 626-628, October 2004 (TR2003-143)

Marks, J.W., "The Usability Problem for Home Appliances: Engineers Caused It, Engineer Can Fix It!", *ACM Symposium on Foundations of Software Engineering (SIGSOFT)*, ISBN: 1-58113-855-5, pp. 95, October 2004

* Nikovski, D.N.; Brand, M.E., "Exact Calculation of Expected Waiting Times for Group Elevator Control", *IEEE Transactions on Automatic Control*, ISSN: 0018-9286, Vol. 49, Issue 10, pp. 1820-1823, October 2004 (TR2004-148)

Li, X.; Porikli, F.M., "A Hidden Markov Model Framework for Traffic Event Detection Using Video Features", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, Vol. 5, pp. 2901-2904, October 2004 (TR2004-130)

Porikli, F.M., "Nonlinear Warping Function Recovery by Scan-Line Search Using Dynamic Programming", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, Vol. 3, pp. 1807-1810, October 2004 (TR2004-031)

Kong, H.S.; Nie, Y.; Vetro, A.; Sun, H.; Barner, K., "Adaptive Fuzzy Post-Filtering for Highly Compressed Video", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, Vol. 3, pp. 1803-1806, October 2004 (TR2004-131)

Wang, B.; Wang, Y.; Selesnick, I.; Vetro, A., "An Investigation of 3D Dual-Tree Wavelet Transform for Video Coding", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, Vol. 2, pp. 1317-1320, October 2004 (TR2004-132)

Xie, L.; Kennedy, L.; Chang, S-F; Divakaran, A.; Sun, H.; Lin, C-Y, "Discovering Meaningful Multimedia Patterns with Audio-Visual Concepts and Associated Text", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, Vol. 4, pp. 2383-2386, October 2004 (TR2004-128)

Divakaran, A.; Peker, K.A.; Chang, S-F; Radhakrishnan, R.; Xie, L., "Video Mining: Pattern Discovery Versus Pattern Recognition", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, Vol. 4, pp. 2379-2382, October 2004 (TR2004-127)

Dalley, G.; Freeman, B.; Marks, J.W., "Single-Frame Text Super-Resolution: A Bayesian Approach", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, Vol. 5, pp. 3295-3298, October 2004 (TR2004-129)

Chen, J.; Khisti, A.; Malioutov, D.M.; Yedidia, J.S., "Distributed Source Coding Using Serially-Concatenated-Accumulate Codes", *IEEE Information Theory Workshop (ITW)*, pp. 209-214, October 2004 (TR2004-111)

* Lee, J.C.; Dietz, P.H.; Leigh, D.L.; Yerazunis, W.S.; Hudson, S.E., "Haptic Pen: A Tactile Feedback Stylus for Touch Screens ", *ACM Symposium on User Interface Software and Technology (UIST)*, ISBN: 1-58113-957-8, pp. 291-294, October 2004 (TR2004-133)

* Lee, J.C.; Dietz, P.H.; Maynes-Aminzade, D.; Raskar, R.; Hudson, S.E., "Automatic Projector Calibration with Embedded Light Sensors", *ACM Symposium on User Interface Software and Technology (UIST)*, ISBN: 1-58113-957-8, pp. 123-126, October 2004 (TR2004-036)

Digham, F.F.; Mehta, N.B.; Molisch, A.F.; Zhang, J., "Joint Pilot and Data Loading Technique for MIMO Systems Operating with Covariance Feedback", *IEEE International Conference on 3G Mobile Communication Technologies (3G)*, pp. 24-28, October 2004 (TR2004-105)

Feris, R.; Raskar, R.; Tan, K.; Turk, M., "Specular Reflection Reduction with Multi-Flash Imaging", *Brazilian Symposium on Computer Graphics and Image Processing (Sibgrapi)*, ISSN: 1530-1834, pp. 316-321, October 2004 (TR2004-098)

Radhakrishnan, R.; Divakaran, A.; Xiong, Z., "A Time Series Clustering Based Framework for Multimedia Mining and Summarization Using Audio Features", *ACM SIGMM International Workshop on Multimedia Information Retrieval*, ISBN: 1-58113-940-3, pp. 157-164, October 2004 (TR2004-046)

Kuwahara, N.; Hata, T.; Nozawa, T.; Vetro, A., "Network Transcoder with Seamless Switching Function", *SPIE Multimedia Systems and Applications*, Vol. 5600, pp. 1-10, October 2004 (TR2004-146)

Wu, M.; Vetro, A.; Chen, C.W., "Multiple-Description Image Coding with Distributed Source Codeing and Side Information", *SPIE Multimedia Systems and Applicaitons VII*, Vol. 5600, pp. 120-127, October 2004 (TR2004-147)

Molisch, A.F., "UWB Wireless Channels - Propagation Aspects and Interplay with System Design", *European Microwave Conference (EuMC)*, Vol. 2, pp. 1101-1104, October 2004 (TR2005-073)

Yuan, Y.; Gu, D.; Arbaugh, W.; Zhang, J., "High-Performance MAC for High-Capacity Wireless LANs", *International Conference on Computer Communications and Networks (ICCCN)*, ISSN: 1095-2055, pp. 167-172, October 2004 (TR2005-080)

* Chen, W.; Ren, L.; Zwicker, M.; Pfister, H., "Hardware-Accelerated Adaptive EWA Volume Splatting", *IEEE Visualization*, ISBN: 0-7803-8788-0, pp. 67-74, October 2004 (TR2005-009)

* Wolf, P. Woelfel, J.K.; van Gemert, J.; Raj, B.; Wong, D., "SpokenQuery: An Alternate Approach to Choosing Items with Speech", *International Conference on Speech and Language Processing (ICSLP)*, ICSLP 2004, October 2004 (TR2004-121)

Smaragdis, P., "Discovering Auditory Objects Through Non-Negativity Constraints", *Statistical and Perceptual Audio Processing (SAPA)*, SAPA 2004, October 2004 (TR2004-094)

Wren, C.R.; Reynolds, C.J., "Minimalism in Ubiquitous Interface Design", *Personal and Ubiquitous Computing*, ISSN: 1617-4909, Vol. 8, Issue 5, pp. 370-373, September 2004 (TR2002-022)

* Seltzer, M.L.; Raj, B.; Stern, R.M., "Likelihood-Maximizing Beamforming for Robust Hands-Free Speech Recognition", *IEEE Transactions on Speech and Audio Processing*, ISSN: 1063-6676, Vol. 12, Issue 5, pp.489-498, September 2004 (TR2004-088)

Yedidia, J.S.; Pedagani, K.; Molisch, A.F., "New Spreading Transforms for Fading Channels", *Allerton Conference on Communication Control and Computing*, Allerton Conference on Communication Control & Computing, September 2004 (TR2004-120)

* Yu, J.; Yao, Y-D; Zhang, J., "Reverse-link Capacity of Power-Controlled CDMA Systems with Beamforming", *IEEE Transactions on Vehicular Technology*, ISSN: 0018-9545, Vol.

53, Issue 5, pp. 1423-1433, September 2004 (TR2004-150)

Yuan, Y.; Gu, D.; Arbaugh, W.; Zhang, J., "Achieving Packet-Level Quality of Service Through Scheduling in Multirate WLANs", *IEEE Vehicular Technology Conference (VTC)*, ISSN: 1090-3038, Vol. 4, pp. 2730-2734, September 2004 (TR2004-117)

Du, J.; Li, Y.; Gu, D.; Molisch, A.F.; Zhang, J., "Estimation of Performance Loss Due to Delay in Channel Feedback in MIMO Systems", *IEEE Vehicular Technology Conference (VTC)*, ISSN: 1090-3038, Vol. 3, pp. 1619-1622, September 2004 (TR2004-118)

Sudarshan, P.; Mehta, N.B.; Molisch, A.F.; Zhang, J., "Spatial Diversity and Channel Statistics-Based RF-Baseband Co-Design for Antenna Selection", *IEEE Vehicular Technology Conference (VTC)*, ISSN: 1090-3038, Vol. 3, pp. 1658-1662, September 2004 (TR2004-077)

Mehta, N.B.; Digham, F.F.; Molisch, A.F.; Zhang, J., "Rate of MIMO Systems with CSI at Transmitter and Receiver from Pilot-Aided Estimation", *IEEE Vehicular Technology Conference (VTC)*, ISSN 1090-3038, Vol. 3, pp. 1575-1579, September 2004 (TR2004-092)

Tan, K-H; Kobler, J.; Feris, R.; Dietz, P.H.; Raskar, R., "Shape Enhanced Surgical Visualizations and Medical Illustrations with Multi-Flash Imaging", *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, September 2004 (TR2005-068)

Karedal, J.; Wyne, S.; Almers, P.; Tufvesson, F.; Molisch, A.F., "Statistical Analysis of the UWB Channel in an Industrial Environment", *IEEE Vehicular Technology Conference (VTC)*, ISSN: 1090-3038, Vol. 1, pp. 81-85, September 2004 (TR2005-074)

Smaragdis, P., "Non-negative Matrix Factor Deconvolution; Extraction of Multiple Sound Sources from Monophonic Inputs", *International Congress on Independent Component Analysis and Blind Signal Separation (ICA)*, ISBN: 3-540-23056-4, Vol. 3195/2004, pp. 494, September 2004 (TR2004-104)

Siefkes, C.; Assis, F.; Chhabra, S.; Yerazunis, W.S., "Combining Winnow and Orthogonal Sparse Bigrams for Incremental Spam Filtering ", *European Conference on Machine Learning (ECML) and the European Conference on Principles and Practices of Knowledge Discovery in Databases (PKDD)*, September 2004 (TR2004-025)

Sudarshan, P.; Mehta, N.B.; Molisch, A.F.; Zhang, J., "Antenna Selection with RF Pre-Processing: Robustness to RF and Selection Non-Idealities", *IEEE Radio and Wireless Conference (RAWCON)*, pp. 391-394, September 2004 (TR2004-103)

Brand, M.E., "From Subspaces to Submanifolds", *British Machine Vision Conference (BMVC)*, September 2004 (TR2004-134)

Catovic, A.; Sahinoglu, Z., "Hybrid TOA/RSS and TDOA/RSS Location Estimation Schemes for Short-Range Wireless Networks", *Bechtel Telecommunication Technical*

Journal (BTTJ), Vol. 2., Issue 2, pp. 77-84, September 2004 (TR2004-096)

Seltzer, M.L.; Raj, B.; Stern, R.M., "A Bayesian Framework for Spectrographic Mask Estimation for Missing Feature Speech Recognition", *Speech Communication*, Vol. 43, Issue 4, pp. 379-393, September 2004 (TR2004-086)

Du, J.; Li, Y.; Gu, D.; Molisch, A.F.; Zhang, J., "Space-Time LDPC with Layered Structure for MIMO Systems", *International Symposium on Wireless Personal Multimedia Communications (WPMC)*, September 2004 (TR2004-119)

Raj, B.; Seltzer, M.L.; Stern, R.M., "Reconstruction of Missing Features for Robust Speech Recognition", *Speech Communication*, Vol. 43, Issue 4, pp.275-296, September 2005 (TR2004-087)

Wu, J.; Horng, H.; Zhang, J.; Molisch, A.F.; Xiao, C., "Adaptive Transmit Diversity with Quadrant Phase Constraining Feedback", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Vol. 3, pp. 2262-2266, September 2004 (TR2004-109)

Ramachandran, I.; Nakache, Y-P; Orlik, P.V.; Molisch, A.F.; Zhang, J., "Symbol Spreading for Ultrawideband Systems Based on Multiband OFDM", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Vol. 2, pp. 1204-1209, September 2004 (TR2005-075)

Raskar, R.; van Baar, J.; Willwacher, T.; Rao, S., "Quadric Transfer for Immersive Curved Screen Displays", *Eurographics (EG)*, August 2004 (TR2004-034)

Wang, H.; Raskar, R.; Ahuja, N., "Seamless Video Editing", *IEEE International Conference on Pattern Recognition (ICPR)*, ISBN: 1051-4651, ISSN: 0-7695-2128-2, Vol. 3, pp. 858-861, August 2004 (TR2005-066)

Porikli, F.M., "Clustering Variable Length Sequences by Eigenvector Decomposition Using Hmm", *International Workshop on Structural and Syntactic Pattern Recognition (SSPR)*, ISSN: 0307-9743, Vol 3138/2004, pp. 352, August 2004 (TR2004-085)

Leigh, D.L.; Lanning, T.; Lesh, N.; Ryall, K.; Miyashita, H.; Makino, S., "Exhaustive Generation and Visual Browsing for Radiation Patterns of Linear Array Antennas", *International Symposium on Antennas and Propagation (ISAP)*, August 2004 (TR2004-057)

* Matusik, W.; Pfister, H., "3D TV: A Scalable System for Real-Time Acquistion, Transmission and Autostereoscopic Display of Dynamic Scenes", *ACM Transactions on Graphics (TOG) SIGGRAPH*, ISSN: 0730-0301, Vol. 23, Issue 3, pp. 814-824, August 2004 (TR2004-067)

* Raskar, R.; Beardsley, P.A.; van Baar, J.; Wang, Y.; Dietz, P.H.; Lee, J.H.; Leigh, D.L.; Willwacher, T., "RFIG Lamps: Interacting with a Self-Describing World via Photosensing Wireless Tags and Projectors", *ACM Transactions on Graphics (TOG) SIGGRAPH*, ISSN:

0730-0301, Vol. 23, Issue 3, pp. 406-415, August 2004 (TR2004-062)

* Raskar, R.; Tan, K-H.; Feris, R.; Yu, J.; Turk, M., "Non-photorealistic Camera: Depth Edge Detection and Stylized Rendering Using Multi-Flash Imaging", *ACM Transactions on Graphics (TOG) SIGGRAPH*, ISSN: 0730-0301, Vol. 23, Issue 3, pp. 679-688, August 2004 (TR2004-050)

Lee, J.C.; Avrahami, D.; Hudson, S.E.; Forlizzi, J.; Dietz, P.H.; Leigh, D.L., "The Calder Toolkit: Wired and Wireless Components for Rapidly Prototyping Interactive Devices", *Designing Interactive Systems (DIS)*, ISBN: 1-58113-787-7, pp. 167-175, August 2004 (TR2004-076)

Singh, R.; Raj, B., "Classification in Likelihood Spaces", *Technometrics*, ISSN: 0040 1706, Vol. 46, No. 3, pp. 318-329, August 2004 (TR2004-089)

Smaragdis, P., "Exploiting Redundancy to Construct Listening Systems", *Perspectives on Speech Separation*, Pierre Divenyi (ed.), Kluwer Academic Publishers (New York), July 2004

* Pfister, H.; Gross, M., "Point-Based Computer Graphics", *IEEE Computer Graphics and Applications*, ISSN: 0272-1716, Vol. 24, Issue 4, pp. 22-23, July 2004

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.

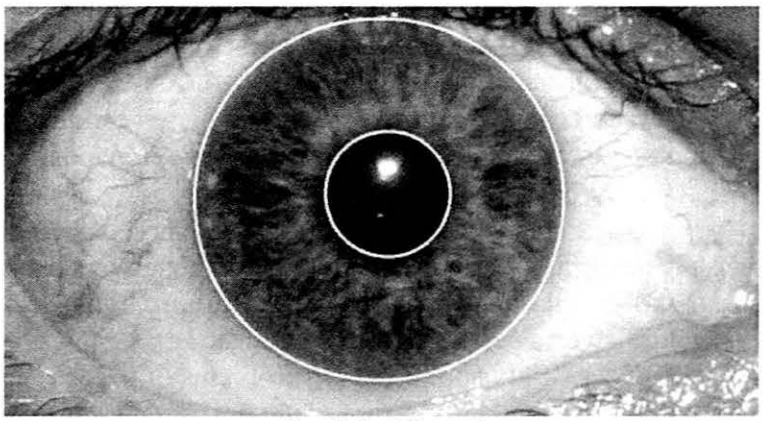
Much of the computer vision research at MERL is focused on of surveillance. For example, MERL has pioneered a state of the art approach to detecting object classes such as human faces in cluttered scenes. This approach uses a powerful machine-learning framework to automatically build very fast object detectors given a set of positive and negative examples of the object class. The same approach has been successfully applied to the problems of pedestrian detection, facial feature finding, face recognition, and gender and race classification. 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, audio-visual event detection, intelligent video browsing systems and fusion of imaging sensors. These systems are being applied to many areas of MELCO's businesses such as surveillance and security, consumer products (cell phones and DVD players) and elevators.

Project Descriptions

Iris Recognition from 1-2 Meters 52
2-D Face Recognition 53
3D Face Recognition 54
Face Scanning 55
Multilinear Face Models 56
Happy and Sad Face Classifiers 57
Pedestrian Detection 58
Moving Cast Shadow Detection 59
Low-Frame-Rate Tracking 60
Waviz Background Models 61
Ensemble Tracking 62
Blind Vision 63
MERL Optic Touch 64
VideoRule - Automatic Integration of Video in Databases 65
Context-Aware Pan-Tilt-Zoom Cameras 66
Single-Axis Multi-Parameter (SAMP) Camera 67
Human Activity Determination 68
Surface Reconstruction 69
Audio-Visual Event Detection for Consumer and Surveillance Video 70
DiamondClassify 71

Iris Recognition from 1-2 Meters



See Color Figure 1

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 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.

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.

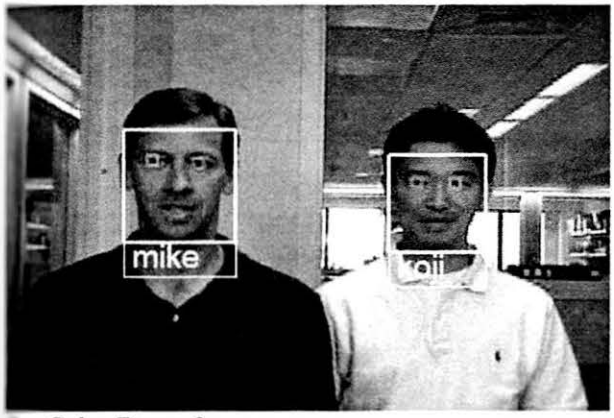
Technical Discussion: We intend to use MERL's state-of-the-art face detection and eye detection technology to enable a camera to be aimed at the iris from up to 2 meters away. With the right camera and lens, a high resolution image of the eye region can be captured at that distance. Once the eye region 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.

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



See Color Figure 2

The goal of this project is to have a computer recognize a person from an image of his or her face. There are many applications for face recognition. Some examples are: 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). The next step is to compare two aligned faces to get a similarity score. We have experimented with two techniques for computing a similarity score - one based on local rectangle features and the other based on Bayesian matching.

Background and Objectives: The problem of face recognition has been studied by a number of researchers. Our objective is to develop a state of the art system and give MELCO some intellectual property in this area. We are currently 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. A number of business units have expressed their desire to build a product around face recognition technology.

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 (rectangle features) that differentiate two face images of the same person from 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.

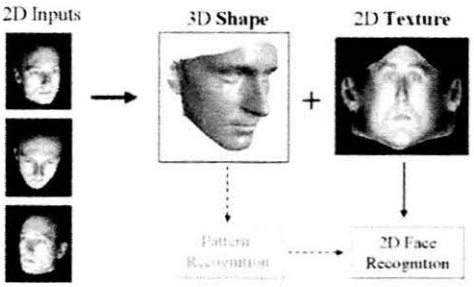
Collaboration: Joint work with the Physical Security Project at Sentansoken. Inaden is planning to productize for access control and face logging. Discussions with Kamaden about access control product. Discussions with Kyoden.

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.

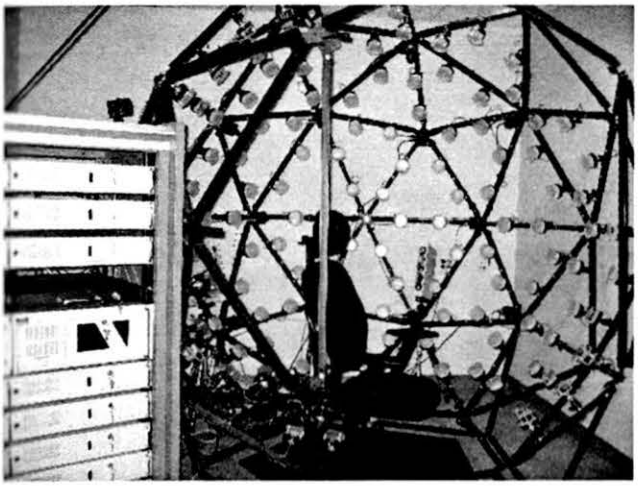
Technical Discussion: We use a linear combination of "eigenheads" obtained by Principal Component Analysis (PCA) of a training set of laser-scanned 3D human faces. The PCA coefficients are used as model parameters. We establish correspondence between faces with an efficient error metric (boundary weighted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

Multilinear Face Models

A multilinear face model is a simple function that can generate a wide variety of realistic face images. The images can vary by many attributes, including identity, expression, speech articulation, and 3D pose. We have developed all the tools needed to estimate such a model and use it to compress and edit video of faces, for example, changing the appearance and performance of an actor in existing video. The image illustrates these tools being used to combine expression, identity, viseme, and 3D pose from the two inset videos, synthesizing the central video.

Background and Objectives: It has long been conjectured that the appearance, expression, mouth shape, pose, etc. of a photographed face could be encoded with a relatively short vector of numbers yet reconstructed with high accuracy. Finding such an encoding is essential for extreme video compression, biometrics, visual user interfaces, etc. We noted that multilinear functions could provide a parameterization of faces that gives the editor separable control over attributes such as appearance, expression, age, etc. Thus one could preserve the fact that every individual smiles in a unique manner, and guarantee that the coding of the smile is the same across all individuals. This would open up interesting applications in film post-production, foreign language dubbing, and surgical modeling. Therefore we set out to develop methods to estimate multilinear models from facial scan data, to recover multilinear facial parameters from ordinary video, and to resynthesize such videos with interesting changes.

Technical Discussion: We estimated a multilinear model from a set of high-resolution 3D face scans taken of several people in various expressions and speech articulations. Because a standard estimator would require a scan of every individual in every expression and articulation, we developed a fast imputative method to estimate multilinear models from incomplete data. We then incorporated the model into an optical flow based visual tracker that estimates identity, 3D pose, expression, etc., from ordinary video. This encodes the facial area of the video, and also gives a detailed estimate of the 3D shape of the face in every frame. We can then mix and match attributes from several such parameter streams, use the multilinear model to obtain the corresponding 3D shape and texture, and render the results back into the video.

Collaboration: Daniel Vlasic and Jovan Popovic, MIT.

Future Direction: We are interested in modeling age variation and in applications on cell-phone/camera platforms.

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

Lab: MERL Research Lab
Project Type: Research

Happy and Sad Face Classifiers

The goal of this project is to have a computer determine if a detected face is happy, sad or neither. The algorithm works on either photographs or video. Each frontal face that is detected in an image is classified according to its facial expression. The same approach can be used to determine other characteristics of a person from his or her face such as age, gender or race as well as other facial expressions.

Background and Objectives: The happy and sad classifiers use the same machine learning framework as the face detector to learn two classifiers for happy and sad faces. The resulting classifiers are very fast. One use for this technology is for marketing purposes, for example, to determine if a customer is enjoying an advertisement or store display. Similarly, age and gender classifiers could be very useful for gathering marketing data.

Technical Discussion: To classify a face as happy or sad, the fast face detector developed at MERL is first used to find all faces in an image. Then the face is normalized to fix its scale, rotation and position. This is done by finding nine facial feature points (such as the corners of the eyes, etc) just as in the face recognition technology developed at MERL. Once the face is normalized, the separate happy and sad classifiers are evaluated on the face image. If either classifier scores above threshold then the face is classified appropriately as happy or sad.

The challenge of learning a happy or sad classifier is to find a set of features that can be used to distinguish smiling (frowning) face images from other face images. This is done using a machine learning framework which is given examples of smiling (frowning) faces and examples of every other facial expression. The AdaBoost learning algorithm is used to select a set of simple rectangle features that best discriminate smiling (frowning) faces from all other facial expressions. If the classifiers are being run on a video sequence of a face, the accuracy is improved by integrating the classifiers scores over time.

Future Direction: In the future, other facial expression classifiers could be trained. Classifiers for other characteristics could also be trained such as gender, age and race.

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

Lab: MERL Technology Lab
Project Type: Research

Pedestrian Detection

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

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

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

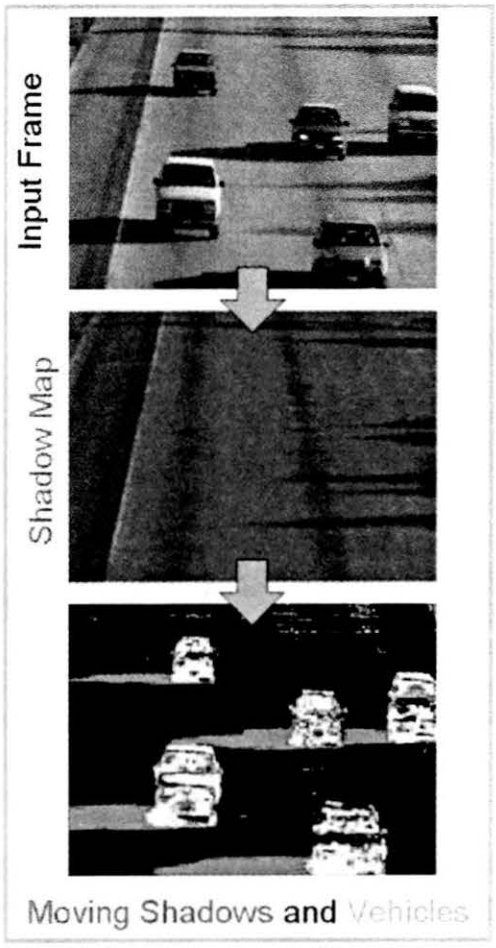
Collaboration: Discussions with the Physical Security System Project at Sentansoken, the Visual Communication Systems Department at Gunden and the Imaging Storage Products Department at Kyoden.

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

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

Lab: MERL Technology Lab
Project Type: Research

Moving Cast Shadow Detection



See Color Figure 3

We developed a novel algorithm to detect and remove moving cast shadows in a video sequence.

Background and Objectives: Cast shadows poses one of the most challenging problems in many vision tasks, especially in object tracking, by distorting the true shape and color properties of the target objects. They correspond to the areas in the background scene that are blocked from the light source. It is essential to eliminate only cast shadows since removal of self shadows, which are the parts of the object that are not illuminated, will result in incomplete object silhouettes.

Technical Discussion: We remove cast shadows by taking advantage of the statistical prevalence of the shadowed regions over the object regions. We model shadows using multivariate Gaussians. We apply a weak classifier as a pre-filter. We project shadow models into a quantized color space to update a shadow flow function. We use shadow flow, background models, and current frame to determine the shadow and object regions.

This method has several advantages: 1) It does not require a color space transformation. We pose the problem in the RGB color space, and we can carry out the same analysis in other Cartesian color spaces as well. 2) Our data driven method dynamically adapts the shadow models to the changing shadow conditions. In other words, accuracy of our method is not limited by the preset threshold values, which is a major drawback of the existing approaches. The accuracy of this method improves as it process more video frames. 3) Furthermore, it does not assume any 3D models for the target objects or tracking of the cast shadows between frames.

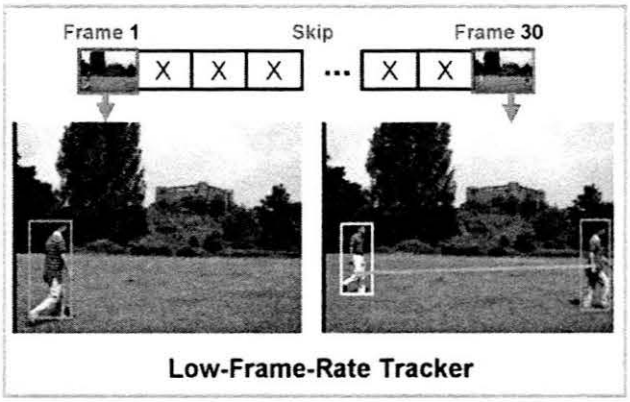
Collaboration: Sentansoken.

Future Direction: We will integrate the shadow removal method for several surveillance and traffic applications, particularly for detection and recognition tasks that require accurate extraction of object boundaries.

Contact: Fatih Porikli
http://www.merl.com/projects/shadow-detection/

Lab: MERL Technology Lab
Project Type: Research

Low-Frame-Rate Tracking



See Color Figure 5

As a part of the Physical Security SK-Pro, we have developed a robust, computationally efficient multi-object tracking method that runs for low-frame-rates as well as fast moving object scenarios.

Background and Objectives: Advanced video surveillance systems are assembled from a large number of cameras. It is desired to achieve real-time tracking performance while keeping the hardware costs on a minimum level in such systems. Therefore, it is necessary to process the vast amount of constantly streaming channels of video on a single CPU at the same time. However, most existing tracking approaches presume they can consume all the available processing power to track objects in a single sequence.

One solution to the challenging problem of processing of multiple video sequences on the same CPU is to sample every input video such that the number of frames per second is decreased proportional to the number of sequences. However, due to the decrease of the frame rate, the tracking algorithm receives video frames at a lower temporal resolution, which causes the objects to appear reciprocally much faster than to the original sequence. Since object movements are usually large and unpredictable, existing tracking approaches fail to detect the target objects in the current frame.

Technical Discussion: We present an object tracking algorithm for low-frame-rate applications. We assign multiple kernels centered on high motion areas. We improve the convergence properties of the mean-shift by integrating two additional likelihood terms. Unlike the existing approaches, the proposed algorithm enables tracking of moving objects at lower temporal resolutions as much as 1-fps frame rate without sacrificing the robustness and accuracy. Therefore, it can process multiple videos at the same time on a single processor. Note that, the low frame rate constraint corresponds to the fast motion of the moving objects. Thus, the proposed method is capable of tracking fast objects even in the original frame rates.

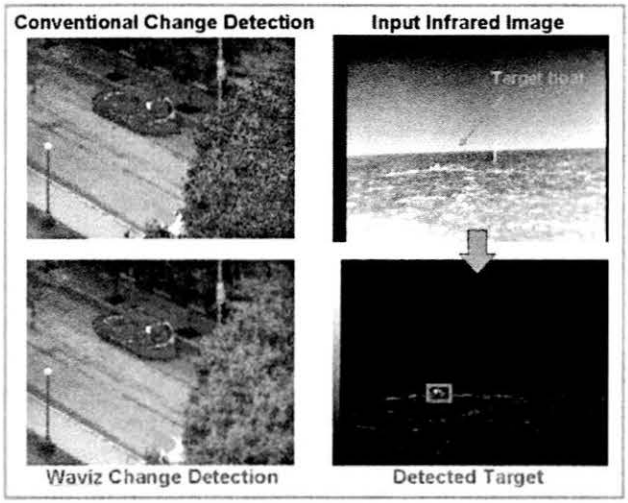
Collaboration: Sentansoken

Future Direction: We will extend our low-frame-rate tracking techniques to detect different types of objects such as vehicles, human, etc. in a wide range of applications using multiple PTZ and omni-directional cameras.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

Waviz Background Models



The main contribution of this work is an algorithm that explicitly harnesses the scene dynamics to improve segmentation. Our algorithm detects new objects based solely on the dynamics of the pixels in a scene, rather than their appearance. Thus we can distinguish a swaying tree from a moving vehicle.

Background and Objectives:
Background subtraction is the most common approach for discriminating a moving object in a relatively static scene. Basically, a reference model (background) for the stationary part of the scene is estimated and the current image is compared with the reference to determine the changed regions (foreground) in the image.

A major shortcoming of all the above background methods is that they neglect the temporal correlation among the previous values of a pixel. This prevents them detecting a structured or periodic change, which is often the case since real-world physics induces near-periodic phenomenon in the environment: the motion of plants driven by wind, the action of waves on a beach, and the appearance of rotating objects.

Technical Discussion: We directly estimate models of cyclo-stationary processes to explain the observed dynamics of the scene and then comparing new observations against those models.

We generate a representation of the background using the frequency decompositions of the pixel's history. For a given frame, we compute the frequency transform coefficients (either FFT or DCT) and compare them to the background coefficients to obtain a distance map for the frame. Then, we fuse the distance maps in the same temporal window of the transform coefficients to improve the robustness against the noise and to remove the trail artifacts. Finally, we apply a threshold to the distance maps to determine the foreground pixels.

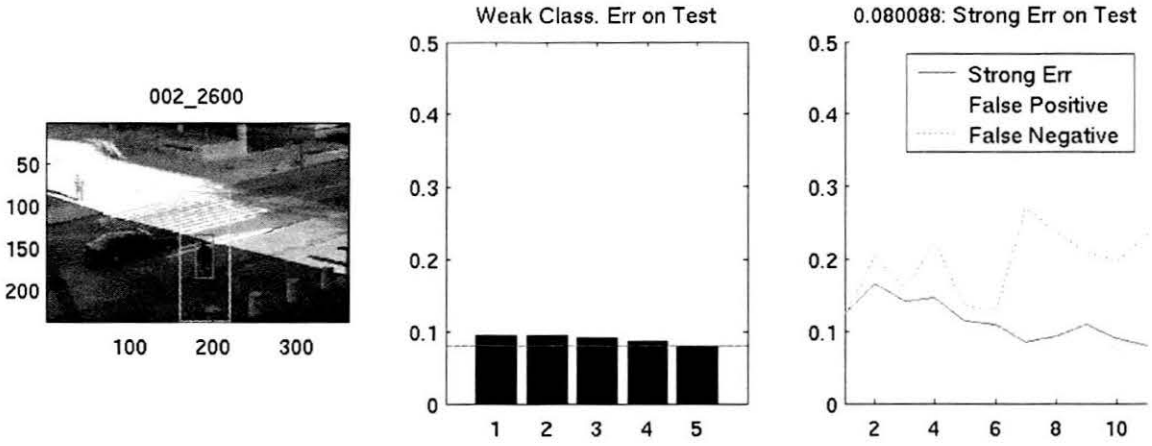
Collaboration: Sentansoken.

Future Direction: This method is very promising in detection of moving objects in periodic backgrounds, e.g. finding people and vehicles in outdoor scenes that depict sea, waves, moving trees, etc. We will improve the computational load of the Waviz Background Modeling method by using recursive frequency coefficient estimation techniques.

Contact: Fatih Porikli, Christopher R. Wren
http://www.merl.com/projects/waviz/

Lab: MERL Technology Lab
Project Type: Research

Ensemble Tracking



We have developed a general tracker that can be used in various scenarios such as surveillance and human-computer interaction. The tracker can track a wide variety of objects (e.g., pedestrians, faces, cars, boats) using either a static or dynamic camera. It is computationally efficient, robust, runs in real time and can work with different image types (e.g., color, gray-scale, infra-red).

Background and Objectives: Tracking algorithms usually focus on the object to be tracked, while neglecting the environment in which the object is moving. Furthermore, most tracking algorithms work with pre-defined features that might not be able to distinguish the object from the background. Our tracker can automatically find the best set of features to separate the object from the background, as well as integrate and update this information over time.

Technical Discussion: We treat tracking as a binary classification problem where the object is to be distinguished from the background. We use AdaBoost to construct, online, an ensemble of weak classifiers that are trained one per frame. In addition, we developed an online update rule for ensemble training that allows us to deal with time-varying distributions that arise from the change in appearance of the object and the background. The result is a dynamic ensemble of weak classifier that allow for robust and computationally efficient tracking.

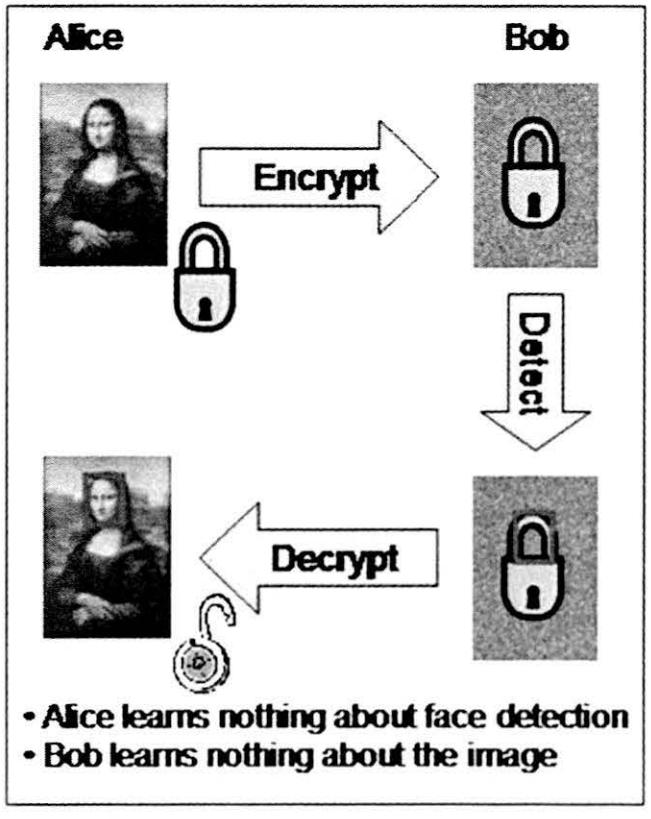
Collaboration: Sentansoken

Future Direction: We are working to integrate the tracker into a future release of MERL's tracking application.

Contact: Shai Avidan
http://www.merl.com/projects/ensemble-tracking/

Lab: MERL Research Lab
Project Type: Research

Blind Vision



Privacy issues are a growing concern among consumers and governments alike. The spread of surveillance cameras only amplifies the problem. We have developed a general framework for secure signal processing to address these concerns. Our goal is to allow two parties, a service provider and a content provider to interact without revealing information to each other. A possible application is secure face detection where the service provider detects faces in the images of the content provider without learning anything about the images. This approach can be extended to other domains such as secure segmentation of medical data, or secure sound classification.

Background and Objectives: Secure multi-party computations allow for multiple parties to share knowledge and resources to achieve common goals while protecting the privacy of their data. Our goal is to use the properties of digital signals such as images or audio files to develop practical protocols for secure signal processing.

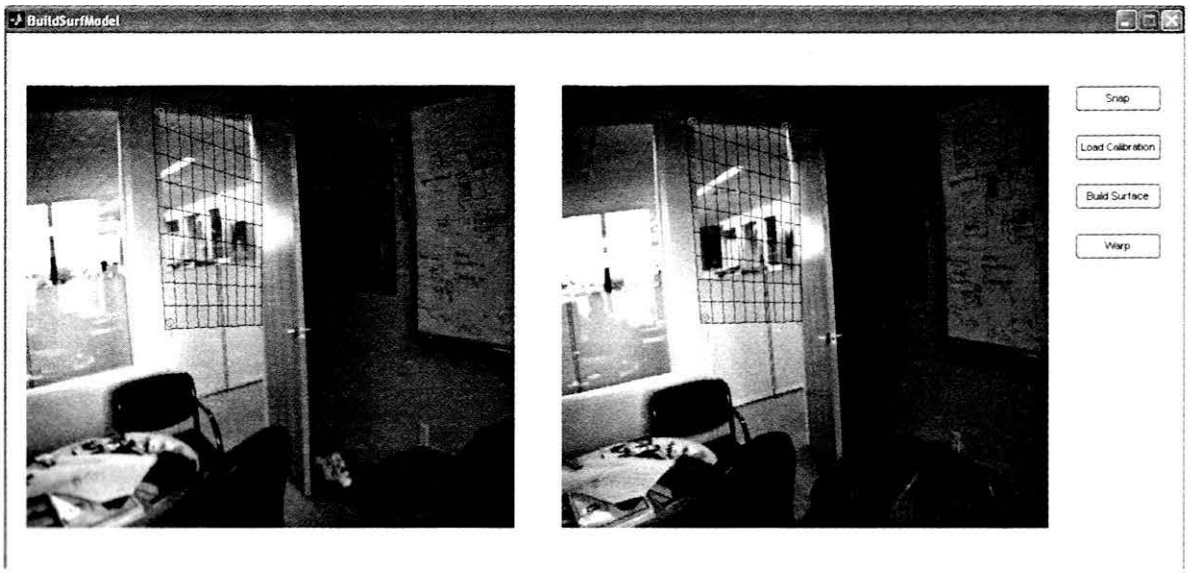
Technical Discussion: We adapt secure multi-party computation protocols to work with digital signals such as images and audio. By taking advantage of the unique properties of digital signals we are able to greatly accelerate existing protocols. In particular we have developed a secure face detection protocol that uses a novel image transformation to distort the content of an image beyond recognition yet let object detection algorithms work on the distorted image.

Future Direction: We currently work on a demo system.

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

Lab: MERL Research Lab
Project Type: Research

MERL Optic Touch



The MERL Optic Touch system will allow a user to detect a break in a virtual boundary around a guarded object from Video camera. We use a stereo camera and an efficient real-time depth segmentation algorithm to detect an event of penetration of a virtual boundary. The boundary can be given by selecting several points on the surface, or with giving a reference point (a center of the object) and a standard 3-dimensional shape, such as sphere for instance.

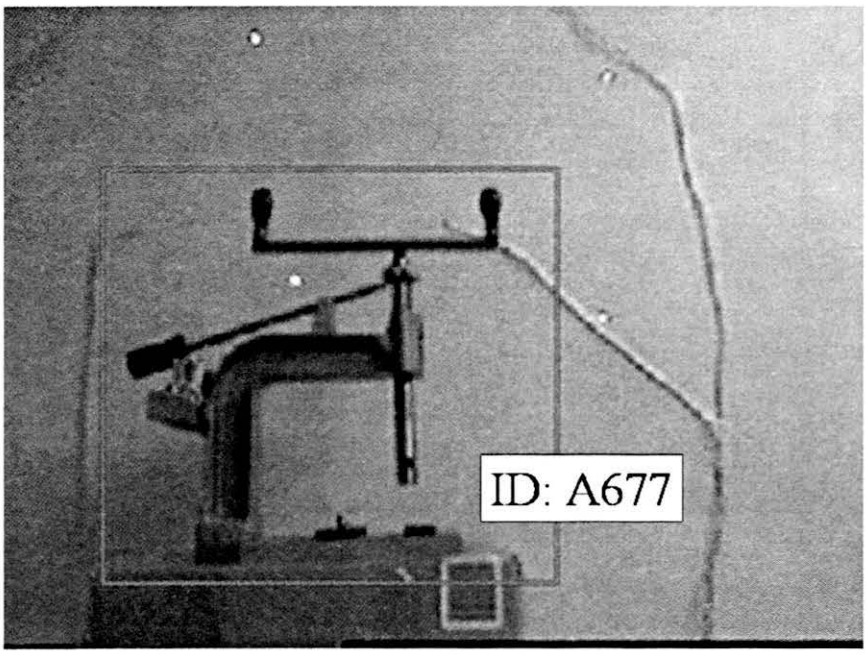
Background and Objectives: The general shortcoming of existing technologies is that they are not applicable in close quarters where projective properties of an extended object can not be ignored in the same way as they are when the object is located at a large distance from the camera. Our solution should make it possible to specify "Trip Surfaces" which may or may not be coincident with a physical surface.

Technical Discussion: The algorithm is based on an issued MERL patent on virtual margin touch detection. The fast depth segmentation algorithm uses precalculated disparity map which is calculated for an analytically specified surface. Any object that occupies the same space as a thin shell about the surface can be detected by a simple table look up. One image in the stereo pair is warped using the calculated disparity map and then subtracted from the other. This subtraction identifies the pixels that do not belong to the guarded surface. Repeating this calculation for a surface located just outside the given surface and identifying pixels in the image which are located outside of one surface shell but not the other allows us to find the event of penetration of the guarded perimeter.

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

Lab: MERL Technology Lab
Project Type: Research

VideoRule - Automatic Integration of Video in Databases



VideoRule is a system for recording video of a plant environment, for storage in an Asset Management database, with automatic cross-referencing between the objects in the video and the object descriptions already in the database. Subsequently an operator can access the database by specifying the name of a particular item in a plant in order to be presented with the video of that item. Alternatively the user can select an object in the video in order to retrieve its description (maintenance history, CAD drawing etc). In addition VideoRule allows the user to make 3D measurements like length, angle, area, volume, from the video.

Background and Objectives: The goal is to provide a support tool for an Asset Management database, to allow video to be easily added to a database with automatic cross-referencing between objects in the video and their description in the database.

Technical Discussion: There are two aspects to the problem: (a) Recovering the identity of objects in the video - we use visual markers (piecodes) on objects at the moment, with a view to using other technologies in the longer term; and (b) Recovering 3D information for the video - we use a hybrid stereo-laser system to recover motion of the video camera, thereby providing 3D context to the video, and allowing user-specified 3D measurements in the video.

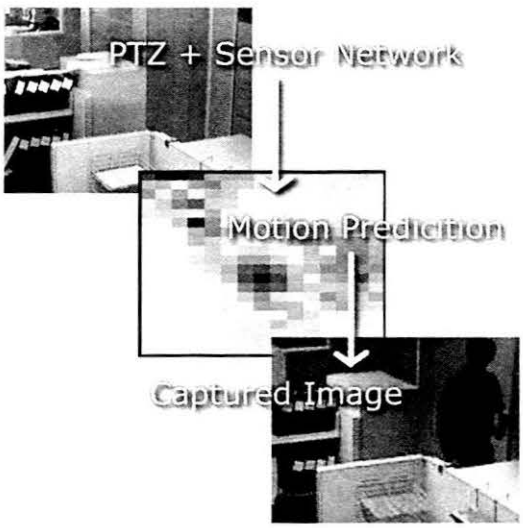
Collaboration: VideoRule Version 1.0 transferred to Dr Shiotani, ATC.

Future Direction: Further evaluation of the current system; automatic change detection.

Contact: Paul Beardsley, Jay Thornton
http://www.merl.com/projects/videorule/

Lab: MERL Research Lab
Project Type: Initial Investigation

Context-Aware Pan-Tilt-Zoom Cameras



This project explores the possibilities of combining Pan-Tilt-Zoom (PTZ) cameras omni-directional sensors and sensor networks. The goal is to create a fully autonomous PTZ system capable of reacting to the context to generate informative video streams with little or no human intervention. The project will explore PTZ control schemes, context sensing modalities, and also associated visualization technology.

Background and Objectives: Pan-Tilt-Zoom (PTZ) camera system provides the ability to recover unparalleled visual detail. However, current PTZ systems have very a limited field of view, and therefore depend on significant attention from a human operator to supply the contextual information necessary to realize the full potential of the sensor platform. This reliance on timely and skilled human intervention significantly limits the potential deployment of PTZ camera hardware. The goal of this project is to create a drop-in replacement for legacy, fixed camera systems that provide the advantages of upgrading to a PTZ system without the prohibitive cost of wiring and staffing that is normally associated with a PTZ installation.

Reviewing recorded video from a PTZ camera is a difficult task, particularly if the camera is not under human control. This project therefore also needs to address the problem of presenting PTZ video in an intelligible format. This presentation format should allow the operator to understand both the context that the PTZ is operating in and well as the specific information the PTZ is providing about the scene.

Technical Discussion: So far we have investigated the problem of automatic calibration. This is essential to reducing installation costs. We have built research prototypes that automatically calibrate PTZ camera to omnidirectional cameras and to sensor networks.

The problem of calibrating PTZ cameras to Sensor networks has forced us to consider a new approach: functional calibration. In functional calibration we seek to recover a description of the relationship between the camera and the sensor nodes that will allow us to make the best use of the PTZ camera. The system directly learns a policy that allows the PTZ camera to capture high-quality video of targets that are sensed by the network. This approach allows us to mix sensors with very different modalities: cameras and motion detectors, for example.

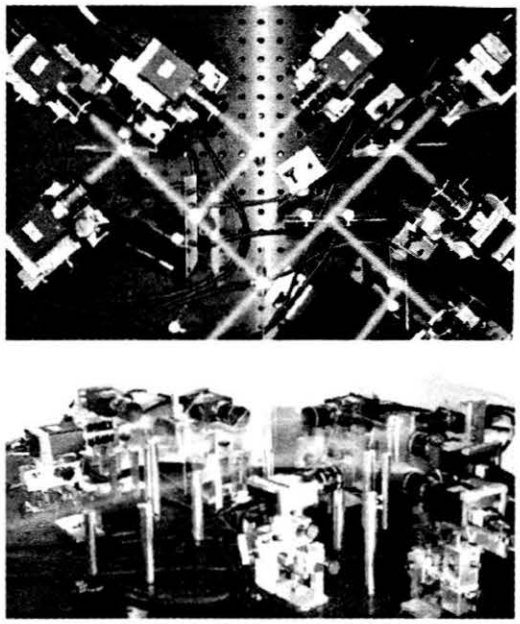
Collaboration: MTL.

Future Direction: Now that automatic calibration and reactive behavior have been successfully implemented, the project is focusing on the harder task of predicting good video opportunities.

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

Lab: MERL Research Lab
Project Type: Initial Investigation

Single-Axis Multi-Parameter (SAMP) Camera



See Color Figure 6

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.

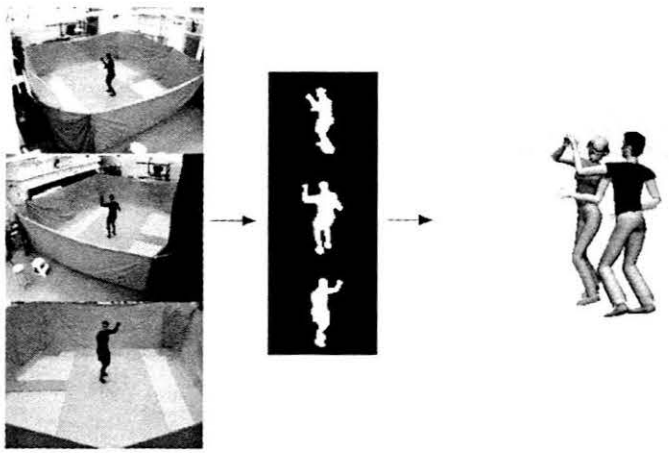
Collaboration: This is a collaboration with Morgan McGuire (Brown University), Shree Nayar (Columbia University), Fredo Durand (MIT), and John Hughes (Brown University).

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, Wojciech Matusik
http://www.merl.com/projects/samp/

Lab: MERL Research Lab
Project Type: Research

Human Activity Determination



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

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

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

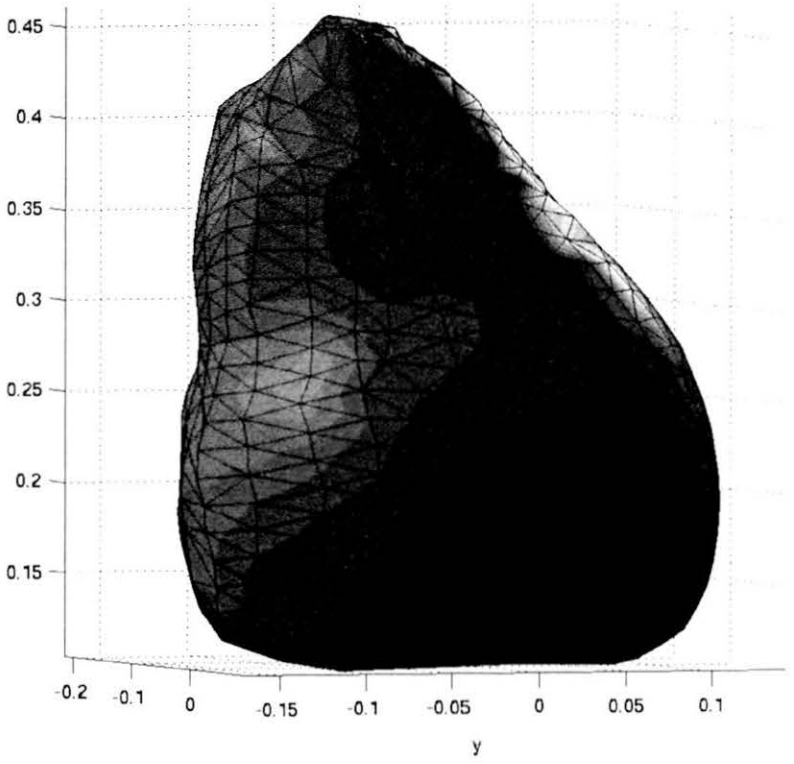
Collaboration: Jessica Hodgins and Liu Ren, CMU.

Future Direction: Investigate application to demographics collection for retail

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

Lab: MERL Research Lab
Project Type: Research

Surface Reconstruction



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

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

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

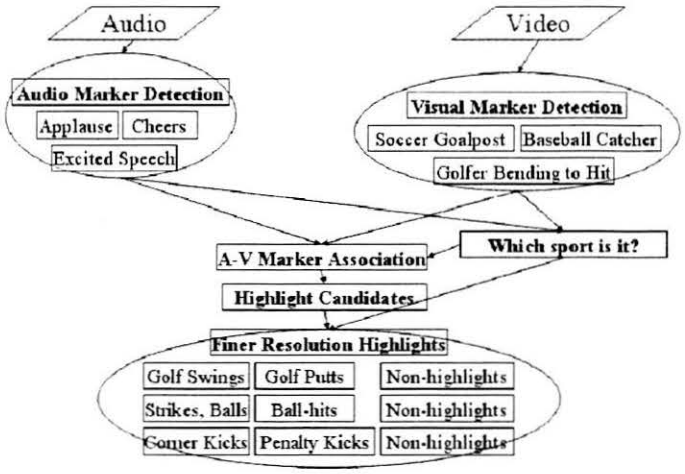
Collaboration: Kongbin Kang (Brown University)

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

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

Lab: MERL Research Lab
Project Type: Research

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" happenings. Therefore, by detecting this set of "interesting" events one can summarize the content. 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. Then, by associating a video marker with an audio marker one can capture the whole event. Modeling the low-level audio-visual features during the time interval between a video marker and an audio marker will enable us in further clustering the interesting events into finer categories.

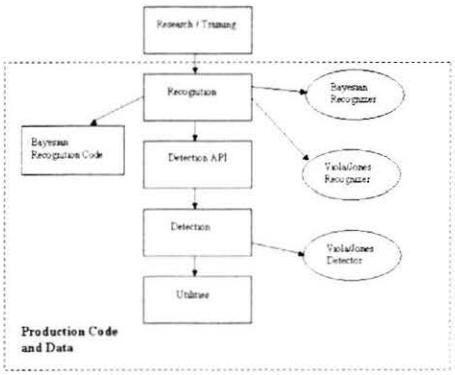
Collaboration: Sentansoken.

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

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

Lab: MERL Technology Lab
Project Type: Research

DiamondClassify



DiamondClassify code hierarchy

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

Background and Objectives: The MERL object detection and face recognition code had grown into a large and unwieldy library, containing both detection and recognition code intended for use in products, and significant amounts of obsolete and experimental research and training code. The primary goal of the DiamondClassify project has been to break this large monolithic library into several smaller, more manageable units by separating the code into libraries intended for use in products, and libraries solely used for research and training purposes.

Technical Discussion: Last year, our initial work on DiamondClassify focused on the product-level portions of the detection and recognition libraries. This year, our effort has focused on the research and training code, which resides in a separate library.

First, the training library was ported to use the existing API provided by the DiamondClassify product-level code. Next, the object detection and face recognition training programs were ported to run on the new training library, on a multiprocessor system. Finally, the face recognition code was modified to run on a single-processor system. As part of this project, significant effort was made to improve performance so that running on a single-processor system would be feasible. As a result, face recognition training now runs approximately 70 times faster on a single processor than in the past.

Finally, the Bayesian matching face recognition algorithm developed at MRL was ported to run within the DiamondClassify API. This allows the choice of recognizer to be made at run time by simply providing a different recognizer file to any program using the face recognition API.

Collaboration: We are working with the Image Recognition Systems Group and the Physical Security System Project at Sentansoken. We are also collaborating with the Internet Media System Department at Johosoken.

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

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

Lab: MERL Technology Lab
Project Type: Advanced Development

Digital Communications

Digital communications and networking are pervasive in today's society. Advanced technologies at the PHY, MAC, and network layers, will provide high speed communication capability for transmission of information with quality of service connecting people anywhere and anytime. From advanced multimedia systems to simple integrated home networking, communications and networking technologies is at the center of a continuing revolution.

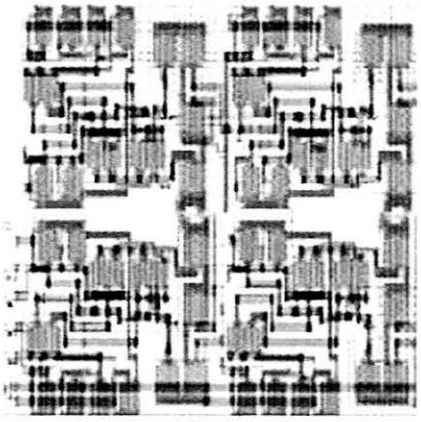
Our goal is to seek new business and technology trends in the area of digital communications and networking. We are conducting fundamental research on communication theory and developing new core technologies for emerging applications. Our focus is on broadband wireless communications, pervasive ad hoc networks, and digital home networking.

Ultra-wideband has attracted much attention. It provides very high data rates for short-range wireless communications. We have become an important player in the field. ZigBee is a promising networking technology for which we have had key contributions to. In broadband mobile communications, we continued to develop Multi-Input-Multi-Output technologies and contributed to the 3GPP standard. In the wireless LAN area, we continued IEEE802.11 development. We are also actively working on DLNA for integrated home networks.

Project Descriptions

RF Analog Logic 74
High-Speed Decoding of Error-Control Codes 75
ZigBee Protocol Stack Development 76
Data Broadcasting in ZigBee 77
Image Over ZigBee Networks 78
ZigBee HVAC Application Profile Development 79
ZigBee for Industrial Plant Monitoring 80
ZigBee/SCP for Home Automation 81
HDTV Transmission over UWB 82
High Precision UWB Ranging and Locating 83
Low Complexity UWB Transceiver with Multiband OFDM 84
Symbol Spreading in MB-OFDM UWB 85
Impulse Radio 86
MIMO-OFDM System for High Speed WLAN 87
Statistical Data Rate Allocation for Layered Space-time Systems 88
2D Normalized Min-Sum LDPC Decoding 89
Frame Aggregation for High Efficient MAC 90
Efficient Block-ACK Mechanism 91
HCCA Enhancements for High Speed WLAN 92
MIMO Multiple Cell System-level Simulator 93
Antenna Selection Signaling & Verification for Super 3G 94
Method for Approximating Lognormal Co-channel Interference 95
DLNA Digital Media Server (DMS) Design 96
Wearable Token for Secure Mobile Devices 97

RF Analog Logic



RF Analog Logic is a project in partnership with MIT, partially funded by DARPA with additional support from Cadence Design Systems, to enable sophisticated new signal processing capabilities for wireless communications systems by implementing statistical inference and machine learning algorithms natively in analog integrated circuits.

Background and Objectives: A very broad class of pattern recognition, signal-processing, and optimization algorithms can be implemented using a few transistors in Analog Logic rather than the 1000's required for the equivalent digital systems, offering better than order-of-magnitude improvements in overall speed, power and dynamic range. The enabling insight is that these algorithms are, at heart, analog continuous-time energy minimization processes, and can be made robust to noise and faults.

Unlike many other more ad hoc approaches to analog computation that have failed, Analog Logic creates a principled design flow to compile optimization programs into robust, scalable circuits. In addition to order-of-magnitude performance wins that have already been demonstrated for error-correction-decoding applications, Analog Logic now offers the opportunity to fully integrate sophisticated signal-processing and pattern recognition capabilities alongside conventional high-performance analog circuitry.

Technical Discussion: We are applying Analog Logic to tackle reducing the effects of interference on RF signal chains, a crucial challenge for robust wideband transceivers. Applications to RF signal processing include synchronization, decoding and demodulation, detection, recognition, interpolation, adaptive filtering and control, prediction and tracking, and RF system optimization. We are designing and simulating our RF Analog Logic circuits with Cadence software and our own JmpLab (Java message passing Laboratory) software for manufacture in the TSMC .18um RF (mixed-mode) process.

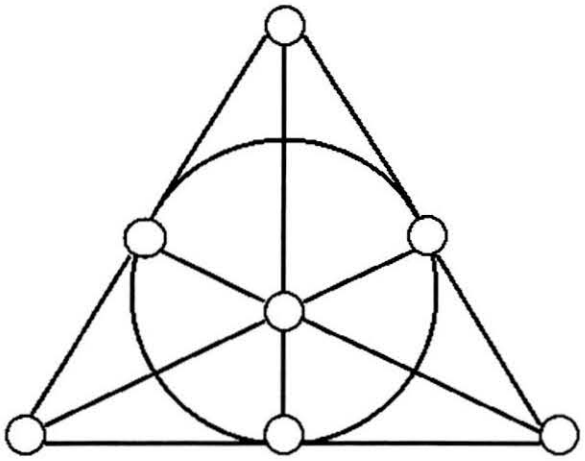
Collaboration: This project is supported in part by DARPA Seedling BAA 04-09 from the Advanced Technology Office (ATO). Cadence Design Systems has provided us with approximately $2.5 million in software licenses through their OpenChoice program to facilitate the development of IP that is interoperable under Cadence. We also have academic collaborations with Professor Neil Gershenfeld's and Professor Rahul Sarpeshkar's groups at MIT.

Future Direction: After testing out our core technology, our goal for the future is to obtain phase 2 DARPA funding. As we begin to apply RF Analog Logic to specific application domains, we will involve both our present collaborators, as well as CR&D in Japan, MELCO divisions, and others.

Contact: Ben Vigoda
http://www.merl.com/projects/analoglogic/

Lab: MERL Research Lab
Project Type: Research

High-Speed Decoding of Error-Control Codes



A simple finite geometry

We are designing high-speed decoding algorithms for error-control codes suitable for VLSI implementation. One of our primary target applications is very high speed (e.g. 40GBit/sec) communication on optical fiber channels.

Background and Objectives: High speed optical fibre channels pose a number of difficult requirements on error-control coding systems. The code must have a low overhead (high rate), it must perform well (i.e., not have an error-floor) down to very low bit-error rates, and the decoding algorithm must be simple enough so that it can easily be implemented in VLSI in order to reach the needed data rates.

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. For the optical fibre application, 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. We have developed a set of decoders based on replica-shuffled versions of iterative bit-flipping schemes that give excellent performance versus complexity/latency trade-offs, and we have shown that error-floors will not exist for these decoders.

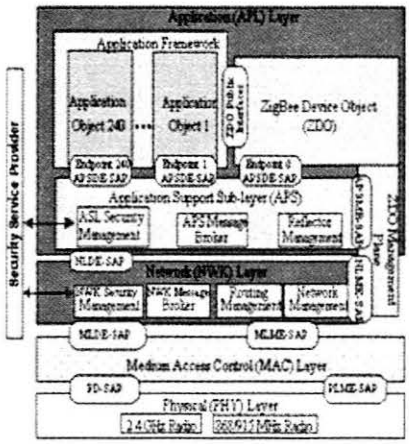
Collaboration: This work is in collaboration with Prof. Marc Fossorier from the University of Hawaii, his students Juntan Zhang and Yige Wang, and the error-control coding team at Johosoken

Future Direction: We plan to begin developing VLSI prototypes for these decoders.

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

Lab: MERL Research Lab
Project Type: Research

ZigBee Protocol Stack Development



ZigBee Stack Structure

ZigBee is a wireless multi-hop networking protocol defined by ZigBee Alliance to facilitate low bit-rate, low cast and low power-consumption sensor networks with features such as security, automatic address assignment and cost-aware routing. The potential applications for this technology include environmental and industrial monitoring, surveillance, home automation, automatic meter reading, tracking of goods and personnel, and military applications. MERL has developed ZigBee network protocol spec v1.0. Currently, we are developing application support sub-layer (APS) and ZigBee device object (ZDO).

Background and Objectives: MERL has been at the forefront of ZigBee related activities including standardization of its specifications, development of network stack, and building up prototype systems. Renesas and MERL collaborated to develop complete ZigBee stack on M16C based target board. The core ZigBee technology developed at MERL will be used to produce value added application systems.

Technical Discussion: ZigBee defines only network and higher layers where as it uses MAC and PHY for IEEE 802.15.4 wireless standard that offers a short range peer to peer wireless connectivity. It allows the use of 804 GHz and 900 MHz frequency bands. ZigBee is a best-effort-delivery protocol. The mechanisms for reliable delivery and packet fragmentation/assembly, if needed, are left to the application layer.

MERL has developed network layer, application support sub-layer (APS), and ZigBee device object (ZDO) while Renesas implemented MAC and PHY. Currently, our code executes on two different target boards from Renesas having a Freescale RF chip and a Chipcon RF chip, respectively. The code has also been ported to M16C-based board produced at Johosoken. M16C MCU is capable of supporting both ZigBee stack and application code simultaneously.

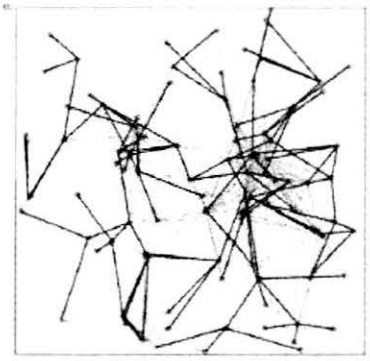
Collaboration: MERL has developed ZigBee stack in close collaboration with Renesas. We have been working closely with Johosoken, Sentansoken, and Jyukanken on ZigBee based systems development. We expect this cooperation to grow in coming months and years.

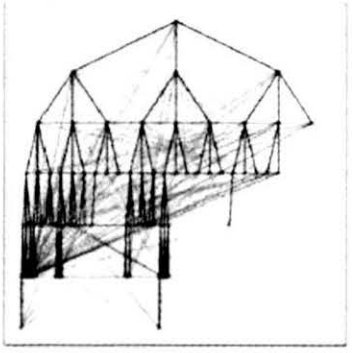
Future Direction: MERL will continue developing other components of ZigBee devices such as security toolkit, device location estimation, and location based application systems. We will also continue implementing new features being included in next version of ZigBee specification.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

Data Broadcasting in ZigBee





ZigBee is an emerging standard for low rate, low power, and short range communication networks. The underlying MAC/PHY technology is IEEE 802.15.4. ZigBee is expected to drive a diverse set of sensor network applications from home automation to environmental and structural monitoring, and from asset management to automated meter reading etc.

Background and Objectives: In this project, our objective is to develop an efficient data broadcasting protocol for Zigbee networks, and contribute it to the ZigBee Alliance for standardization. In ZigBee networks, it is very often a necessity to send a broadcast message to all devices in the network at once. A need to request a report from all the devices, or to command them to sleep for a while are some examples. In a very large and dense network rebroadcasting of a broadcast message at intermediate nodes causes extra traffic in the air. Furthermore, redundant traffic drains devices batteries more quickly and increases end to end latency in message delivery due to collisions and retries. Therefore, an efficient broadcasting protocol is needed to prune out redundant retransmissions. Currently ZigBee uses a best effort flooding based broadcasting mechanism in which every node forwards each broadcast message they receive. Our objective is to lower the number of nodes that forward a broadcast message, while still ensuring that all the devices receive the message, that is the network is fully covered.

Technical Discussion: In ZigBee networks, during network formation devices are assigned 16-bit logical addresses by the coordinator device. The addresses are distributed hierarchically such that it is possible for every device to know the address of its parent and the address space of the child nodes of its parent, and parent neighbors. In the broadcasting protocol we are developing, a device exploits the hierarchical addresses of its neighbors and their link qualities, and makes a decision to whether to forward a broadcast message or not.

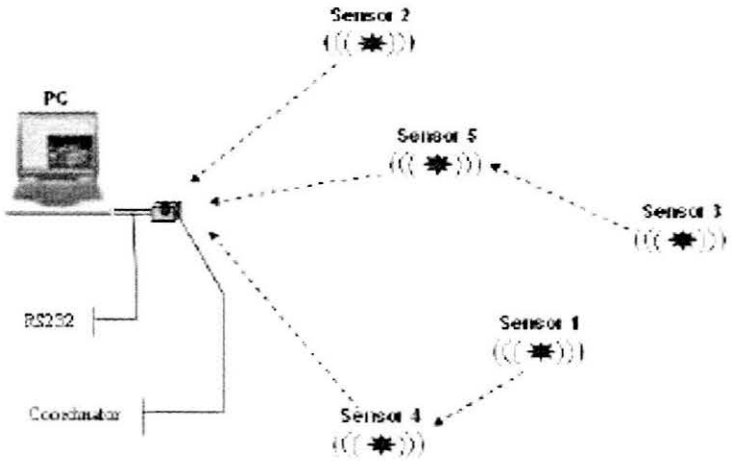
Collaboration: Renesas, Johosoken

Future Direction: Further enhancement and modification of the developed broadcasting protocol according to the changes in the new ZigBee specifications, development of a multicasting protocol for ZigBee networks.

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

Lab: MERL Technology Lab
Project Type: Research

Image Over ZigBee Networks



We have built an initial prototype for testing transmission of images over ZigBee networks that supports multi-hopping. 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.

Background and Objectives: Our objective is to develop a prototype system for JPEG2000 image transmission over ZigBee network. 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.

Technical Discussion: The prototype is implemented on a single M16C microprocessor. It includes ZigBee networking (NWK) and IEEE 802.15.4 MAC layer protocols and application. Fragmentation and reassembly are performed at the application layer. Source images prestored on the end devices and transmitted to the destination per request. Images encoded by using JPEG2000 compression. JPEG2000 has a higher source coding complexity then JPEG, but its compression efficiency is better and it provides resolution and quality scalable coding of a bitstream. One ZigBee device is connected to a general purpose PC via an RS-232 connection and acts as a gateway into the wireless portions of the network. The gateway allows the collection of various data from the network such as network topology, neighbor tables and route information. A simple application runs on the PC and issues requests for an image transfer from one of the devices in the network. The route that the packets traverse is also available to the gateway/PC data sink.

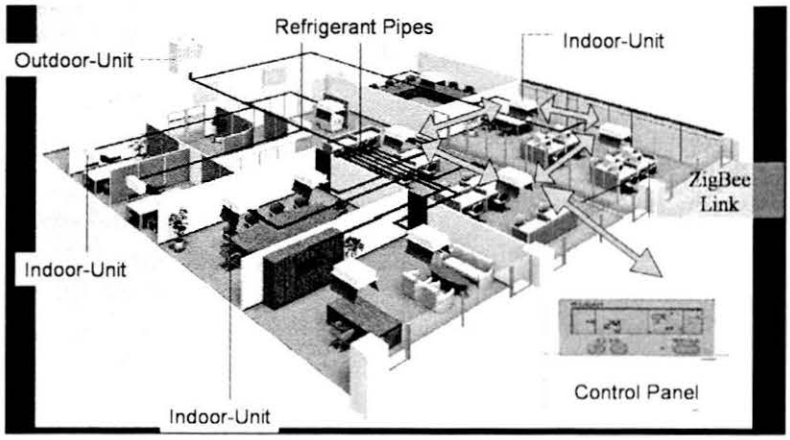
Collaboration: Sentansoken

Future Direction: Perform multihop testing over ZigBee for JPEG2000 images; integrate real time JPEG2000 encoding hardware into ZigBee platform; explore adaptive source coding based on link quality; investigate various flow control mechanisms.

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

Lab: MERL Technology Lab
Project Type: Initial Investigation

ZigBee HVAC Application Profile Development



We have been active in the ZigBee Applications Framework Group to specific profiles that will properly support Mitsubishi Electric's ability to leverage ZigBee wireless personal area mesh networking technologies. We have been assisting MELCO in driving the ZigBee Alliance towards producing and HVAC specific application profile. MERL currently holds the chair position for the HVAC Profile Task Group (HVAC-PTG) and is assisting MELCO in writing the profile specification.

Background and Objectives: MELCO does significant business in the commercial and residential split-ductless HVAC markets. It is desirable to enable the wireless monitoring and control of such systems in order to reduce installation costs as well and improve energy efficiency.

Technical Discussion: MERL's main contribution has been in our efforts to drive the ZigBee Alliance to develop the HVAC application profile. Towards this end MERL has taken a leadership role within the Alliance and is currently chairing a profile task group that is defining HVAC device descriptions and application packet formats specific to HVAC systems. Additionally, the work being done within the HVAC-PTG is being coordinated with other application profile development, specifically, the groups working on Industrial Plant Monitoring (IPM) and Commercial Building Automation (CBA). This effort focuses on the enabling multiple application profiles to run simultaneously over a single ZigBee network and reuse network resources, specifically router nodes, to carry all application traffic.

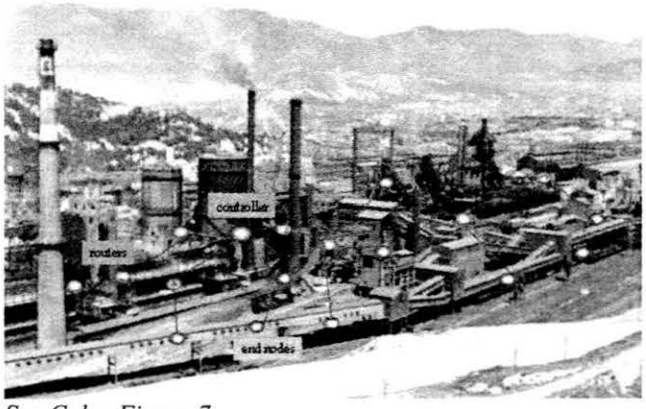
Collaboration: We are working with Kushiro-san, Nakata-san, and Yamamoto-san in the Living Environment Systems Laboratory.

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

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

Lab: MERL Technology Lab
Project Type: Advanced Development

ZigBee for Industrial Plant Monitoring



See Color Figure 7

ZigBee is an emerging standard for low rate, low power, and short range communication networks. The underlying MAC/PHY technology is IEEE 802.15.4. ZigBee is expected to drive diverse set of sensor network applications from home automation to environmental and structural monitoring, and from asset management to automated meter reading etc. In this project, our objectives have been to develop an industrial plant monitoring (IPM) application profile for standardization in ZigBee and also to prototype a monitoring system based on the ZigBee technology. The IPM application profile specification is being drafted in the IPM Profile Task Group, which is being chaired by MERL, and it is expected to be standardized in fall 2005. The IPM profile consists of a set of various sensor device descriptions, and it also specifies a fragmentation and reassembly protocol to support delivery of bulk array data. In parallel with the standardization efforts, we have been developing an asset management platform for a use in a waste water plant monitoring application by Shaden. We have designed and produced sensor boards for ZigBee devices to monitor physical metrics such as motor vibration frequencies, environment and equipment temperature and etc.

Background and Objectives: The standardization of the IPM application profile to provide interoperability among different vendors is important. The early demonstration of the prototype IPM system will facilitate the actual deployment of the monitoring network requested by MELCO external customers. The monitoring system will be a useful tool for maintenance and process control, and for assisting in making timely repair/rebuild and retrofit decisions and optimizations for future operations by means of early faults detection.

Technical Discussion: The prototype devices consist of IEEE 802.15.4 MAC and PHY, ZigBee networking protocol and application support sub-layer (APS). The application devices are programmed on M16C chips.

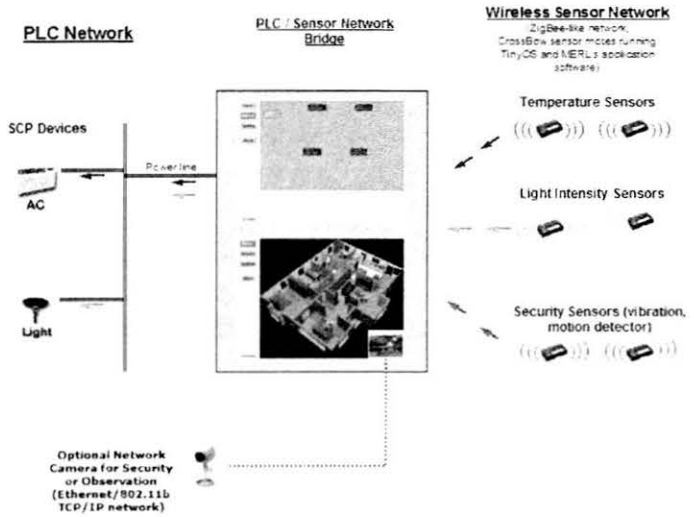
Collaboration: MEUS, Shaden, Sentansoken.

Future Direction: Standardization and testing of the IPM prototype will be completed in December 2005. The draft specification will be submitted to the ZigBee Architecture and marketing group for approval. We would like to extend this project to structural and environmental monitoring applications (e.g., cherry fields). Location awareness will be added to the system features in late 2005. We have started testing the efficiency of link quality based range estimation using IEEE 802.15.4 radios.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

ZigBee/SCP for Home Automation



MERL has developed bridging technology for Simple Control Protocol (SCP), a power-line communications network, and ZigBee, a wireless sensors network. Both of these protocols offer a low cost and low bit-rate connectivity solution primarily to provide the controlling functions in a home environment.

Background and Objectives: Future homes will utilize advanced wired and wireless networks for connectivity. Our objective is to demonstrate some of the underlying communication technologies that form the basis of such networks. Recent developments in power line data transmission and low bit rate radio enable a whole new range of devices to effectively communicate. Typical household devices (e.g., lights, doorbells, appliances, entertainment, security systems) are now on the verge of being seamlessly integrated into home networks by using both wired and wireless technologies. Essentially, the home network will consist of many low cost networked devices in addition to PC, broadband internet connection and access point. This will allow homeowners a complete home network with the capability to control all aspects of their environment as well as introduce new applications for the home.

Technical Discussion: Both the SCP and ZigBee technologies are non-IP based networking protocols. ZigBee is especially suitable for sensors and/or control networks. The SCP, on the other hand, utilizes existing infrastructure. Bridging the two domains will offer the best of both technologies. SCP/ZigBee bridge is implemented on the PC and has the functionality of integrated home control center for both networks. It routes ZigBee sensor data to appropriate SCP device. It allows configuration and control of various network-specific parameters such as routes, parent-child relationships for ZigBee and property routes for SCP. It includes ability to map both networks into UPnP space.

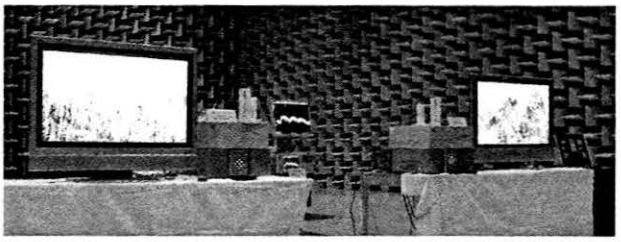
Collaboration: We are working closely with RSO, RT, and Johosoken.

Future Direction: SCP/ZigBee can be implemented on M16C hardware to minimize PC usage.

Contact: Jinyun Zhang, Georgiy Pekhteryev
http://www.merl.com/projects/scpzigbeebridge/

Lab: MERL Technology Lab
Project Type: Advanced Development

HDTV Transmission over UWB



The Ultra-wide-band technology allows high data rate wireless communication over short range and has generated a lot of interests in the past few years. MERL is a founding member and contributor of MB-OFDM Alliance (MBOA) has several contributions to the MBOA standards. One of the foreseeable application areas for UWB technology is wireless AV data stream transmission. The large bandwidth (20Mbps/streams) needed for the HDTV stream makes it difficult to send wirelessly over other narrower band link. The UWB link provides enough bandwidth to support multiple stream transmission.

Background and Objectives: The foci of this project are 1) to implement a complete MB-OFDM UWB physical layer and 2) to build the application (HDTV transmission) on top of the platform. Our target raw data rate is 110Mbps at UWB PHY layer to support simultaneous transmission of multiple HDTV video streams.

Technical Discussion: The UWB transceiver is implemented using FPGAs, high speed A/D, D/A and discrete RF components. Many DSP functional blocks (FFT, channel estimation, interleaver etc) are specially designed to handle the 528MHz sample rate. The transceiver interfaces the PC through PCI bus. The overall baseband sub-system is implemented in multiple FPGAs with total gate count over 18 million. We achieve better than 10e-6 BER at 110Mbps over 4 meters range. The receiver dynamic range is >30 dB. Application software was developed to demonstrate simultaneous transmission of multiple High Definition video streams over UWB. The user-case for the Demo is a Multimedia Client-Server scenario where we have one Media Server (transmitter) and several Media Players (receivers). Since the current implementation of the UWB connection is unidirectional, we use TCP/IP communications over 802.11a/b/g or 100Base-T PHY as a feedback channel for command and control. Multimedia content is prerecorded on the Server as MPEG2-TS HD files. Each Media Player is connected to an HDTV Monitor via 1394 interface. It uses 1394 AV/C protocol for command and control and ISO 61883-4 protocol for the MPEG2-TS data transmission. An HDTV's built-in NetCommand feature allows the display of a list of available content and the selection of certain items for viewing. In addition, simple commands like "Play" and "Stop" are implemented.

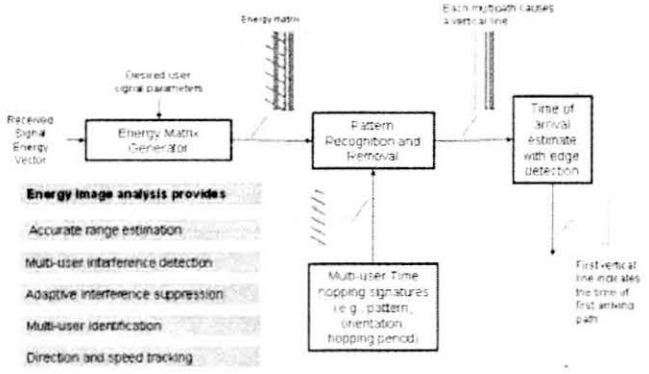
Collaboration: We worked with Johosoken on RF hardware design and some of the baseband DSP designs.

Future Direction: To improve the overall value of the current prototype, we are working on three fronts in parallel: performance improvement, specification compliance and feature enhancement. We are also working on LSI implementation issues.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

High Precision UWB Ranging and Locating



In the new decade, the demand for location aware applications in both cellular and short range communication networks will continuously keep increasing. As the required positioning accuracy gets higher, the precise tracking of the first arriving signal component to the accuracy level of several centimeters is becoming important. Especially in UWB based networks with highly dispersive channels, the time of arrival (TOA) estimation is extremely challenging, where the leading-edge path is not necessarily the strongest path.

In this research, we study and design novel low cost coherent and non-coherent receiver architectures and time of arrival tracking methods for UWB systems that will enable high precision ranging. The advanced techniques are being contributed to the emerging IEEE 802.15.4a standards.

Background and Objectives: UWB radios have bandwidths more than 500MHz, which offers good object penetration capability and a high time resolution for improved ranging accuracy. Therefore, the UWB technology will soon bring the location awareness capability to short range cooperative communication networks; and the important task will become designing a simple and yet efficient TOA tracking algorithms and receiver architectures suitable for resource limited devices. Our objective is to investigate and design novel precision ranging techniques for low cost non-coherent receivers and high performance coherent receivers.

Technical Discussion: Besides the challenge that the line of sight signal may not be the strongest signal, interference from simultaneously operating users can be deleterious to the TOA detection performance. MERL has been the pioneer to introduce the concept of converting collected received signal energy samples into a 2D energy image at the receivers, and then to apply image edge detection filters to remove interference and to more accurately identify the leading signal energy. The energy image analysis based ranging estimation is under consideration by IEEE 802.15.4a TG as part of the non-coherent receiver architecture.

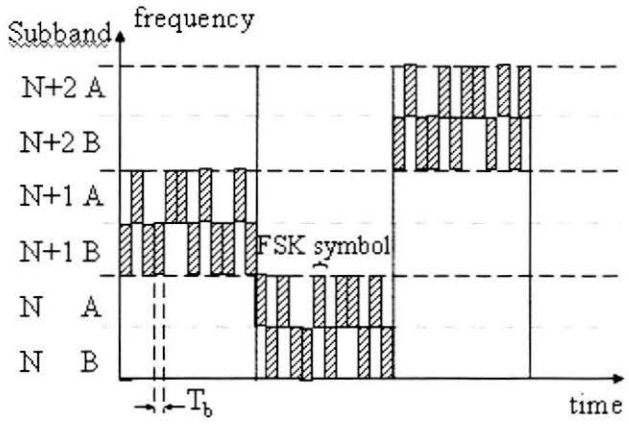
Collaboration: Johosoken

Future Direction: Adopting the energy image based ranging estimation to coherent receiver architectures, high precision ranging with mobile devices, identifying interfering sources and tracking their range, prototyping of the invented ranging technologies, and developing location aware sensor network platforms.

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

Lab: MERL Technology Lab
Project Type: Research

Low Complexity UWB Transceiver with Multiband OFDM



We propose a low-complexity transceiver for ultrawideband communications with moderate (1-15Mbit/s) data rate. This transceiver is based on time-frequency-interleaved frequency-shift-keying (TF-FSK), and shows a high degree of compatibility with multiband-OFDM, the currently envisioned standard for high-data-rate (>100Mbit/s) UWB communications. For dual-mode devices (able to communicate with high data rate UWB devices AND moderate data rate UWB devices), the major part of the multiband-OFDM transceiver can be reused for the TF-FSK transceiver, assuring a high degree of compatibility at almost no extra cost. Depending on the data rate, coverage ranges of up to 30m with LOS connection are possible.

Background and Objectives: The IEEE has established a standardization body, IEEE 802.15.3a, for defining a physical-layer standard based on UWB transmission. The goal of this standard is achieving a data rate of 110Mbit/s at 10m distance, and higher data rates at shorter distances. The only baseline proposal currently (May 2004) still under consideration is the multiband-OFDM proposal. This proposal uses a combination of time-frequency (TF) interleaving and OFDM.

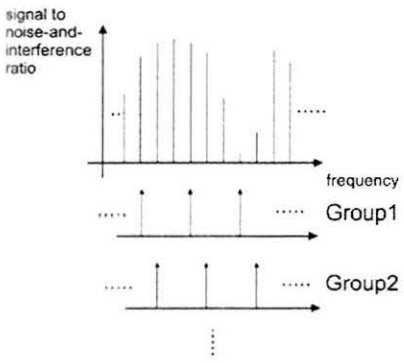
Technical Discussion: This low complexity transceiver is based on time-frequency interleaved FSK Using incoherent demodulation. A simple pulse-based scheme is used. Each symbol is represented by transmitting energy in either the lower or the upper half of a 528 MHz band - in other words, we use very wideband FSK on top of the time-frequency interleaving. In order to keep the signal structure as similar to OFDM as possible, multiple contiguous symbols are transmitted within one 528 MHz band. The duration of the transmit symbol is taken shorter than the bit duration to reduce the amount of intersymbol interference, while retaining the total transmit power. Coverage distances of some 30m can be achieved for 3Mbit/s data rates, based on evaluations with the IEEE channel model.

Future Direction: The performances can be improved by adjusting several parameters. Their optimal values and the performances need to be confirmed with a complete theoretical analysis. Concurrently we should smooth the spectrum of the FSK signal and equally distribute the energy amongst bits to close the gap with an implementable system. We continue our efforts to propose this low data rate scheme to the MultiBand OFDM Alliance.

Contact: Yves-Paul Nakache, Andreas F. Molisch
http://www.merl.com/projects/lowcomplexityuwb/

Lab: MERL Technology Lab
Project Type: Research

Symbol Spreading in MB-OFDM UWB



The IEEE 802.15.3a group is defining an international standard to enable reliable, cost-effective, high data rate communication for Wireless Personal Area Networks (WPANs) using UltraWide Bandwidth (UWB) signals. The targeted data rates are from 110Mbps to 480Mbps. MERL has made key contributions to the development of the PHY specifications of the MultiBand OFDM Alliance (MBOA) Special Interest Group, one of the two groups involved in the activity of 802.15.3a. The MBOA ultimately launched its specifications in collaboration with the WiMedia alliance. Some of the targeted applications include wireless home video, projectors, external cards and dongles for low-cost Consumer Electronics and PC devices.

In April 2005, MERL's contribution "Symbol Spreading Technique" for high data rates was included in the final draft of the MBOA standard for UWB wireless communication. By using multiple OFDM subcarriers to encode each symbol, MERL's technology exploits the frequency diversity inherent in a UWB channel thereby reducing transmission errors caused by interference and fading.

Background and Objectives: The essence of the MBOA/WiMedia proposal is to divide the available spectrum into several subbands of 500MHz. Within the current subband, Orthogonal Frequency Division Multiplexing (OFDM) is used to communicate 100 streams of data simultaneously using 100 different subfrequencies (tones). The frequencies are close together and interfere with each other, but the modulation used for the individual data streams are chosen to be mathematically 'orthogonal' so the data streams can nevertheless be reliably decoded by a receiver. A problem with the above approach is that each individual OFDM tone is subject to fading and interference from other sources. At high data rates the coding protection was not sufficient to fight bad channel conditions of dense multipath environments. Tones are attenuated by the channel fading and the information is lost.

Technical Discussion: The essence of MERL's proposal is to mix the data streams together so that multiple tones are used for each data stream (each tone combines information from several data streams). This allows reliable communication even if communication on a few tones is blocked.

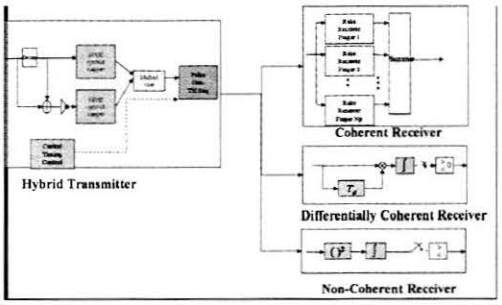
Collaboration: MERL proposed its original scheme to the MBOA to quickly work in collaboration with other companies eager to improve this method. I2R, Broadcom and Infineon worked with us and contributed to the design of a more efficient spreading code.

Future Direction: Follow the standardization process led by WiMedia to guarantee the inclusion of the symbol spreading technique in the future PHY specification versions.

Contact: Yves-Paul Nakache
http://www.merl.com/projects/symbolspreading/

Lab: MERL Technology Lab
Project Type: Advanced Development

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: In 2002 the American frequency regulator (FCC) has allowed unlicensed ultrawideband (UWB) communications. Since that time, there is growing interest in UWB, both for high-data-rate communications as well as for sensor networks with geolocation capabilities. Additionally, a new standard group, IEEE 802.15.4a, is developing a standard for low data rate radio with location capability based on UWB impulse radio. Our focus has thus been the development of modulation techniques for low data rate transceivers and contributing our ideas to the IEEE 802.15.4a standards body.

Technical Discussion: In the realm of embedding and/or sensor network the additional cost of adding wireless connectivity to existing devices is an extremely important issue. In order for large scale market adoption wireless solutions must offer performance at exceptionally low cost. While impulse radio systems show good potential, they have several problems: the collection of energy from many multipath components, increased acquisition times, and difficulties in shaping the spectrum to fill the FCC mask as closely as possible. In order to solve these problems coherent receivers tend to be costly and complex (high sampling rate, many rake finger for energy collection, and advanced signal processing is required). One approach is to sacrifice performance and use simple receiver architectures based on non coherent energy collection or differential transmitted reference receivers. However, loss in performance may limit the applicability of such receivers. We have investigated and developed a coded modulation technique that we call Hybrid Impulse Radio (H-IR) that combines aspects of BPSK, transmitted reference, and Pulse Position Modulation (PPM). The key feature of H-IR modulation is that is allows reception at variety of receivers types ranging from simple energy detectors to complex coherent rake receivers. The H-IR technique is currently under consideration for use by IEEE 802.15.4a standards body.

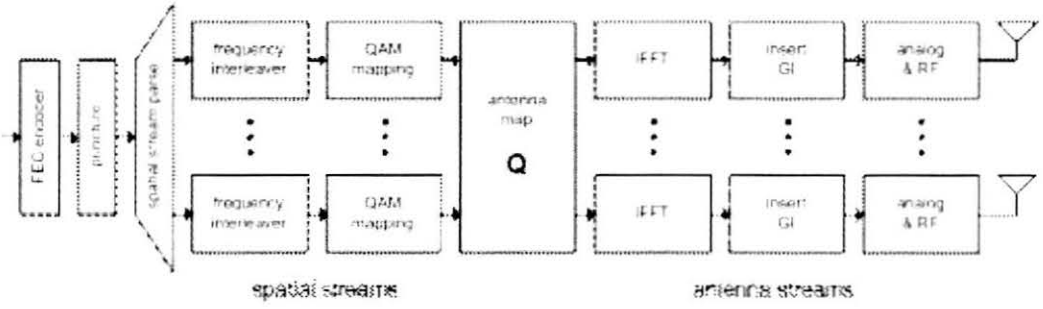
Collaboration: Johosoken

Future Direction: We will investigate further enhancements to the H-IR scheme such as the inclusion of higher order modulation for greater bit rates and advanced coding and decoding techniques for improved performance

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

Lab: MERL Technology Lab
Project Type: Research

MIMO-OFDM System for High Speed WLAN



Wireless LANs have been a major success for the wireless industry, now being an indispensable part of laptop computers as well as many personal digital assistants and similar devices. They allow much higher data rates than cellphones and, being restricted to so-called "hot spots", access points are easier and cheaper to set up. The 802.11 family of standards (WiFi) have greatly contributed to this proliferation of wireless LANs. Within this family, a new standard (802.11n) is being developed for data rates in excess of 100 Mbit/s. In order to achieve this high data rate in the small available spectrum, so-called MIMO systems, which have multiple antennas at both link ends, are being used.

Background and Objectives: Our objective is to contribute to the development of high-throughput wireless LANs, particulary the IEEE 802.11n standard. The main goal is to find transmission schemes and implementations that improve the performance (compared to "standard" MIMO techniques) without significantly increasing the costs of the system.

Technical Discussion: In MIMO-OFDM, data are transmitted on different frequencies (subcarriers), as well as from different antennas. One important question is how the data can be distributed most efficiently over the subcarriers and antennas, and how they should be encoded to protect against transmission errors. One of the results for our research was a scheme for joint encoding and space/time "layering", i.e., distributing the data to the different antennas. This scheme uses the highly efficiency LDPC (low density parity check) codes, and introduces correlation between the signals from different antennas that allows a more efficient iterative decoding. Other contributions include an adaptation of MERL's antenna selection scheme to the framework of 802.11n, and contributions to the TGnSync industry alliance that is developing a joint proposal for the 802.11n standard.

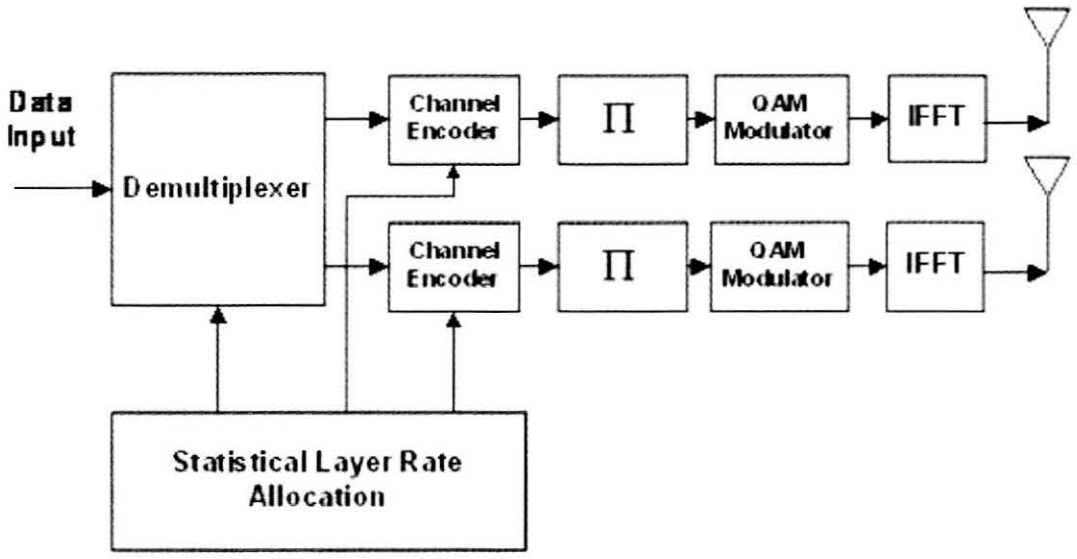
Collaboration: Johosoken.

Future Direction: We will continue our activities within the 802.11n standardization, and develop a MATLAB implementation of the draft standard as a basis for further improvements and additional investigations.

Contact: Andreas F. Molisch
http://www.merl.com/projects/mimo-ofdm/

Lab: MERL Technology Lab
Project Type: Advanced Development

Statistical Data Rate Allocation for Layered Space-time Systems



MIMO systems are wireless systems with multiple antennas at transmitter and receiver. They allow much higher spectral efficiencies and data rates than conventional, single-antenna systems, by transmitting different data streams from the different transmit antennas. One key problem is what data rates should be assigned to the different transmit antennas. We develop a scheme that needs to know only average channel state information.

Background and Objectives: The goal of our investigation is to find a layered transmission scheme for MIMO that adapts its rate to the channel state while requiring little feedback from the receiver to the transmitter.

Technical Discussion: Our scheme employs a modified layered structure for multiple-input multiple-output (MIMO) systems, where the layer detection order is fixed and the data rate for each layer is allocated based on the detection order and channel statistics. Using a Gaussian approximation of the layer capacities, we derive the optimum data rate allocation. The amount of backoff from the mean layer capacity is proportional to the standard deviation of the layer capacity. This scheme considerably improves the throughput compared to the case without adaptive rate allocation. At the same time, it requires only average channel state information, which simplifies transceiver structures and decreases feedback requirements compared to other rate-adaptive schemes.

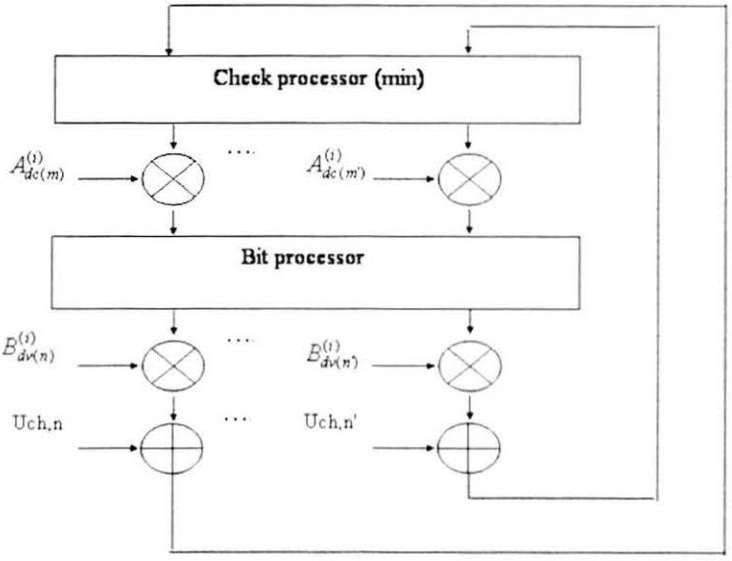
Collaboration: Johosoken.

Future Direction: The concept has been developed, and been borne out by simulations. Possible implementations in proprietary wireless LAN systems are considered.

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

Lab: MERL Technology Lab
Project Type: Research

2D Normalized Min-Sum LDPC Decoding



The proposed idea here is to use two-dimensional (2-D) normalization Min-Sum algorithm to improve the performance of decoding of irregular low density parity check (LDPC) codes, in which messages delivered by check node and then messages delivered by bit node in Min-Sum decoding are both normalized. Simulation results and theoretical analysis show that the proposed decoding has a better performance compared with Min-Sum and the conventional normalization Min-Sum decoding methods.

Background and Objectives: Recently, LDPC codes have been intensively studied due to its superior error correction performance. It has been accepted or considered by many industry standards such as DVB, IEEE802.16 and IEEE802.11n. One of the concerns related to LDPC is its decoding complexity and performance. So, we developed a 2-D normalization Min-Sum decoding algorithm to improve the performance, and reduce the complexity at the same time.

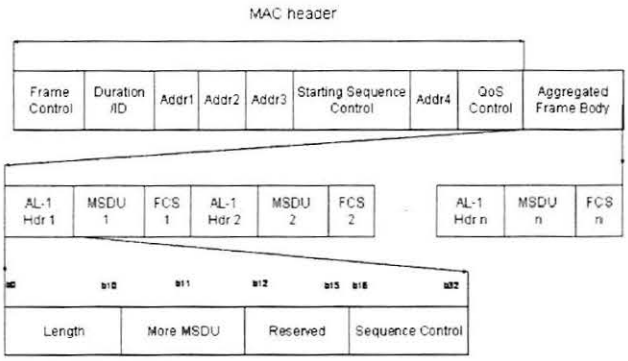
Technical Discussion: In 2-D normalized Min-Sum scheme, belief messages outgoing from both check and bit node processors are normalized. The normalization factor of a check (bit) node processor depends on the degree of that node. To circumvent brute force search, parallel differential optimization and density evolution are used to obtain the 2-D optimal normalization factor pair. Theoretical analysis and simulations with the 2-D normalized Min-Sum decoder show that the proposed decoding scheme provides better performance, less complexity and decoding speed trade-offs than conventional BP or Min-Sum decoders.

Future Direction: The project has been completed.

Contact: Daqing Gu, Jinyun Zhang
http://www.merl.com/projects/min-sum/

Lab: MERL Technology Lab
Project Type: Research

Frame Aggregation for High Efficient MAC



A frame aggregation scheme is proposed for high efficient MAC. The proposed scheme performs frame aggregation at both MAC service data unit (MSDU) level and MAC Protocol Data Unit (MPDU) in a highly efficient manner and still provides tremendous flexibility in the meantime. The frame aggregation scheme achieves a substantial efficiency improvement, without causing appreciable additional complexity. Moreover, the frame aggregation scheme is used in conjunction with a simple management procedure so that backward compatibility with legacy MAC standards is guaranteed.

Background and Objectives: IEEE802.11 standard working group is currently developing a high throughput WLAN standard called 802.11n, in which a throughput of 100 Mbps is achieved at MAC SAP (session access point). This new standard requires both efficient MAC and high-speed PHY. Therefore, the legacy MAC protocol must be enhanced to improve its efficiency. One of the problems with the legacy MAC is the overhead associated with MAC header. One of the efficient ways to reduce the overhead caused by MAC header is to use larger packets. So, aggregating a number of frames together for transmission is becoming popular for high efficient MAC.

Technical Discussion: We propose a highly efficient and flexible frame aggregation scheme, which can substantially improve the channel efficiency. The proposed frame aggregation can occur at two levels, namely the MSDU level and MPDU level. At the MSDU level, frames with identical destination address are aggregated into a single MSDU with proper auxiliary fields inserted into the resulted aggregation. In contrast, MPDU frames with different destination addresses can be transmitted in an aggregated manner at the MPDU level, sharing a single PLCP preamble. Thus, excessive overheads at both MSDU level (e.g., MAC header for each MSDU) and MPDU level (e.g., PLCP header) can be eliminated, and a significant increase in efficiency can be achieved. Besides, the aggregation at each level can be used together or separately. Hence, this two level aggregation can operate flexibly, based upon the applications and real traffic scenarios.

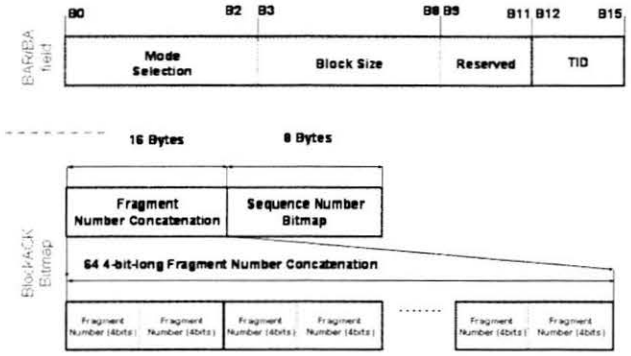
Collaboration: This project is in collaboration with Johosoken.

Future Direction: In future, we will continue to develop high efficient MAC protocols, and contribute to IEEE802.11 standard body. We are also going to simulate and implement frame aggregation protocols.

Contact: Daqing Gu, Jinyun Zhang
http://www.merl.com/projects/frame-agg/

Lab: MERL Technology Lab
Project Type: Research

Efficient Block-ACK Mechanism



The upcoming IEEE 802.11n standard requires that a throughput of 100 Mbps must be achieved at MAC-SAP. Various mechanisms in the legacy IEEE 802.11/IEEE 802.11e MAC protocols entail immense protocol overhead and eventually result in serious throughput performance degradation. Therefore, direct application of the legacy MAC protocol in the IEEE 802.11n standard is not plausible, if the protocol efficiency is not increased significantly. We hereby propose Efficient Block Acknowledgement (BlockACK) mechanism, which performs BlockACK in a highly efficient manner and still provides tremendous flexibility in the meantime. To fulfill the same task, the proposed efficient BlockACK consumes far less bandwidth than the legacy BlockACK mechanism in IEEE 802.11e.

Background and Objectives: The per frame ACK mechanism used in the legacy IEEE 802.11 presents a huge waste of the wireless resource. This is why BlockACK is proposed in IEEE 802.11e. But unfortunately, the BlockACK mechanism itself is not entirely optimized with respect to the bandwidth efficiency. The BlockACK bitmap field in the BlockACK frame has a fixed length of 128bytes. Thus, even when a few frames are transmitted and require BlockACK, a 128-byte-long BlockACK bitmap has to be included in the BlockACK anyway. This causes significant waste and inflexibility. Lots of effort has been spent to develop a new and efficient BolckAck mechanism.

Technical Discussion: We propose an efficient BlockACK mechanism, in which the BlockACK bitmap is a variable size field. It can either signal the packets that have been received correctly or erroneously, whichever takes lesser bytes to transmit. The rate adaptation will optimize the transmission rate to achieve a reasonably low packet error rate (PER). Thus, the probability of the persistent occurrence of errors is expected to be low, which implies that most of the packets should be correctly received and the corresponding BlockACK message only need to indicate which frames have not been correctly received. It is easy to show that at any time the proposed scheme only need 25% of the legacy BlockACK bitmap field to convey the same amount of information.

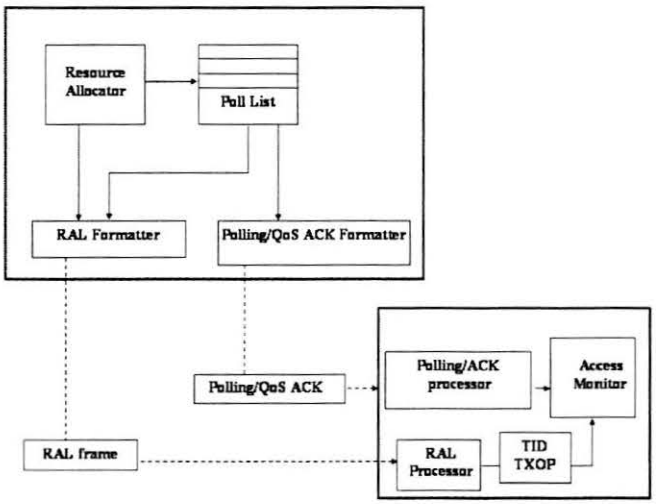
Collaboration: This project is collaborated with Johosoken.

Future Direction: We will continue to work on high efficient MAC proposal for high throughput WLAN in future. Implementation and simulation for BlockACK will be conducted, too.

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

Lab: MERL Technology Lab
Project Type: Research

HCCA Enhancements for High Speed WLAN



An enhancement and simplification of the legacy 802.11e HCCA (HCF controlled channel access) has been proposed. The proposed approach employs a multipolling mechanism to reduce polling and handshaking overhead by disseminating polling information at the beginning of the contention free period (CFP). Moreover, we propose a novel deferring mechanism under which special frames are counted instead of time slot or time unit used in current multipolling systems. The mechanism solves the hidden terminal problem and at the same time maintain the high channel utilization of multi-polling system. The proposed mechanism has good backward compatibility with legacy 802.11e protocol as well as low implementation complexity.

Background and Objectives: In HCCA mechanism, each time the AP only polls one STA, there is no hidden terminal problem in the network since only that STA can transmit when polled. Moreover, since AP monitors the STA's activity on a per poll basis. If the polled STA does not response to AP's poll, AP will immediately poll next STA in polling list after detecting the timeout of STA's transmission therefore the waste of channel resource is negligible. On the other hand, although in IEEE 802.11e the polling message can be piggybacked within acknowledgment and data, the polling overhead is still high. Moreover, in 802.11e, since normally QoS Poll will not be piggybacked with QoS ACK unless they are addressed to the same destination.

Technical Discussion: The proposed solution works for data transmission during contention free period (CFP), more specifically, the polling information for all the STAs are disseminated first at the beginning of each CFP. The PC/HC is responsible to poll the next STA in the polling list by the end of the transmission of current STA. Each STA retrieves its transmission sequence defined by the number of simplified polling message /piggybacked with QoS CF-ACK it should observe before starting transmission and assigned TXOP through the multipolling message at the beginning of CFP. The STA will access the channel only when it has observed the number of simplified polling messages it retrieved.

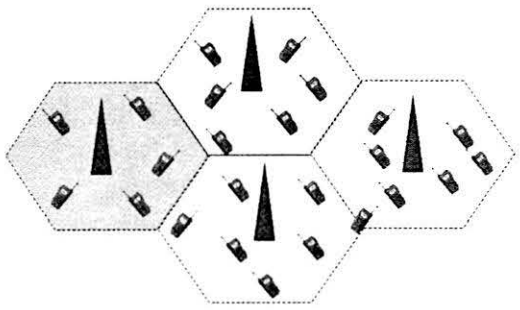
Collaboration: This project is in collaboration with Johosoken.

Future Direction: This is the project for IEEE802.11n high throughput WLAN. We will continue to work with Johosoken to develop a high efficient MAC for IEEE802.11n standard.

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

Lab: MERL Technology Lab
Project Type: Research

MIMO Multiple Cell System-level Simulator



We are developing a system-level simulator to measure the performance gains of various MIMO schemes in a multi-cell and multi-user setting. The simulator can handle several MIMO transmission schemes and receivers proposed for beyond 3G cellular systems. And, it fully supports the advanced spatial channel model proposed in the 3GPP (third generation partnership project) standardization body.

Background and Objectives: MIMO technology is expected to result in three improvements in 3GPP: higher peak rates, greater coverage (cell area), and greater average spectral efficiency. While a large amount of research has focused on the performance of MIMO systems on the link-level, i.e., between one transmitter and one receiver, or for the case with channel-aware schedulers, one transmitter and many receivers, the gains of MIMO are less well understood at the cellular system-level, which consists of several cells and mobile stations that can interfere with each other. Our aim is to develop a system-level simulator for MIMO and evaluate its performance for different system parameter settings and various transmitter and receiver architectures.

Technical Discussion: The system simulator being developed can handle several MIMO transmission schemes proposed in 3GPP such as per-antenna rate control, space-time transmit diversity (STTD), sub-group rate control with double STTD for 4 transmit antennas, etc. Various MIMO receiver algorithms with different performance vs. complexity trade-offs are also implemented. Channel-aware schedulers such as the proportional fair scheduler, round-robin scheduler, and the Maximum net rate scheduler another are also tailored for MIMO.

The system simulator is being developed to investigate MIMO performance over Wideband code division multiple access (WCDMA) physical access technology, which is used in third generation systems, as well as orthogonal frequency division duplex access technology, which is a candidate for next generation cellular systems.

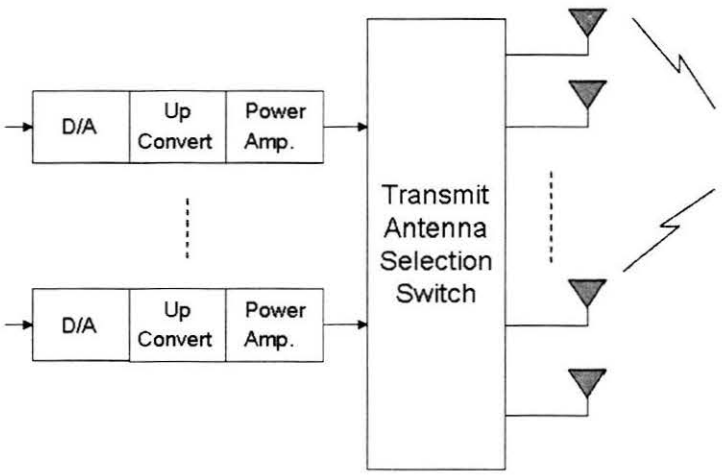
The MIMO system simulator fully interfaces with our implementation of the spatial channel model (SCM) specified by 3GPP. The SCM provides a single unified spatio-temporal model for several environments such as suburban macrocell, urban macrocell, and urban microcell. The SCM represents the frequency-selective MIMO channel as a superposition of clustered constituents, which are specified in terms of powers, angles of arrival and departure, and time of arrival, which are specified stochastically and are correlated with each other. The SCM provides a good balance between realistic, experimentally verified, spatial and temporal channel models and modeling complexity.

Future Direction: We plan to incorporate several MIMO-OFDM techniques into the simulator and evaluate their performance. Investigating the impact of channel estimation techniques on link and system capacity is another area of interest.

Contact: Neelesh Mehta
http://www.merl.com/projects/multi-cell/

Lab: MERL Technology Lab
Project Type: Advanced Development

Antenna Selection Signaling & Verification for Super 3G



We investigate antenna selection techniques at the transmitter to reduce MIMO complexity. In systems with error-prone feedback channels, the design of feedback signaling to enable transmit antenna selection and the receiver algorithms to cope with errors becomes critical.

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 means of a selection switch, which routes the signals to the best subset of available antennas.

In frequency division duplex systems, the receiver estimates the instantaneous channel state, and the transmitter relies on feedback from the receiver to adapt to the channel variations. Under ideal conditions, such as instantaneous error-free feedback, antenna selection has been shown to achieve the highest possible value of diversity order for space time trellis codes, orthogonal space time block codes, etc. Diversity order is a measure of the ability of the transmission scheme to cope with adverse fading. However, the performance of antenna selection under non-ideal conditions such as feedback errors, delay, and sub-optimal selection due to incomplete channel knowledge has received relatively little attention.

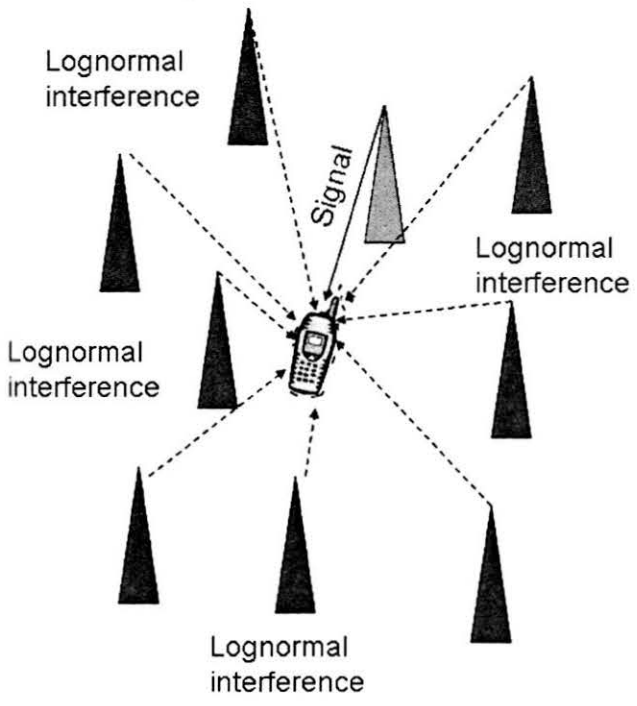
Technical Discussion: Errors rates in the feedback channel can be as high as 5%. Subject to the feedback channel's bandwidth and latency constraints, we investigate how to optimally design the feedback mechanisms so as to minimize the average symbol error rate of spatial diversity systems with transmit antenna selection and the average mutual information rate of spatial multiplexing systems with transmit antenna selection. The problem of verification, in which the receiver decodes the transmitted data using knowledge of the feedback channel's properties, is also investigated.

Future Direction: Our aim is to develop a common framework to design and analyze the feedback mechanisms and the corresponding receivers for various MIMO schemes.

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

Lab: MERL Technology Lab
Project Type: Initial Investigation

Method for Approximating Lognormal Co-channel Interference



The sum of lognormal random variables (RV) arises in many problems in wireless communications, such as in modeling co-channel interference from neighbouring base stations in cellular systems. While several methods are available in the literature to analytically approximate the sum by a single lognormal RV, each has its own disadvantage and none is unarguably better than the rest. We are developing a novel and general method to find this approximation. Our method provides the parametric flexibility required to overcome the drawbacks of the classical methods.

Background and Objectives: In the analysis of wireless systems, one often encounters the sum of lognormal random variables (RV). For example, it characterizes the total co-channel interference power from interfering base stations in a cellular system. The lognormal distribution also arises in ultra wide band systems, optics, and reliability theory. While exact closed-form expressions for the lognormal sum probability distribution function are unknown, several analytical approximation methods have been proposed in the literature to approximate it by a single lognormal RV. However, all the methods have their own advantages and disadvantages; none is unquestionably better than the others.

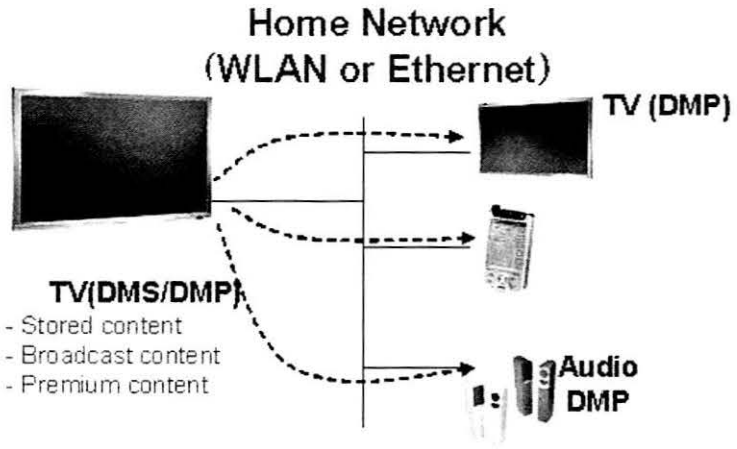
Technical Discussion: The Fenton-Wilkinson method, which equates the moments to determine the parameters, is inaccurate in the head portion (small values of the RV) of probability distribution. On the other hand, the Schwartz-Yeh method, which equates the moments of the logarithms of the RVs, is inaccurate in the tail portion of the probability distribution. We are developing a general method that uses a short Gauss-Hermite expansion of the lognormal moment generating function to determine the approximation parameters. The method offers considerable flexibility in accurately matching different regions of the probability distribution.

Future Direction: The method can be potentially be applied to problems in which the sum of random variables, with an arbitrary distribution, is approximated by a lognormal RV. We are investigating such problems. We are also trying to modify the approximation process so as to optimize any metric of interest.

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

Lab: MERL Technology Lab
Project Type: Research

DLNA Digital Media Server (DMS) Design



Consumers are using an increasing amount of digital media (photos, music, and video) on devices in the Consumer Electronics (CE), mobile, and Personal Computer (PC) domains. They want to conveniently enjoy the content across different devices and locations in the home. The DLNA developed a common set of industry design guidelines that provides the information needed to build interoperable networked platforms and devices for the digital home.

MERL has developed DLNA v 1.0 compliant Digital Media Server (DMS). It has been certified by UPnP IC and DLNA D2G. Also we contributed to the standard development.

Background and Objectives: Our objective is to contribute to the development of reusable software for common platforms of digital information appliances with audio-visual capabilities. We focus on set of technologies to control the transportation of audio-visual data and the operation of the appliances through home networks.

Technical Discussion: The basic function of a DLNA DMS is to implement the ContentDirectory and the ConnectionManager services as defined by the UPnP architecture. the ConnectionManager service helps DMP (Digital Media Player) determine what type of content is available on DMS. the ContentDirectory service helps DMP find content. In addition to UPnP services DLNA DMS provides media streaming transport via utilization "HTTP GET" method as specified in the DLNA Guidelines 1.0.

In order to make network experience seamless to the end user DMS as a DLNA device suppose to implement functionality recommended by IETF draft "Dynamic Configuration of IPv4 link-local addresses". Also known as "Auto IP" feature. Our target platform is Linux OS.

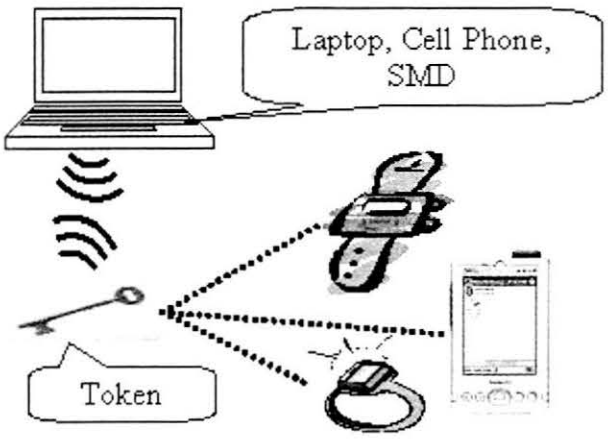
Collaboration: Johosoken

Future Direction: Implement some features of DLNA v1.5 specification, improve performance of DMS by implementing more efficient algorithms for parsing and streaming content.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

Wearable Token for Secure 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. In order to verify the efficiency of the proposed approach, we implemented a system of mobile devices consisting of a personal digital assistant (PDA), a laptop, and a USB flash memory. With the addition of our protocol, the prototype exhibits secure authentication with zero user interaction during the normal authentication phase.

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 be able to protect confidential information on a mobile memory stick as well as provide safe exchange of that data to authorized parties.

Technical Discussion: Data security for any type of system becomes necessary as that data becomes more available. Security for data involves a family of protocols that prevent adversaries from unauthorized access to data by use of encryption, authentication, key distribution and key exchange algorithms. Keys are required for encryption and decryption of messages. Key distribution plays an important role in this area and related fields of study. Other areas of security involve prevention of denial of service protocols.

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: Research

Digital Video

The field of Digital Video embraces techniques that span across several disciplines including traditional electrical engineering areas such as signal processing, communication and networking, as well as computer science areas, such as data analysis, content understanding, and databases. Digital Video enriches our everyday lives by enabling various forms of communication and entertainment.

At MERL, we focus on technology that not only improves current video-centric systems, but also establishes a vision for next-generation systems for consumer, business, and government markets. The three areas of concentration are described below.

- Compression: In the past, we have focused mainly on encoder optimization of traditional coding schemes. More recently, we have engaged in the design of new schemes such as multiview video coding, distributed source coding and efficient storage of broadcast video.

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

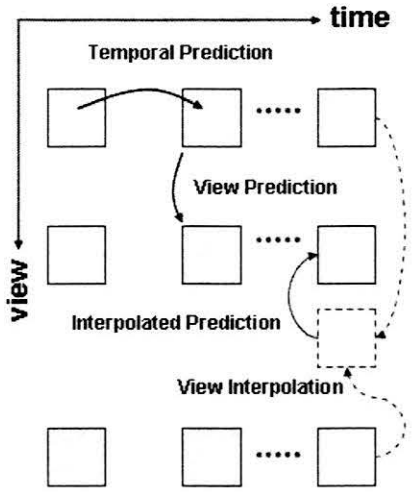
- Analysis: Multimedia storage and retrieval, including video indexing, summarization and audio-visual event detection are active areas of research at MERL. We apply this expertise to integrated systems such as DVD recorders, PVR devices, and surveillance systems.

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

Project Descriptions

Multiview Video Coding 100
Video Compression Using Syndromes 101
Efficient Storage of Broadcast Video 102
H.264/AVC Encoder Optimization 103
Compressed Domain Segmentation 104
DTV Standards in the US 105
3D TV 106
Video Summarization for PVR's 107
ROI-Based Streaming of JPEG 2000 108
Post-Filter for Image/Video Artifact Reduction 109
MPEG Transcoding for Surveillance 110

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 JM 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.

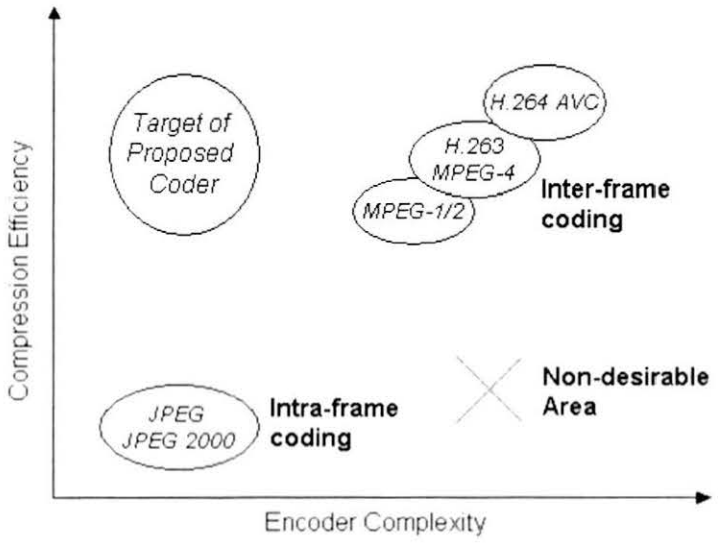
Collaboration: This project is done in collaboration with the Multimedia Information Coding & Transmission Technology Department at Johosoken.

Future Direction: We are working on improving the quality of our multiview video codec and remain active in MPEG standardization activities.

Contact: Emin Martinian
http://www.merl.com/projects/multiview-coding/

Lab: MERL Technology Lab
Project Type: Research

Video Compression Using Syndromes



This work targets the class of applications in which low-complexity video encoding is desired at the transmitter and a greater amount of processing power could be used at the decoder. Applications that may require this type of balance between encoder and decoder complexity include mobile uplink of images and visual sensor networks.

Background and Objectives: Conventional video coding schemes are well suited for broadcast video distribution in which there are many receivers and few encoders. In this model, computationally demanding motion estimation techniques are employed in the encoder to exploit temporal correlation among video frames and achieve good compression efficiency. However, there are a growing number of applications that require a computationally light-weight encoder to conserve battery power, but can tolerate a higher-level of computation at the receiver, while still demanding good compression efficiency. This paradigm shift in application needs is the primary motivation for our work.

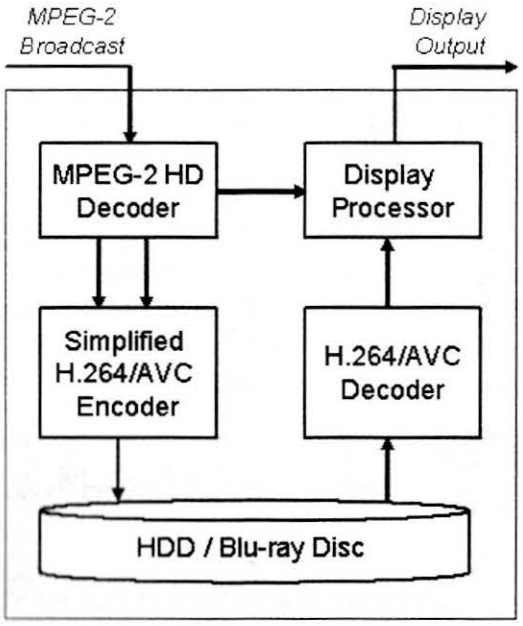
Technical Discussion: For each frame of the video, the transmitter will encode and send two sets of bits. The first set of bits correspond to a low-quality reference that is generated using intra-frame coding techniques. The second set of bits are the syndrome bits, which are essentially a signature of the image data that the decoder will use to recover the high-quality frame. The way that the syndrome bits are computed corresponds to the error-correcting code that is used by the system. Motion estimation is performed in the decoder between the low-quality reference frames to provide improved evidence for the syndrome decoder. With this evidence and the syndrome bits, belief propagation is used to iteratively solve for the reconstructed values of the image.

Future Direction: Improve performance of current algorithm, consider standardization and investigate business opportunities.

Contact: Emin Martinian
http://www.merl.com/projects/video-syndromes/

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 an efficient motion re-estimation and mode decision, which would typically account for more than 80% of a full H.264/AVC encoder. To perform the mode decisions, the MPEG-2 input modes may be mapped to H.264/AVC modes in a simple manner. However, to achieve better performance with some added complexity, we employ the fast mode decision architecture and techniques that have been developed as part of our encoder optimization project <http://www.merl.com/projects/avc-optimization/>. After computing the mode, efficient motion re-estimation algorithms that intelligently reuse the motion information available in the incoming MPEG-2 video stream are applied.

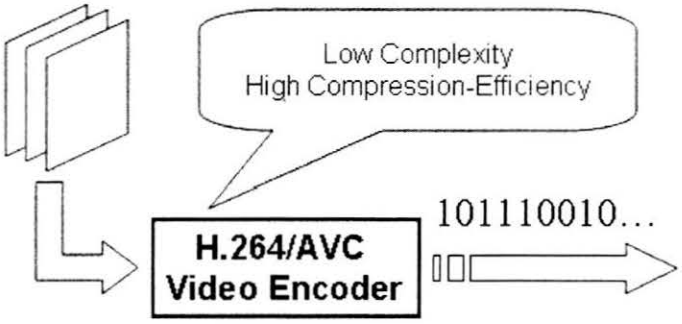
Collaboration: This project is done in collaboration with the Multimedia Information Coding & Transmission Technology Department at Johosoken.

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

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

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

Technical Discussion: Rate-distortion optimization using the Lagrange Multiplier method is widely used in video coding. In macroblock coding-mode decision, the rate-distortion optimization evaluates the Lagrange cost for each candidate coding-mode for a macroblock and selects the mode that yields the minimum cost. The computation of the optimized mode decision is very intensive. First, our approach uses efficient transform-domain techniques to compute the distortion and rate such that the Lagrange cost is calculated with much lower complexity. Second, we use a simple cost function (much simpler than Lagrange cost) to rank the candidate modes, and then we only evaluate the Lagrange costs of the few top modes to decide the mode. These approaches significantly reduced the computational complexity of the encoding while maintaining virtually the same compression efficiency.

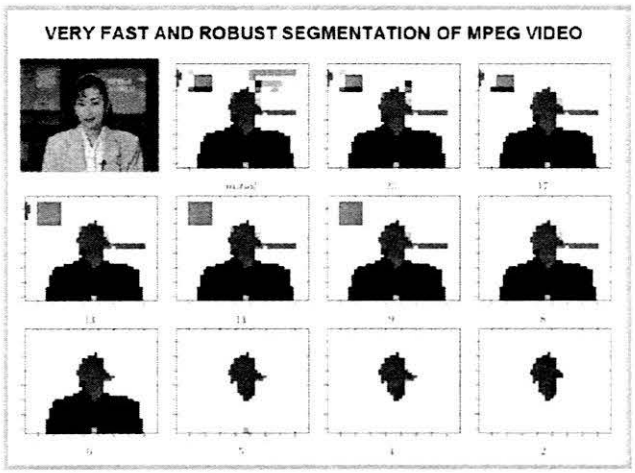
Collaboration: This project is done in collaboration with the Multimedia Information Coding & Transmission Technology Department at Johosoken.

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

Compressed Domain Segmentation



See Color Figure 8

We developed a novel compressed domain video object segmentation method for MPEG or MPEG-like encoded videos.

Background and Objectives: Computational superiority is the main advantage of the compressed domain processing. In addition to computational advantage, the compressed domain video process possesses two important features, which are very attractive for object analysis. First, the texture characteristics are provided by the DCT coefficients with the need of only partial decoding. Second, the motion information is readily available without incurring cost of complicated motion estimation process for not intra only MPEG encoded videos.

Technical Discussion: In the proposed method, we first exploit the macro-block structure of the MPEG encoded video to decrease the spatial resolution of the processed data, which exponentially reduces the computational load. Further reduction of complexity is achieved by temporal grouping of the intra-coded and estimated frames into a single feature layer. The video segmentation is achieved by using the combination of DCT coefficients for I-frames and block motion vectors for P-frames. A frequency-temporal data structure is constructed. Starting from the blocks where the AC-coefficient energy and local inter-block DC-coefficient variance is small, the homogeneous volumes are enlarged by evaluating the distance of candidate vectors to the volume characteristics. Affine motion models are fit to volumes. Finally, a hierarchical clustering stage iteratively merges the most similar parts to generate an object partition tree as an output. The experimental results have shown that the proposed compressed domain video segmentation method provides the similar results as by using spatial domain process with much less computational complexity.

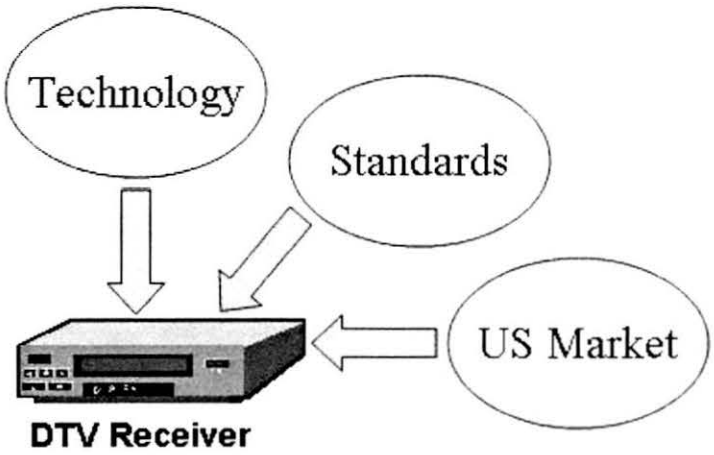
Collaboration: Sentansoken

Future Direction: As future work, we plan to use the compressed domain processing as a precursor to improve the accuracy and computational speed of uncompressed domain segmentation.

Contact: Fatih Porikli, Huifang Sun
http://www.merl.com/projects/compress-segment/

Lab: MERL Technology Lab
Project Type: Research

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 the ATSC, SCTE and CEA, as well as mandates issued by the FCC. Current topics of interest include broadcast flag and associated content protection systems, advanced audio-visual codecs such as Windows Media and H.264/AVC for video, DTV tuner mandate, DTV closed caption, receiver performance guidelines, as well as the standard for unidirectional receiving devices that has been agreed upon between Cable and CE industries.

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

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

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

Lab: MERL Technology Lab
Project Type: Advanced Development

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.

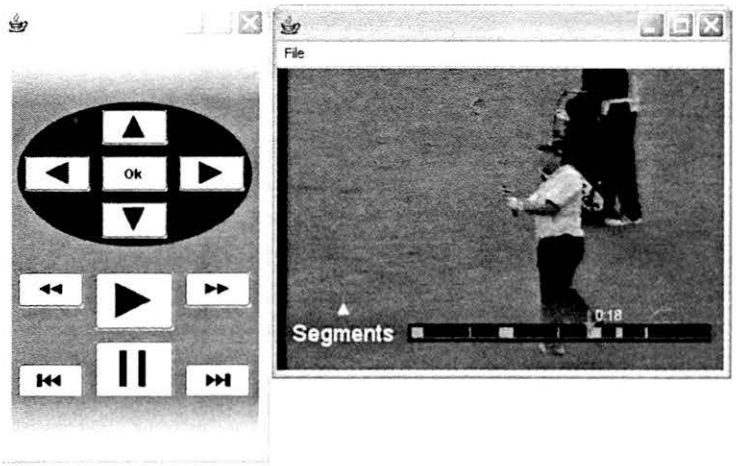
Collaboration: Microlens Technologies, North Carolina, USA.

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



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.

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.

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. We will collaborate with our MERL colleagues in developing convenient user interfaces for the PVR application and running user studies to test the effectiveness of the techniques.

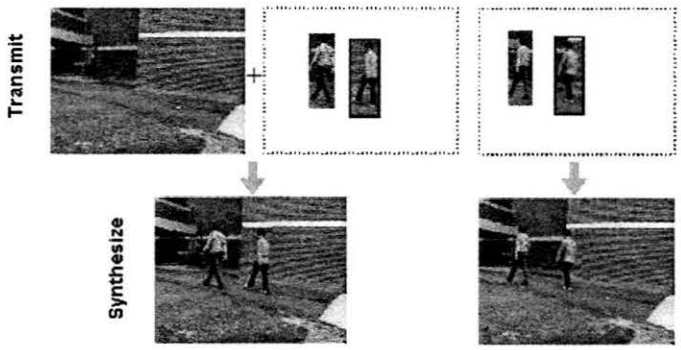
Collaboration: Sentansoken, Johosoken.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

ROI-Based Streaming of JPEG 2000



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

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

Technical Discussion: Several ROI coding techniques for JPEG 2000 images have been proposed in the past few years. These methods can be classified into two categories: static and dynamic ROI coding. The static ROI coding method may be problematic when the application is in an interactive environment since the ROI information is only available during decoding. In the existing dynamic ROI coding method, the ROI was handled by dynamically inserting layers. This causes the packet header tag-tree re-encoding, since the inserted layer does not have the position tag. This is an undesirable feature for real-time image transmission applications. In this project, we have proposed a dynamic ROI coding method, which makes full use of the empty packet property to suppress the non-ROI packets and effectively elevate the priority of the ROIs. The proposed method is very efficient and performs the entire progressive streaming functionalities without packet header tag-tree re-encoding.

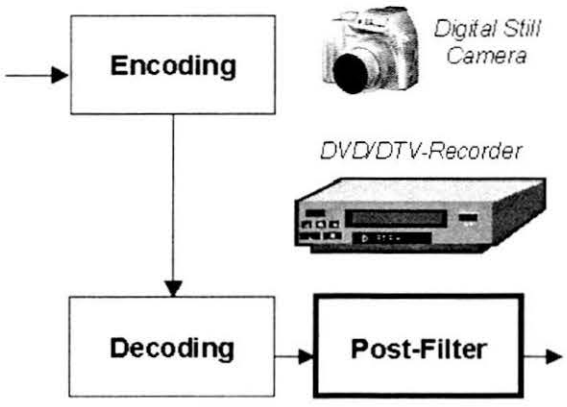
Collaboration: This project is done in collaboration with the Physical Security Group at Sentansoken.

Future Direction: Besides exploring business opportunities for the proposed techniques, we are considering novel error control coding techniques for robust transmission of scalable bitstreams.

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

Lab: MERL Technology Lab
Project Type: Research

Post-Filter for Image/Video Artifact Reduction



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

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

Technical Discussion: We have proposed a new adaptive approach for both blocking and ringing artifacts removal that applies pattern classification techniques to first identify different type of artifacts and then performs the filtering accordingly. Our strategy is as follows: 1) form an edge map based on the local statistics; 2) according to the edge map, detect the blocking artifacts and classify the coding blocks into different categories, e.g., smooth, texture, edge; 3) apply a 1-D low-pass filter to reduce the blocking artifacts and a 2-D fuzzy identity filter to reduce the ringing artifacts. Since the fuzzy filter is applied to the edge blocks only and is able to preserve edges, the filtered images look sharp and clean.

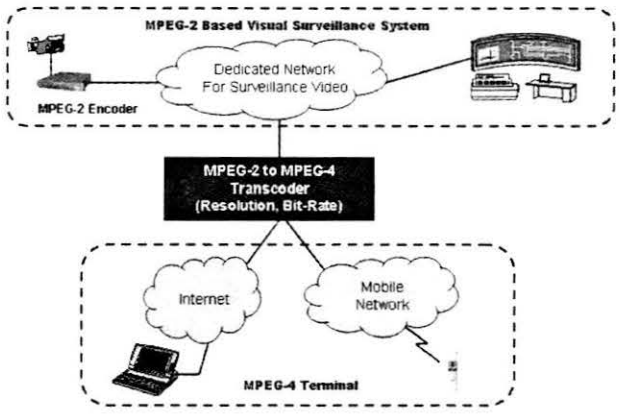
Collaboration: This project has been done in collaboration with the Multimedia Processor Department at Johosoken as well as the University of Delaware.

Future Direction: Continue to research methods on spatial-temporal filtering.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

MPEG Transcoding for Surveillance



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

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

Technical Discussion: Our MPEG-2 to MPEG-4 transcoding software is able to achieve reduced bit-rates, spatial resolutions, and temporal resolutions. The transcoding is done in an efficient way such that multiple bitstreams can be transcoded with general-purpose processors. Conventional cascaded approached decode the original bit stream to the spatial-domain, perform some intermediate processing, and then finally re-encode to a new bitstream. Our proposed architectures simplify this process, while still maintaining the picture quality. Many advanced features have been added, such as the ability to handle adaptive GOP input, improved error robustness in the input and output streams, as well as adaptive field/frame filtering to achieve a reduced spatial resolution.

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

Future Direction: This technology has already been successfully deployed as a product for surveillance systems. We are considering requirements for new application domains.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

Off-the-Destktop Interaction and Display

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

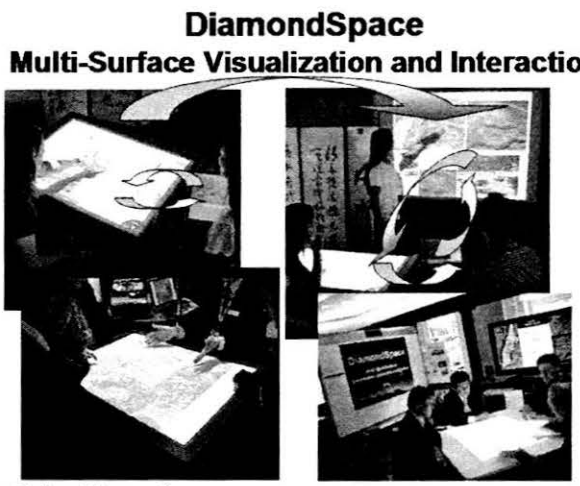
At MERL, we are exploring new user interface devices, paradigms, and interaction techniques that go beyond today's single-user mouse and keyboard interfaces. Our research investigates four areas of newly emerging technologies: (1) spoken-language and intelligent dialog-based interfaces for automotive, handheld devices, and home appliances; (2) new interaction technologies for shared table and wall displays as well as for handheld projectors and TVs; (3) callibration technologies for large multi-projector displays; and (4) scalable fonts for electronic displays of all sizes.

In the speech area, MERL's researchers are working with MELCO business units to bring innovative spoken-language interfaces into automobiles, cell phones, and kiosks. The SpokenQuery projects have developed new speech interaction technologies analogous to spoken Google. Other projects include noise-robust speech recognition, intelligent dialogs, and speech-only communications devices (ComBadge). New input and interaction technologies have been developed for shared table and wall displays. Examples include DiamondTouch and DiamondSpace. Other new forms of interaction have been developed for handheld projectors (HoverPen) and for browsing video on TVs (TimeTunnel and Content-Aware Media Browsing). MERL has continued its outstanding record of projector research this year with multi-projector display callibration for dome displays. And finally, MERL's Saffron has continued on the path to be the world leader in scalable font technology.

Project Descriptions

DiamondSpace 112
DiamondTouch 113
DTFlash: Rapid Prototyping for Multi-User Multi-Touch Applications 114
Temporal Magic Lens 115
DiamondTouch Applications 116
DT Controls 117
HoverPen - New Method for Text Input with a Handheld Device 118
Multi-Projector Imagery on Curved Screens 119
Timetunnel Interface for Video Browsing 120
Content-aware Media Browsing & Digital Video Playback Interface 121
Noise Robust Speech Recognition 122
SpokenQuery 123
FormsTalk: Multimodal Mixed-Initiative Form Filling 124
DiamondHelp: Collaborative Help for Networked Home Products 125
Human-Robot Interaction for Hosting Activities 126
ComBadge 127
Saffron: High Quality Scalable Type for Digital Displays 128

DiamondSpace



See Color Figure 9

This research investigates human computer interaction and data visualization techniques involving multiple interactive display surfaces, including tabletops, electronic walls, and laptop/desktop displays, such that people can explore, understand, utilize and manipulate information from many data sources and types. Application domains include collaborative business and work spaces, emergency response and mission control centers, urban planning rooms, as well as geospatial visualization and analysis.

Background and Objectives: DiamondSpace incorporates many of our earlier research efforts (including PDH, DiamondTouch, DiamondSpin, and UbiTable), and serves as a unifying project for a set of our current research efforts.

Technical Discussion: The DiamondSpace project embarks on four areas of research challenges: (1) Three multi-surface Visualization and Interaction (MVI) conceptual models of data sharing, visualization, manipulation: asynchronous, synchronous and coordinated/multi-view. (2) Interaction Techniques: A number of novel interaction techniques for large touch surfaces have been studied including methods for coordination and conflict resolution for group collaboration, a draggable popup widgets for tools, menu and data called CoR2Ds, a multi-level touch input model for real-time data exploration called Glimpse, and spatial multiplexing called modal spaces as one way to effectively utilizing large display surfaces. Our ongoing research includes interaction techniques that span across multiple displays and interaction surfaces, as well as bimanual touch-gestures. (3) Multi-lens and bifocal displays: We are developing DTLens, a new two-handed interaction technique for multiple people to manipulate, visualize and annotate multiple lenses simultaneously on digital tabletops. (4) User Evaluation: In order to evaluate our research and quantify the advantages of multiple display surfaces and collaborative work practices, we have studied the nature and role of non-speech audio feedback, gradients of privacy and sharing, and overall impact on group interactions.

Collaboration: DiamondSpace is partially funded by the U.S. National Geospatial and Intelligence Agency (NGA) under Contract Number HM1582-05-C-0028 within the ARDA/GI2Vis program Phase 3, and by Mitsubishi Electric Kamakura Works. Our academic collaborators, post docs and graduate student interns have come from Stanford, University of Washington, University of Toronto, University of Calgary and University of Paris.

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

Lab: MERL Research Lab
Project Type: Research

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 (e.g., two hand input) and distinguishes between simultaneous inputs 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

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

DTFlash 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.

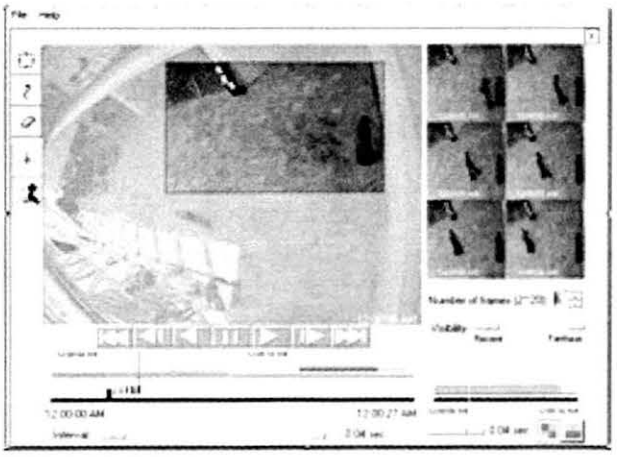
Also of note, DTFlash applications work as regular web pages, allowing for simple deployment and ushering in a new dimension of multi-user enabled web pages which eliminate the need to take turns with the mouse. Flash is also based on vector graphics and optimized for small downloads, so DTFlash applications 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.

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

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

Lab: MERL Technology Lab
Project Type: Initial Investigation

Temporal Magic Lens



A Temporal Magic Lens is a novel interaction technique that combines spatial and temporal querying and browsing for video, and can be used for video surveillance data, home video, or any other recorded digital content. It provides both dynamic and static views of video data, it can be used for viewing real-time video, as well as for browsing and searching archival footage, and it can be used with or without the presence of meta-data. Its framework provides a foundation for other researchers to apply their technology (e.g., meta-data extraction, object recognition, compositing techniques, and interactive timelines) to the domain of video.

Background and Objectives: With more and more video data available, the task of understanding, analyzing, summarizing, or even finding an event of interest becomes a daunting task. Existing automatic techniques may reduce the amount of data by converting a raw video stream into a set of abstract objects (e.g., events, people, trajectories), but there is still a lot of data for a person to deal with. Temporal Magic Lens provides much needed interface and interaction support to deal with the abstract or meta-data extracted by the automatic techniques, as well as with the raw video data. It combines spatial and temporal query components into a single query and presentation mechanism.

Technical Discussion: A Temporal Magic Lens has four key components. (1) The spatial query indicates the spatial region of interest, and is defined as a contiguous set of pixels in any shape. It defines the physical boundaries of the magic lens, and is also the area in which the query results are displayed. (2) The temporal query indicates the temporal region of interest, and may be specified by manual or automated methods. A temporal window may extend into the past, future, or some combination of the two. (3) Rendering and compositing techniques are used to blend multiple frames to create a single frame, encapsulating or summarizing the query results. (4) Drill-down mechanisms provide easy interactions for a user to better understand and explore the summary data provided in the composite view, including methods to determine which frames objects/content came from, and easy visual indications of temporal distance.

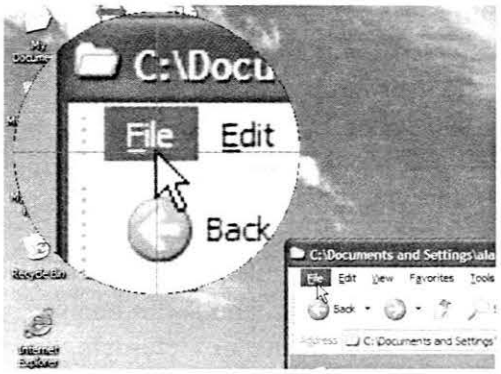
Collaboration: This work is being done by the Data and Sensors group in collaboration with members from the Computer Vision Applications & Devices group.

Future Direction: We have completed an initial prototype implementation. Future work includes exploring the application of this technique simultaneously to input from multiple sources and/or PTZ cameras.

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

Lab: MERL Technology Lab
Project Type: Research

DiamondTouch Applications



The goal for this project is to develop novel applications to provide a shared focus of attention for collaborating users, exploiting DiamondTouch, a novel multi-user input device. (See DiamondTouch Hardware and DiamondTouch SDK for more details on the underlying technology, and DTFlash for a novel authoring environment.) Possible applications include command-and-control rooms, business or technical meetings, and a variety of casual applications (e.g., musical instrument control, home coffee table, games and entertainment, etc).

Background and Objectives: To date, most software applications are intended for single users and designed to utilize traditional input devices (a single keyboard and mouse). In contrast, DiamondTouch is well-suited for shared-display groupware applications; it enables many people to simultaneously interact with the surface without interfering with each other. Our multi-user applications are developed with the DiamondTouch SDK.

Technical Discussion: DiamondTouch applications come in a variety of styles. Some exploit DT's multi-user nature, while others may utilize its multi-touch capability -- each user may touch the unit in more than one place. Furthermore, DiamondTouch's ability to provide identification information for each touch (which users are touching where) is critical in many applications. Our efforts this year were focused on providing better backwards compatibility with legacy software by adding new capabilities to DTMouse. DTMouse is an application to provide mouse emulation capabilities for DiamondTouch. In addition to basic mouse operations, it provides a zooming feature for pixel-level precision input. We are exploring alternative methods as well. Other applications include DTMap, a multi-user map application, and a number of entertainment and gaming applications to showcase DiamondTouch. Most recently we have begun to explore multi-display applications (e.g., table plus wall) for DT.

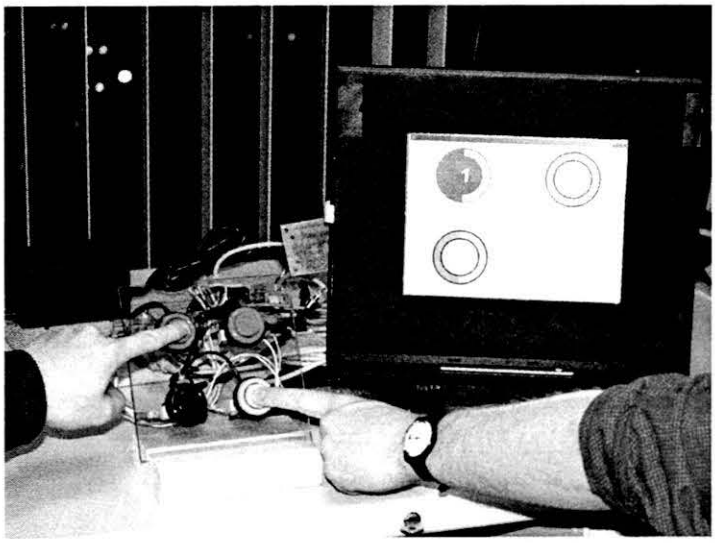
Collaboration: DiamondTouch is a joint project of MERL's Technology and Research Laboratories. MERL is collaborating with MELCO partners from Kamaden, Johosoken, Shaden and Sentansoken. We also have active collaborations with universities who will explore DiamondTouch as a collaborative input technology.

Future Direction: Many of the applications we develop will be used by our MELCO collaborators and their customers. We will work with them to refine the applications, and develop new functionality. Our efforts in this area continue under the DiamondSpace umbrella. We also plan to exploit DTFlash as a rapid prototyping environment for new DiamondTouch application development.

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

Lab: MERL Technology Lab
Project Type: Advanced Development

DT Controls



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 - such as window, seat, and climate controls. By placing receivers in the seats, 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.

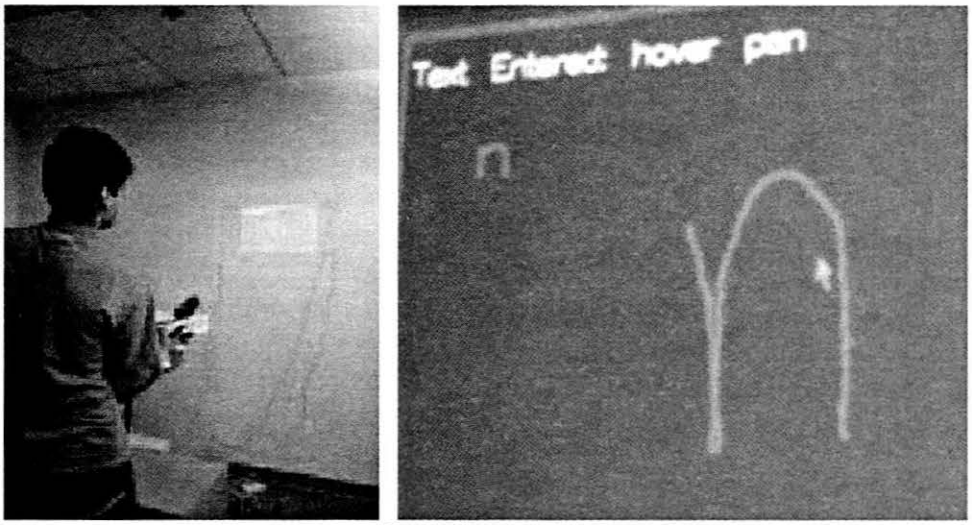
Collaboration: DT Controls should be directly applicable to existing business at both Shahon and Shakaihon.

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 Technology Lab
Project Type: Research

HoverPen - New Method for Text Input with a Handheld Device



See Color Figure 10

HoverPen demonstrates a new way to do text input with a handheld device. The device consists of a projector-camera that is directed at a nearby surface. The user creates script by forming letters 'in mid-air'. The user sees the written script projected to a stable position on the surface, along with the automatically recognized text corresponding to the script.

Background and Objectives: This project extends previous work on interactive handheld projectors to create a complete solution for doing mouse interaction and text entry with a handheld projector. The objective is to demonstrate that a projector can be a valuable addition to handheld devices like cellphones, either supplementing or even replacing the small built-in screen with a larger projected display, while still permitting all the existing user interaction.

Technical Discussion: The handheld device contains a projector and camera. The camera is used to recover the motion of the device relative to the projection surface. This is all the information necessary to project a stable projection of the script that the user is forming by writing in mid-air. We employ a unistroke recognizer to convert the script to recognized text. This could easily be replaced with a cursive script recognizer.

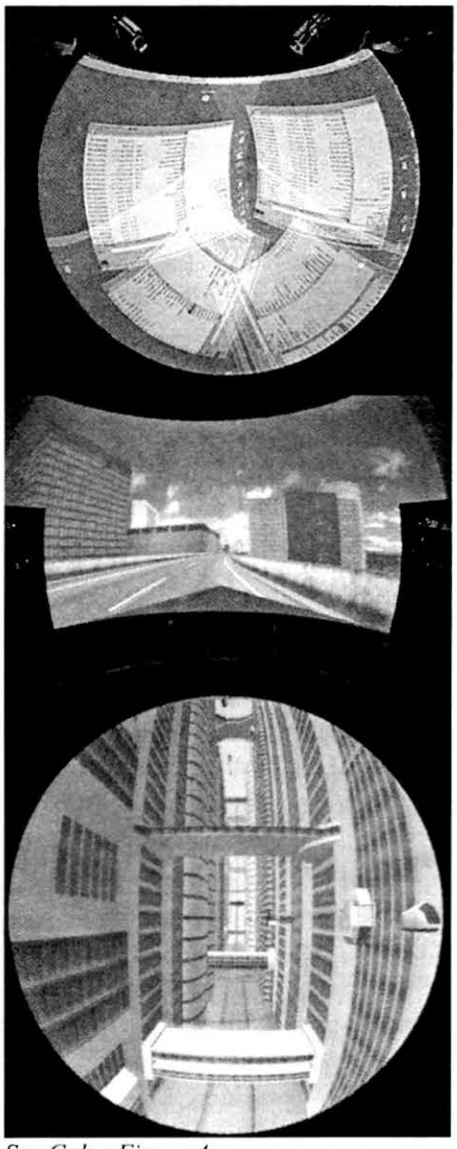
We have built or are building several applications which use this technology. Current work is on an application that allows a user to attach digital text to physical items in the world, and later retrieve the attached text.

Future Direction: The basic technology is mature. We are currently seeking real-world applications and connections with MELCO.

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

Lab: MERL Research Lab
Project Type: Initial Investigation

Multi-Projector Imagery on Curved Screens



See Color Figure 4

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

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

Technical Discussion: We define a new quadric image transfer function and show how it can be used to achieve sub-pixel registration while interactively displaying two or three-dimensional datasets. Accurate estimation of geometric relationship between overlapping projectors is the key for achieving seamless displays. They influence the rendering algorithms and also determine soft edge blending efforts. Our technical contributions are as follows: Simplification of quadric transfer; Calibration methods; Software blending scheme using quadric transfer; Fast rendering strategy exploiting graphics hardware; and Sweet-spot selection and head-tracking support. Accurate alignment of projectors has been demonstrated on 2 and 3 meter diameter dome screens with up to 6 projectors.

Collaboration: Mr Ogata, Mr Wada, MPC; Mr Tanaka, Mr Ashizaki, JohoSoken.

Future Direction: Calibration for large and wrap-around curved screens.

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

Lab: MERL Research Lab
Project Type: Research

TimeTunnel Interface for Video Browsing

TimeTunnel is our technology for browsing stored digital video in fast-forward and rewind modes. It provides temporal context by displaying trails of overlapping sequential future images in reduced size. In collaboration with MDEA, we have implemented TimeTunnel as a new feature of high-end models of MDEA's big-screen high-definition televisions equipped with a hard-disk-based Personal Video Recorder (PVR).

Background and Objectives: Mitsubishi Electric manufactures consumer audio-visual devices such as television sets and personal video recorders and also creates software to record and manage video content. TimeTunnel is an innovative user interface that can boost the effectiveness and attractiveness of these products and differentiate them from the competition. In addition to the fast-forward playback familiar from existing video players and recorders, TimeTunnel simultaneously shows a number of upcoming frames from the video stream in the periphery of the screen, giving the consumer a view into the future of the video they are traversing. User testing shows that significant improvements in user accuracy in fast-forwarding to a desired location are possible with TimeTunnel. Our objective, in collaboration with MDEA, was to implement TimeTunnel as a new feature in MDEA's 2005 product, with no adverse impact on the existing system. We met this objective by exploiting the target platform's high-quality video scaling and frame sampling functionality.

Technical Discussion: The major technical challenge lies in obtaining periodic sequential samples of reduced-size images that are in the future with respect to the current image. In a previous PC-based prototype, our implementation reads ahead in the MPEG-2 stream and extracts dc-images (1/8 scale images easily extracted from the compressed data) from the I-frame pictures. But the MDEA television uses the IEEE 1394 Audio/Video Control disc interface, which does not support read-ahead. Also, buffering the MPEG-2 stream before decoding would add complexity to the critical path for decoding. Instead, we observe that the platform already supports high-quality video downscaling and sampling of decoded video images, which our implementation uses to obtain the images. The TimeTunnel display is rendered using the platform's graphics overlay and scaling facilities. This implementation incurs no additional manufacturing costs because it is purely a software enhancement.

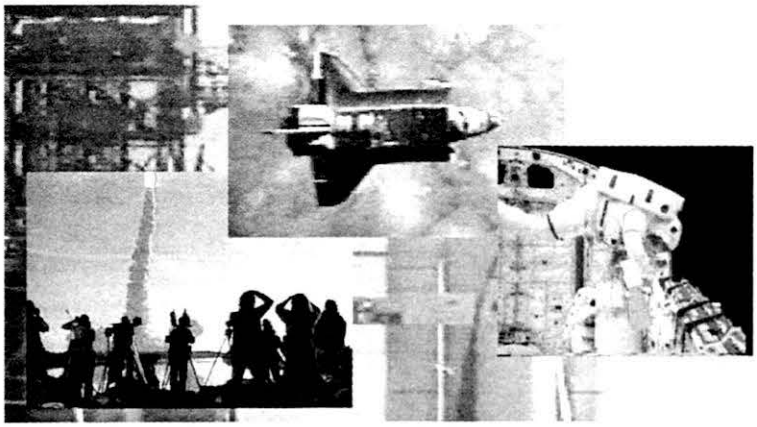
Collaboration: MDEA, ID-Ken.

Future Direction: We plan to continue our relationship with MDEA and continue improving TimeTunnel. We hope to transfer this technology to other MELCO products as well.

Contact: Samuel Shipman
http://www.merl.com/projects/timetunnel2/

Lab: MERL Technology Lab
Project Type: Advanced Development

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.

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 are simply scaled to fill the larger display surface. Furthermore, each frame of the video is displayed in place of the previous frame. 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.

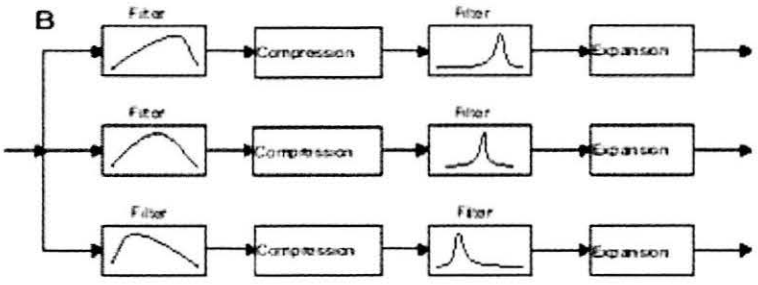
Technical Discussion: The input video is a series of frames that are normally displayed sequentially. The output video is the same series of frames that have been scaled, rotated, filtered, and displayed in parallel on different regions of the display in a manner that not only preserves the continuity 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: We plan to adapt our prototype video playback interface for use in a media browsing tool. Short summaries of recorded videos will not only give the viewer a taste of what the available programs include, but also present these previews in a manner consistent with their style and content.

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

Lab: MERL Research Lab
Project Type: Research

Noise Robust Speech Recognition



The aim of the project is to develop a low-power feature computation module for noise-robust speech recognition. Such a device can be used with embedded recognizers in devices ranging from cellphones to monitoring devices that are meant to remain always on.

Background and Objectives: Automatic speech recognition (ASR) is a computation heavy process that places a significant power burden on the device on which it is being performed. Consequently, it is currently infeasible to perform ASR continuously on embedded devices, and other devices with a low power budget. Current technology restricts the use of ASR to limited interactions that run only for short periods of time. One of the largest consumers of power in an ASR system is the "front-end", i.e. the module that computes cepstral feature vectors from the speech signal. In state-of-art systems the least expensive front-end modules still consume 250 milliwatts. These low-power front ends also usually compromise on performance to minimize power consumption. Another problem is ASR is that it is highly susceptible to noise - high background noise levels reduce accuracy.

In this project we have designed a "front-end" processor that can compute cepstral features from speech using only about 250 micro watts of power. In addition, these features will be highly robust to noise, contrary to the conventional approach that compromises noise robustness for power. Such a device can greatly enhance our ability to deploy low-power ASR systems in an "always-on" mode.

Technical Discussion: Human speech perception is known to be highly robust to the degrading influence of noise. It stands to reason that any technique that mimics the processes in the human ear can be expected to be more noise-robust than techniques that are based on simple signal processing strategies. However, the former are also known to be computationally highly expensive.

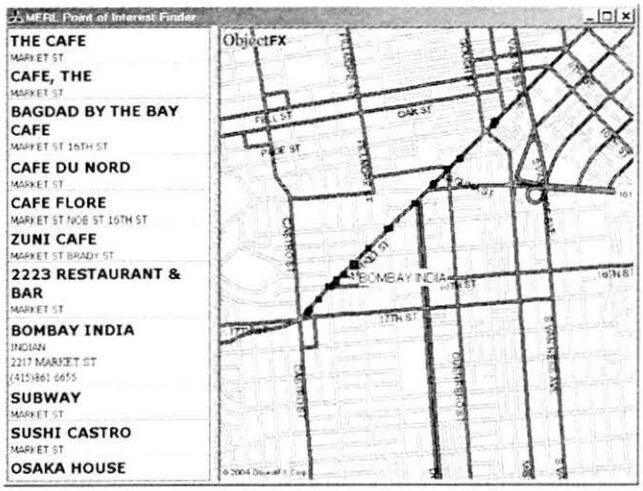
In this project we have designed a signal processing scheme that can be implemented as an analog filterbank that mimics some of the processes in the human ear by companding the speech signal. This procedure is observed to enhance spectral peaks in speech - i.e. the most important characteristics of the speech signal.

Collaboration: Rahul Sarpeshkar and Lorenzo Turicchia, MIT.

Contact: Bhiksha Raj, Bent Schmidt-Nielsen
http://www.merl.com/projects/robustsr/

Lab: MERL Research Lab
Project Type: Research

SpokenQuery



SpokenQuery is technology for accessing databases using a verbal description of the desired information. It is particularly useful in applications where hand and eyes free operation is desired. These include: automotive navigation and entertainment systems, home entertainment systems, cellphone information systems, hand held industrial systems, call centers, and information kiosks.

Background and Objectives: An important part of many products is choosing an item or a small set of items from a very large set. Examples include: selecting point of interest, choosing a television program, or purchasing a song. However, current user interfaces to these tasks require either typed input, or multiple menu selection, and there are many situations where this is not acceptable. For example, typing queries would obviously not be acceptable while driving an automobile. The objective SpokenQuery is to enable this selection using a verbal description of the desired item. The result is a list of items that are judged to be "pertinent" to the query. This is similar to current Information Retrieval systems (e.g. Google), the list is not exact but should contain a significant number of useful items. SpokenQuery solves several problems associated with searching a large database using a standard spoken user interface. With SpokenQuery, there is no grammar to memorize, no tedious menu tree to navigate, and the system is more robust to misrecognition and out-of-vocabulary speech. The ability to access information and to select entertainment, products and services in the automobile, on the cellphone, and on televisions will enable large new markets.

Technical Discussion: SpokenQuery is implemented with a combination of speech recognition and information retrieval techniques. SpokenQuery is implemented as a post-processor for a speech recognition engine. Therefore, it is not tied to any particular speech recognizer. Since SpokenQuery does not make use of any language specific semantics it is also not tied to any language. Finally, SpokenQuery dramatically lowers the cost of building a spoken user interface because the index files are created automatically from the contents of a database. This feature also enables applications of spoken search in dynamic data such as news, traffic, and weather.

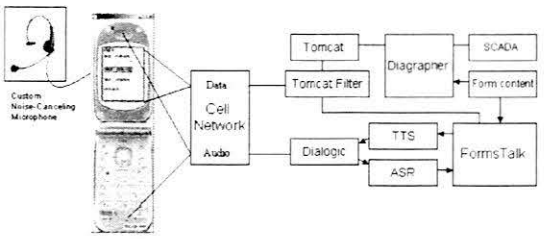
Collaboration: AdvisoryAgent (Sentansoken), Noise Robust Speech Recognition.

Future Direction: Port to Japanese and other languages; port to embedded and client server devices; improve accuracy performance; reduce memory and processor footprint; and produce prototype products and services for MELCO Business Units

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

Lab: MERL Technology Lab
Project Type: Research

FormsTalk: Multimodal Mixed-Initiative Form Filling



FormsTalk is middleware for building form-filling applications that support a mixture of speech and non-speech interaction modes. Examples of non-speech modes include touch screen, keyboard and mouse, and telephone keypad. FormsTalk also supports flexible, mixed-initiative interaction, in which either the user or the system can take the lead, depending on circumstances. In 2004 we built a set of extensions to FormsTalk to provide remote access to industrial plant control data via cell phone.

Background and Objectives: Form-filling is a common framework for many different kinds of applications. A multimodal interface approach improves the accessibility of these applications by allowing users to choose whichever mode (speech, touch, etc.) is the best match for their capabilities and the current task. To date, developing such interfaces has tended to be very labor-intensive, with a lot of application-specific code. The goal of FormsTalk is to reduce the amount of application-specific code, so that most of the labor for a new application is involved in authoring the content of the forms, and deployment on different platforms requires a minimum of additional effort. To date, FormsTalk applications have been deployed on PCs, phones, and kiosks in three languages.

Technical Discussion: FormsTalk is built on top of DiamondTalk, which is an application-independent Java architecture for building conversational, multimodal spoken-language interfaces. DiamondTalk allows us to easily substitute different speech recognition and generation engines (e.g., from different vendors), as technologies and applications change. FormsTalk also uses Collagen (see project description) as its dialogue manager, which provides its mixed-initiative capabilities.

A key part of FormsTalk is a modular architecture which supports the use of special purpose components in application domains where these are necessary. Extension of FormsTalk to the current cell phone application included implementation of components allowing integration with an Apache web server and with a Dialogic telephony board. These components can be reused in future applications. We also designed a special purpose interaction protocol allowing both the voice and data channels of the cell network to be used with an off-the-shelf phone handset. Smoothly integrating the differing capabilities of the two channels was a fundamental part of our work this year.

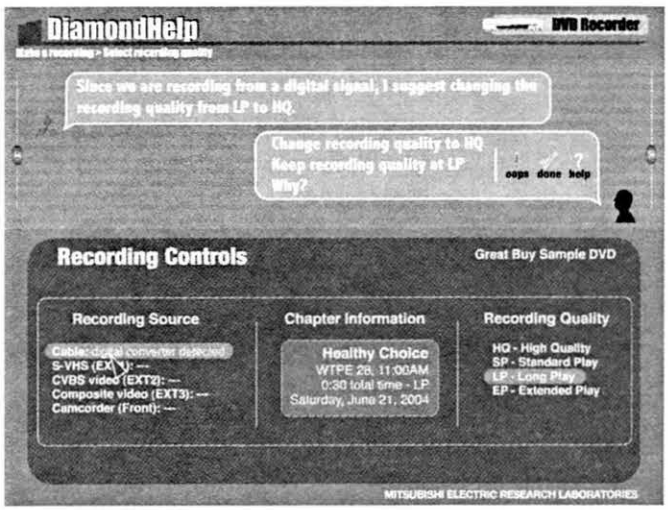
Collaboration: FormsTalk is the user-interface portion of the "Broadband Web Services" project led by the System Technology Dept. of SentanSoken. We are also members of the W3C Multimodal Interaction and W3C Voice Browser working groups, which are developing relevant standards.

Future Direction: Investigate other applications domains for FormsTalk.

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

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.

DiamondHelp is currently a finalist in two international design competitions: the competition associated with the Third Interational Conference on Appliance Design, and INDEX2005.

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

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

Lab: MERL Research Lab
Project Type: Advanced Development

Human-Robot Interaction for Hosting Activities

We are developing a collaborative robot that not only holds conversations with people, but also uses physical gestures and movement to interact with them, thereby "engaging" the human in the interaction.

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 or tasks they need help performing. In museums and stores, it directs people to locations of interest and can demonstrate objects of interest.

Technical Discussion: Our robot collaborates with people to perform the hosting task of demonstrating the iGlassware system developed at MERL. The robot can seek out a person to interact with for a demonstration. People interacting with the robot do not need any training to interact. In recent work, the robot also recognizes and responds to their head nods, in addition to the previously developed interpretation and production of head gestures to indicate interest in the robot and environment. Over 100 people have completed demonstrations with the robot. Studies of people interacting with the robot show that they find the robot's gestures more natural than an unmoving conversational partner, and that they direct their attention more to the robot than the unmoving conversational partner.

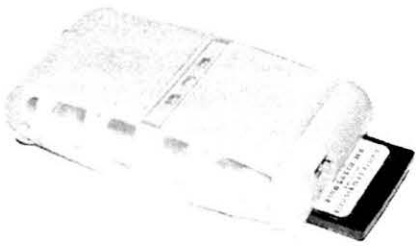
Collaboration: We have collaborated with members of the MIT Vision Group (http://www.ai.mit.edu/projects/vip/projects.htm) to use the Watson vision system and HMM algorithms to interpret head nods of people interacting with the robot.

Future Direction: We will conduct user studies of head nodding during conversation. Using a now fully mobile robot, we will program it to approach people to engage in conversation, and use its body, as well as head movements and conversation, to engage users in new demonstration abilities.

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

Lab: MERL Research Lab
Project Type: Research

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: We are working quite closely with the speech applications group in MERL MTL. Through collaboration with Eric Brewer and the TIER project (Technology and Infrastructure for Emerging Regions) at UC Berkeley, we are conducting a user study in Pudupalayam village in Tamil Nadu in India during July 2005. This work and its continuation will assess the technical, social, and business aspects of ComBadge.

Future Direction: We are developing automated services that can be contacted via ComBadge messages. Initially, control of household devices and a voice portal to traffic, weather, and stock quotes will be implemented. Also, we are investigating 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

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 as the new type rendering engine in Flash Player 8 and Studio 8 authoring suite. The Macromedia Flash Player is the most widely deployed multimedia player, reaching close to 100% of all Internet-enabled desktops.

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 for future development is to adapt Saffron for use in embedded applications and on mobile devices. This includes enhancing performance and reducing memory requirements. Continued licensing efforts to deploy Saffron in the marketplace.

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

Lab: MERL Research Lab
Project Type: Research

Sensor and Data Systems

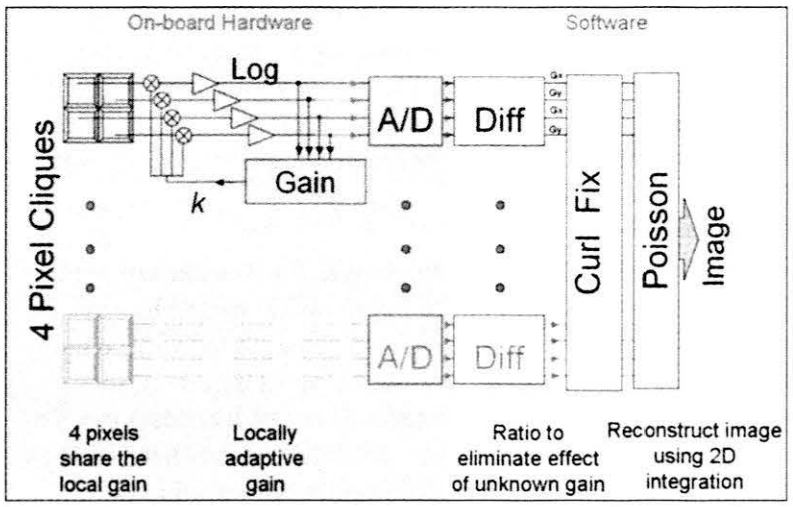
Until recently, computer applications have largely presented the user with human-generated content that has been processed according to fixed or human-generated algorithms. Now it is becoming possible for a computer system to collect data from the local environment on its own, process that data adaptively,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

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

Project Descriptions

Gradient Camera 130
Acoustic Doppler Sensors for Surveillance 131
Acoustic Doppler for Denoising Speech Signals 132
Low-Power, Low-Bandwidth Communication for Sensor Networks 133
Advance Elevator Call 134
Ultra-Low Resolutions Sensors in Panoptes 135
RFID Application 136
LED-Based Sensors 137
Real-Time Audio Buffering for Telephone Applications 138
Sound Recognition 139
Data Mining by Random Walks 140
Dimensionality Reduction 141
Intelligent Multi-Dimensional Data Summarization 142
Human-Guided Antenna Design 143

Gradient Camera



We propose a gradient-based camera that relies on ratios of neighboring pixel intensities rather than absolute intensity measurements. A gradient camera is similar to existing intensity cameras in electro-optical structure, but by measuring only the local changes in the image, we gain some significant advantages. It is most similar to the locally-adaptive gain cameras. However, unlike locally-adaptive gain cameras where the gain is different for each pixel, in gradient camera the local gain is same for each clique of four neighboring pixels. Since the ratio between neighboring pixels in the clique is invariant to the gain, we do not need to record the gain. Gradient camera needs little or no exposure metering to capture high contrast scenes, hides effects of quantization well and distributes noise as low-frequency error rather than masking high frequencies.

Background and Objectives: Capturing images of a scene with large variations in intensity requires a camera with a high dynamic range. This is sometimes achieved with image sensors with logarithmic response curve. On the other hand, successful biological visual systems (including the human eye) use sense change more acutely than absolute intensity. Our goal is to build a gradient camera than captured subtle detail as well as large variations in intensity.

Technical Discussion: Quantizing sensed intensity differences between adjacent pixel values permits an ordinary A/D converter to measure detailed high contrast, high dynamic range (HDR) scenes. Once we have computed pairwise ratios of intensities, we reconstruct the original image using a 2D integration by solving a Poisson equation. The pairwise ratios eliminate effect of unknown gain within each clique. We measure alternating "cliques" of sensors (small groups) that locally determine their own best exposure. This intrinsically differential design suppresses common-mode noise, hides and smoothes quantization, and can correct for its own saturated sensors. Simulations demonstrate these capabilities in side-by-side comparisons.

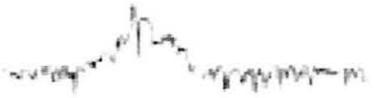
Collaboration: Jack Tumblin, Northwestern University, Amit Agrawal, U of Maryland.

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

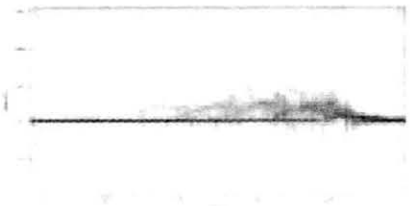
Lab: MERL Research Lab
Project Type: Research

Acoustic Doppler Sensors for Surveillance

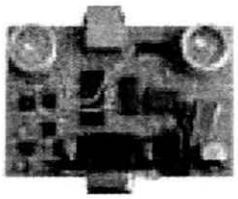
Doppler spectrum at peak stride



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 expanding the technology 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.

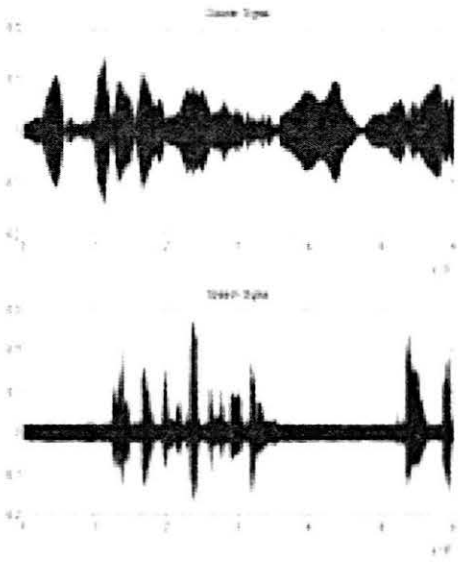
The Doppler based surveillance technology can be used by itself, or, more effectively, to augment other surveillance sensors such as cameras and microphones.

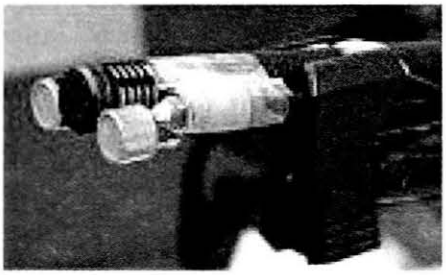
Future Direction: The current technology is suitable for constrained spaces such as corridors. Our next step is to expand the technology to normalize for angle of approach, and enable a wider range of spaces. We are also in the process of combining the Doppler sensors with cameras for superior surveillance.

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

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 both 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 use the correlations between these motions and the speech signal produced by the speaker to reduce the effect of extraneous noises, since the latter do not affect the Doppler signal to the same degree as the speech signal

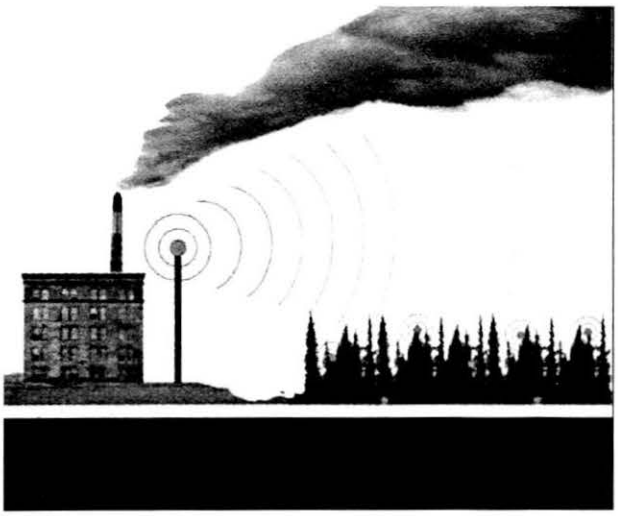
We have thus far achieved significantly improved speech activity detection using the Doppler signal for speaker-independent scenarios, as well as improved denoising in situations where the identity of the speaker is known.

Future Direction: The current technology achieves superior speech activity detection in speaker independent scenarios, as well as improved denoising when the identity of the speaker is known and training data are available for the speaker. Future work will be aimed at improved speaker-independent denoising.

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

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 100m, a data rate of 1 bps 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 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 moveable 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 achieveing 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 circuit (to prevent the hub's transmitter from jamming its co-located receiver) and novel modulation schemes.

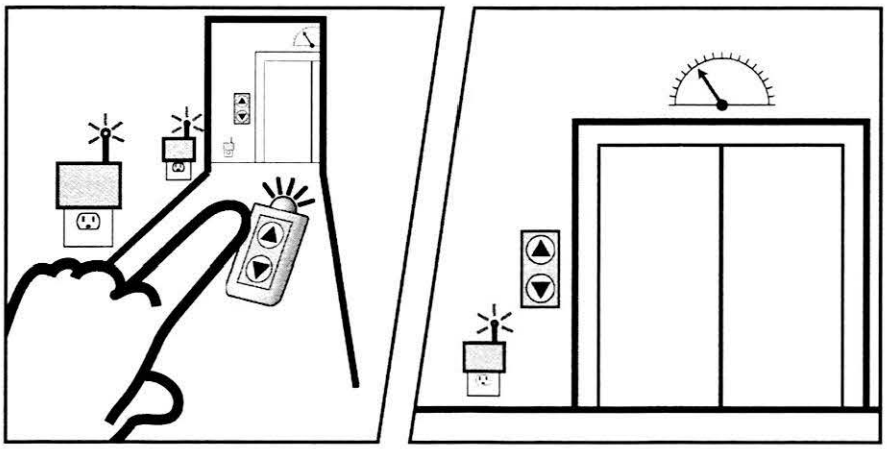
Collaboration: There is currently no formal collaboration with any labs or business units in Japan. However, we have contacted some and had informal conversations.

Future Direction: In the future we will investigate inexpensive, low-power communication systems with significantly longer range (up to 1 km) and slightly higher data rates (about 5 bps).

Contact: Darren Leigh, Joseph Katz
http://www.merl.com/projects/waesn/

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

Ultra-Low Resolutions Sensors in Panoptes

One bit per pixel per square meter: that is one of the primary lessons learned in a pilot study into the usefulness of ultra-low resolution sensors (ULRS) for understanding human behavior in built environments carried out last year at MERL. At this resolution it is possible for automatic systems to learn to recognize patterns of human behavior and exploit those patterns to improve building efficiency, elevator response times, and effectiveness of security measures. This project will prototype a hardware implementation of a ULRS platform to facilitate future research, and also to point the way toward a commercially feasible implementation.

Background and Objectives: Work began in the Spring of 2004 on the design and implementation of an Ultra-Low Resolution Sensor package. Those plans are coming to fruition in 2005 with a prototype run of 500 sensor nodes. These sensors will be installed alongside smaller, camera-based systems to facilitate direct comparisons of the sensor modalities.

The ULRS package will also allow cost-effective examination of significantly larger areas by allowing data collection over a thousand square meters, as opposed to one hundred square meters in previous studies. In this way, the facility will offer the unique opportunity to study the behaviors of people on a building-wide scale.

From the commercial perspective: efficiency, convenience and security are the promises of smart buildings. Smart buildings are those that can sense their inhabitants and adjust to accommodate their needs. Unfortunately cameras and computer vision approached require significant computation and high-bandwidth communication. Our experiments indicate that with clever sensor and network design, ultra-low resolution sensors will provide a light-weight, low-bandwidth solution to this problem.

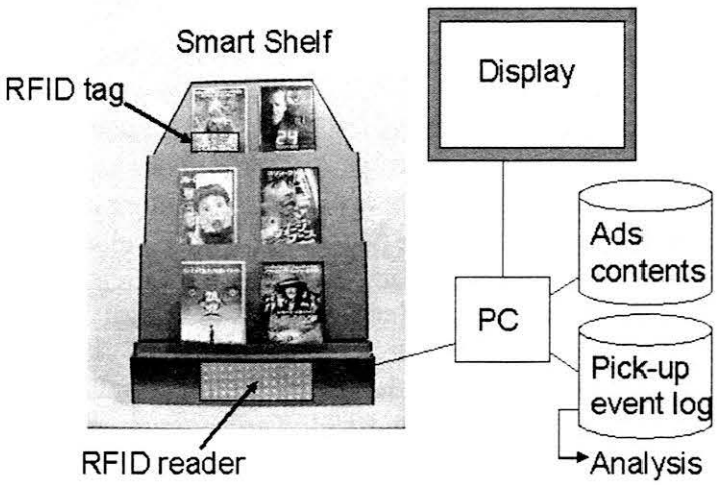
Technical Discussion: Through collaboration with the Massachusetts Institute of Technology's PlaceLab project, we are currently implementing two test facilities that will each cover approximately 800 square meters of office space. One facility will cover all of the MERL public spaces and hallways. The other will cover another, equivalently-size office space. The nodes, which will sense and report nearby motion, are a combination of PlaceLab wireless communication boards and a MERL-designed sensor package. The communication boards operate in the 2.4Ghz band, and could possibly be replaced with Zigbee hardware in the future.

Collaboration: MTL, MIT PlaceLab.

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

Lab: MERL Research Lab
Project Type: Research

RFID Application



We developed an RFID-based "Smart Shelf" that detects the presence or absence of each item on the shelf. When an item is removed from the shelf for examination, relevant advertising is displayed. The shoppers' behavior can be recorded in a database and used for marketing purposes. We conducted field trials in real stores to evaluate the reliability of the data and the applicability to marketing.

Background and Objectives: The RFID market is rapidly expanding in applications such as supply chain management, logistics, and other retail systems. Although the first phase of RFID adoption has been occurring in pallet and case level logistics so far, greater benefit is expected in item level adoption. However, in addition to technical issues like RF interference and tag cost issues, it is unclear how to create profitable systems which directly benefit consumers.

To better understand and overcome these problems, we have been integrating RFID into retail application concept systems. This integrated system will serve as a test bed for identifying the RF issues in real store environments and proposing new applications to benefit both consumers and retailers.

Technical Discussion: The smart shelf system uses RFID tags which are attached to the items and RFID readers which are embedded underneath the shelf. The readers periodically interrogate the tags so that the system can manage real-time shelf inventory. Furthermore, the system can detect the removal of an item when a customer picks it up. The "pick-up" event changes the content of a computer-driven display to show advertising and information related to the item. Thus, the system provides benefits beyond simple inventory management by attempting to directly influence sales. At the same time, the "pick-up" events are recorded in the database and can be used for off-line marketing analysis and real-time analysis for interactive sales promotion on the shelf display.

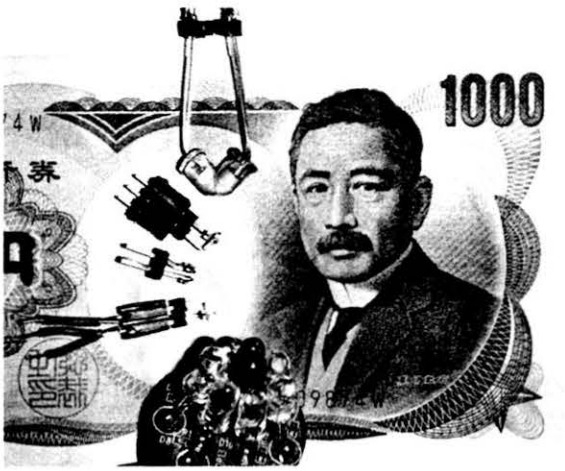
Collaboration: DM dept of Johosoken, MDIS, MDIT.

Future Direction: Our next step is to integrate the real-time recommender into the concept system for enhancing the interactive sales promotion feature.

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

Lab: MERL Technology Lab
Project Type: Research

LED-Based Sensors



We have developed an ultra-cheap LED based sensing system that employs LEDs for both emission and detection. These devices are used in combination with a novel microprocessor interface, which enables us to measure changes in back-reflected light using a single digital I/O pin and directly producing a PWM signal (no power-hungry A/D converter needed). The system holds great promise for following color changes in chemochromic sensing materials, and surface-state and contamination-detection on a wide range of surfaces.

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

Technical Discussion: A wide range of chemical species can be detected using colorimetric reagents. Often these reactions are carried out in solution, but many can be translated into the solid-state by processing in an appropriate polymer matrix. Therefore, almost any colorimetric reaction could be immobilized onto the surface of a dual-LED to create a solid-state sensor. Uncoated Dual-LED sensors have also been used successfully to follow colorimetric reactions in solution, with detection possible at very low indicator dye concentrations (~ 5 parts per billion). In addition to the sensing capabilities of this system, it holds the added bonus of offering short-range wireless communication of data.

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

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

Contact: William Yerazunis
http://www.merl.com/projects/LED_chemical_sensors/

Lab: MERL Research Lab
Project Type: Research

Real-Time Audio Buffering for Telephone Applications

On ordinary phones, if you miss something, it is gone forever. With real-time audio buffering, you can hear what you have been missing. Best of all, the operation is intuitive, requiring little or no action on the user's part.

Background and Objectives: The cell phone market is extremely competitive. Many phones have similar capabilities and prices. To stand out in the crowd, unique features are needed. We propose real-time audio buffering as an inexpensive feature users will appreciate.

Real-time audio buffering senses when the phone is removed from the ear. When the phone is returned, the missed incoming audio is played back, but at a rate faster than real-time. Very quickly, the phone catches back up to real-time, and the user misses nothing, despite the distraction. The ear sensor also allows us to create an elegant "instant-replay" function. Briefly move the phone away from the ear, and it automatically replays the last few seconds. Repeating this gesture skips successively further back in time.

Technical Discussion: Real-time audio buffering uses a capacitive proximity sensor to determine when the phone is near the ear. The sense electrode is created by coating the inside of the ear piece with a conductive coating as might normally be applied for shielding purposes. In addition, the remainder of the case interior is similarly coated and connected to the phone's internal ground. Holding the phone up to an ear capacitively couples the ear sense electrode to the phone's ground through the user (via ear and hand). By measuring the change in capacitance, we can easily detect the phone proximity to the ear.

The audio section constantly stores incoming audio in a circular buffer. Normally, the "play pointer" immediately tracks the "record pointer". When the phone is removed from the ear, the record pointer continues adding incoming audio, but the play pointer is halted until the phone is returned to the ear. To help give the user the context of the conversation, it helps to move the play pointer a little further back in time each time the phone is removed from the ear. This implements the instant-replay function as a side benefit. When the play pointer is behind the record pointer, we speed up play back via a pitch-preserving technique and skip silent intervals.

Recently, we have noted other uses for the ear sensor, in addition to the buffering function. Intelligent display control improves both power consumption and functionality by automatically turning off the display when the phone is pressed against the ear, and turning it on when removed.

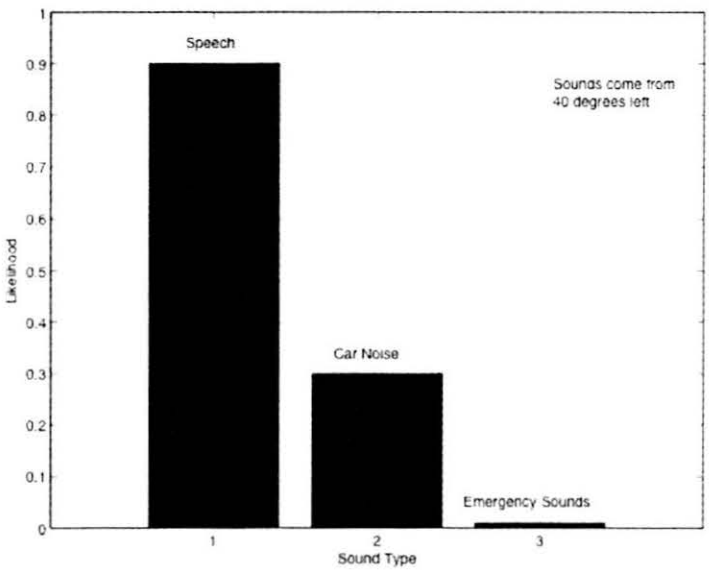
Collaboration: Working with MDEN, we have created a robust demonstration system which has been supplied to certain customers for evaluation of the buffering function.

Future Direction: We are awaiting customer feedback to determine our next steps.

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

Lab: MERL Research Lab
Project Type: Advanced Development

Sound Recognition



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

We anticipate this technology to provide the basis for various types of audio sensing which can find applications in surveillance, factory automation, entertainment media analysis, and other domains where audio feedback can be critical.

Background and Objectives: Sound recognition has long been a neglected field of research. So far it has been dominated by speech recognition work that has found limited success for anything other than speech sounds. Our work is focused on generalized sound recognition, a framework equally capable of working with any type of sound regardless of its nature and its recording conditions. Our objective is to provide lightweight and cost-effective monitoring capabilities using sound.

Technical Discussion: The sound recognition project is already a couple of years old and in that time we have extended this project to encompass more than just recognizing sounds. We have developed new methods that can identify sounds coming from different positions (thereby eliminating the need for localization), we have developed approaches that conceal data between two cooperating parties in order to perform secure recognition, and we have scaled our algorithms to work on embedded machines such as cell phones and PDAs.

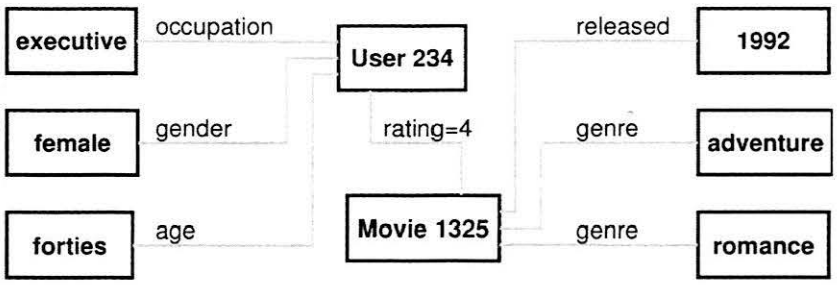
Collaboration: Iwamoto-san, Hashimoto-san, Sentansoken.

Future Direction: Our current direction of research is looking at integrating the sound recognition framework with PTZ cameras to assist their movements and also lighten the load of human operators monitoring them. Preliminary technical work is completed and we are currently finishing developing our first prototypes.

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

Lab: MERL Research Lab
Project Type: Research

Data Mining by Random Walks



To model consumer behavior from relational databases, we view these databases as graphs that associate people, products, and attributes. We develop statistics of random walks on these graphs that are highly predictive of future associations (e.g., purchases). These statistics enable product recommendation services that optimize both the satisfaction of the consumer and the long-term profit of the vendor.

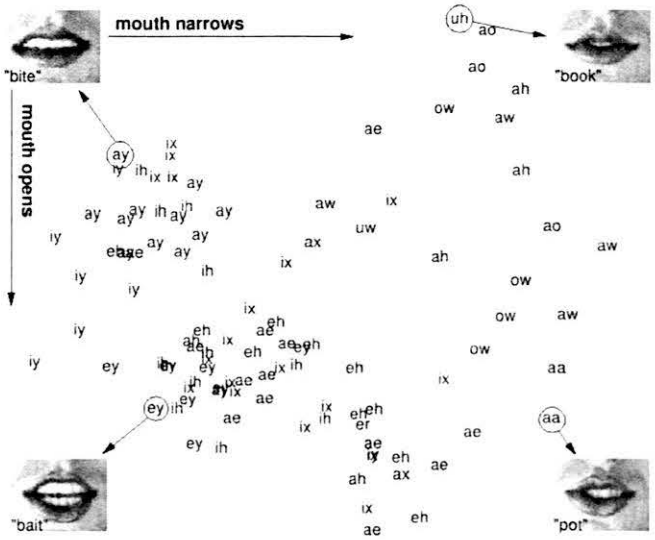
Background and Objectives: The overwhelming majority of the world's data is in relational databases. These are hard to mine, because almost all statistical methods are specialized for tabular data and other very simple data structures. For example, collaborative filtering usually operates on tables in which consumers rate products. We sought to develop a collaborative filtering framework that can incorporate and exploit all the messy structure and side information in relational databases. When available, information about age, gender, occupation, product category, manufacturer, profit margin, etc., is naturally incorporated into the inference process.

Technical Discussion: We model a random walk on the graph with a Markov chain. The expected time to wander from a person to a product and back is called the commute time. This considers all possible paths between the person and the product, including whether they ever bought a related product, whether the manufacturer appeals to their market demographic, etc. Small commute times indicate likely associations, but this statistic is heavily biased to favor generically popular items, i.e., it is more likely to recommend heavily advertised products rather than products for which a person has a special affinity. To correct, we embed the Markov chain in a Euclidean space, interpreting commute times as squared distances. Angles in that space indicate how two items are correlated in their random walks -- how similar they are in their relationships to all other items. These correlations also factor out generic popularity, so that recommendations are tailored to the individual's tastes and attributes. Real-data experiments indicate that correlations are far more predictive of consumer behavior than commute times and all other conventional statistic used for collaborative filtering. In principle, correlations require inversion of an impractically large matrix, but we developed very fast methods to approximate the relevant submatrix of the inverse. This also makes it practical to compute the recommendation that optimizes the long-term profit of the vendor.

Future Direction: We are now developing very fast exact methods with the intent to model millions of customers, products, and attributes.

Contact: Matthew Brand
Lab: MERL Research Lab
http://www.merl.com/projects/randomwalks/
Project Type: Initial Investigation

Dimensionality Reduction



We are developing manifold models of high-dimensional signals that capture much more information than classic dimensionality reduction methods, yielding superior performance in compression, classification and regression tasks. This technology applies to a vast array of multimedia applications, and allows us to manipulate and edit data in new ways. The image visualizes part of a manifold of speech acoustic features. The embedding organizes vowel sounds (TIMIT phoneme codings) by mouth shape. Phonemes are better separated on this manifold than in acoustical feature space; that may lead to improved speech recognition.

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

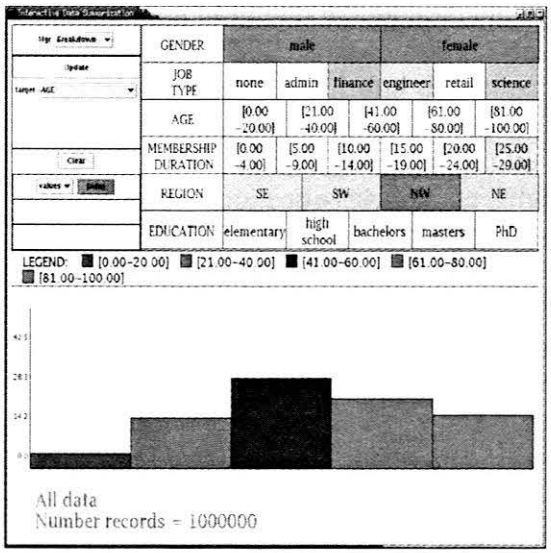
Technical Discussion: We illustrate with an example of faces. Although facial data typically has millions of degrees of freedom (e.g., face images have 106 pixels), faces probably have a relatively small number of degrees of freedom. The manifold of faces, which collects all possible faces, is a thus low-dimensional surface in image space. We seek a function that relates the face manifold's intrinsic coordinate system to face images. Because the manifold is not known, it must be estimated from sample data. We have developed estimators of the manifold dimensionality, coordinate system, and mapping functions that give much more accurate data compression and reconstruction than linear models, while using fewer dimensions. On the theoretical side, we have shown that there are classes of extrinsically curved manifolds that can be recovered from finite samples with perfect (not approximate) isometry.

Future Direction: We are investigating stable methods for very large data sets (10^6 points or more) and real-time methods for integration into signal-processing pipelines.

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

Lab: MERL Research Lab
Project Type: Research

Intelligent Multi-Dimensional Data Summarization



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

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

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

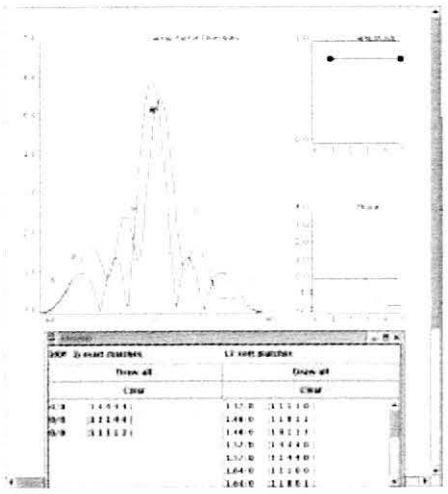
Collaboration: This project is joint work with the Data Management Technology Department at Johosoken.

Future Direction: This project has transitioned to our colleagues at Johosoken as they are better-suited to integrate the prototype with existing MELCO tools. They will also add new features to increase the usability of the initial prototype tool.

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

Lab: MERL Technology 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: FY04 efforts focused on the design of phased-antenna arrays with non-uniform spacing. We have designed and developed a novel interface that combines Overview and Detail components to provide access to data at different levels of abstractions. The Overview pane supports visual query techniques across a series of linked views of the two-dimensional data. Antenna performance data (e.g., radiation patterns) and other data sets involving linear ordered sequences (e.g., finance, weather, census data) are especially well-suited for visual query techniques; the graphical nature of the input and output methods support users in their information-seeking tasks, in this case designing an antenna with particular performance characteristics. After narrowing the search, the Detail pane provides spreadsheet access to the antenna parameters and performance characteristics, along with a graphical editor for exploring trade-offs in antenna element placement.

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 use machine learning techniques to explore the space of sparse phased-array antennas. Our goal is to find a guiding/unifying principle for designing such arrays that have low sidelobe levels.

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

Lab: MERL Technology Lab
Project Type: Research

Color Figures

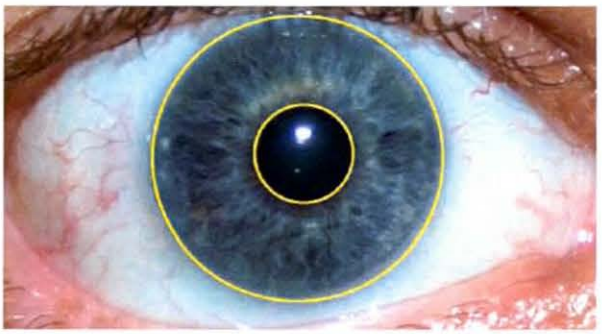


Figure 1 – see page 52

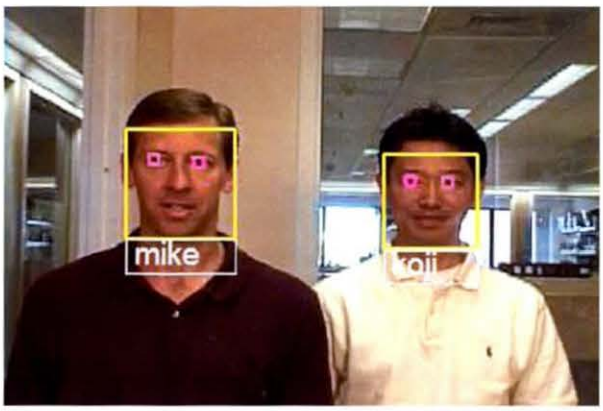


Figure 2 – see page 53

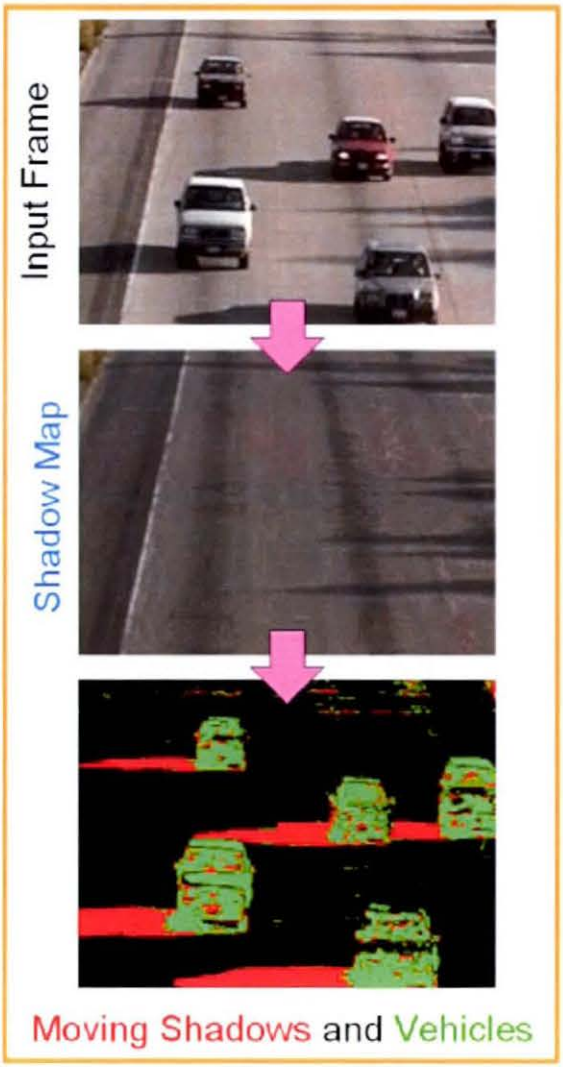


Figure 3 – see page 59

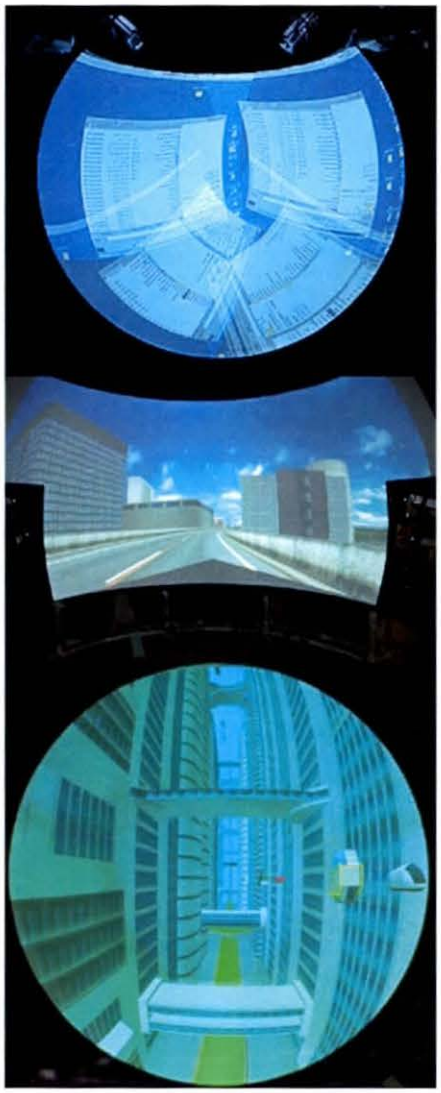


Figure 4 – see page 119

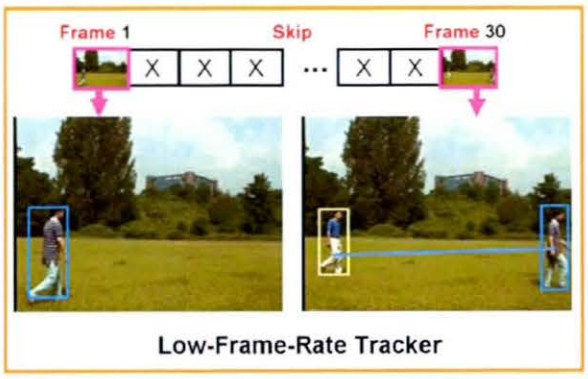
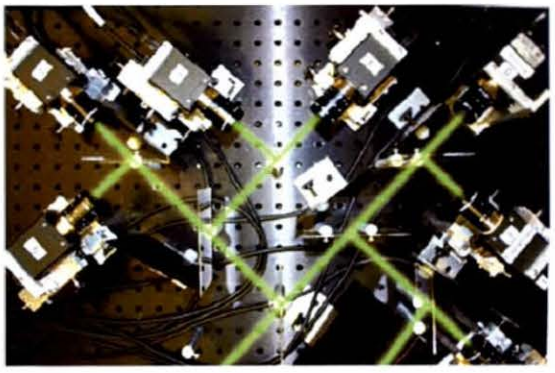


Figure 5 – see page 60



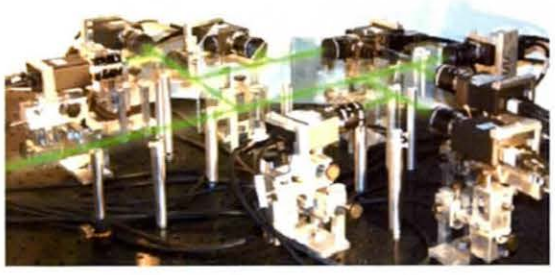


Figure 6 – see page 67

Figure 7 – see page 80

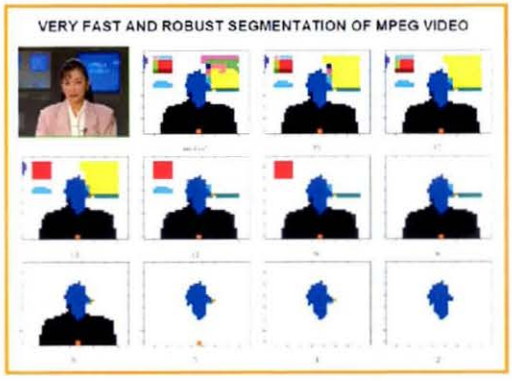


Figure 8 – see page 104

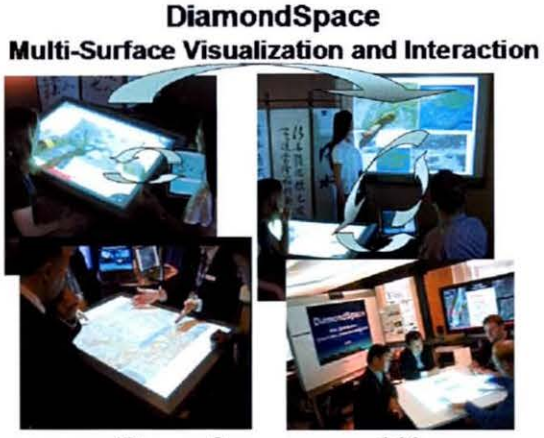


Figure 9 – see page 112

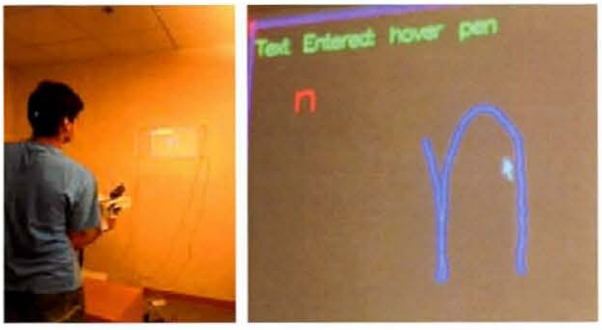


Figure 10 – see page 118