Mitsubishi Electric Research Laboratories (MERL)

Annual Report

July 2006 through June 2007

TR2007-00

(Published November 2007)

Welcome to Mitsubishi Electric Research Laboratories (MERL), the North American corporate R&D arm of Mitsubishi Electric Corporation. 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, Janet McAndless, Richard C. Waters

Table of Contents

Mitsubishi Electric Research Laboratories

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

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

- Generating new technology in areas of importance to MELCO.
- Significantly impacting MELCO's business—using our technical expertise in partnership with organizations in MELCO to produce new and improved products.

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.
- To be the prime source of new technology for MELCO in our areas of expertise.

MERL focuses on five principal technology sectors:

Digital Communications - featuring wired & wireless transmission technology & networking.
Multimedia – featuring speech interfaces and the encoding, decoding, & analysis of video.
Sensor and Data Systems – featuring HCI, data analysis, & equipment condition monitoring.
Imaging – featuring computer vision algorithms and the observation of people in images.
Mechatronics – featuring advanced simulation and control of factory automation systems.

An Algorithms group supports all five sectors, developing fundamental algorithms.

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.

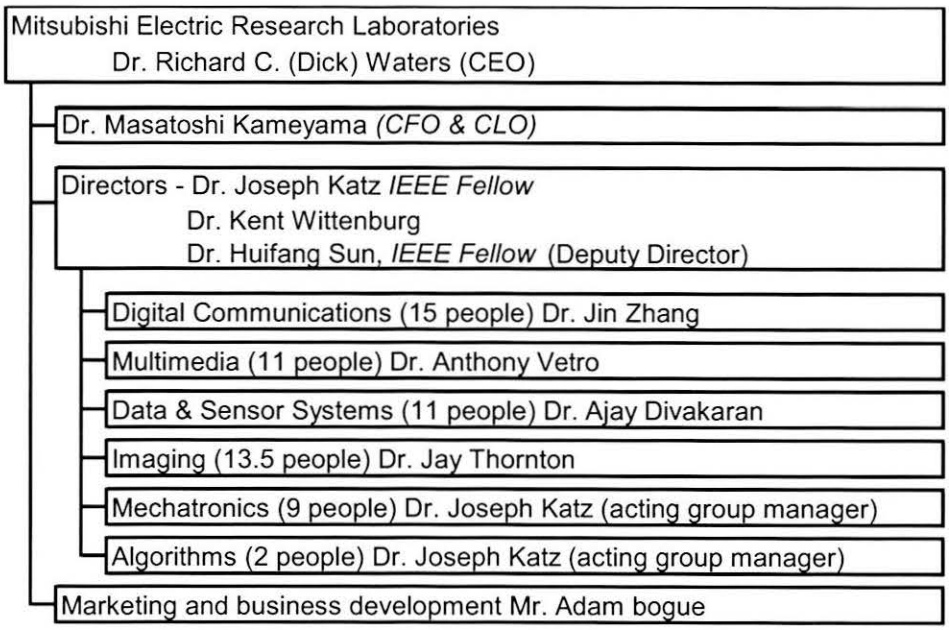
MERL's output ranges from papers and patents, through proof-of-concept hardware and software prototypes, to modules for industry-first products. The headquarters operation includes a small marketing and business development department to help realize the full market potential of our work and an in-house 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 is organized as six groups centered on technology areas, which collaborate closely to achieve groundbreaking results. The six members of the top management team work closely together, guiding all aspects of MERL's operation.



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 December 1999, he became CEO of MERL as a whole. In addition to his duties at MERL, Dick is currently a member of the board of directors of the Computing Research Association.

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

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

Joseph Katz *Ph.D., California Institute of Technology, 1981*
Vice President & Director

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

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

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

After four years as a Professor at Fairleigh Dickinson University, Huifang Sun moved to the Sarnoff Research Laboratory in 1990 becoming Technology Leader for Digital Video Communication. In 1995, Huifang joined MERL as the leader of our video efforts, becoming a Deputy Lab Director in 1997. In recognition of his productive career in video processing Huifang was made an IEEE Fellow in 2001. He was made a MERL Research Fellow in 2003.

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

Adam Bogue had 15 years of industry experience before joining MERL. This included 3 years at GenRad Inc and 7 years at Active Control eXperts Inc. beginning as Director of Sales and Marketing and ending as Vice President, Core and New Business Unit. Adam came to MERL in June of 2000 to lead our Marketing and Business Development effort.

Mitsubishi Electric

Number 196 on Fortune magazine's most recent list of the world's 500 largest corporations, Mitsubishi Electric Corporation has $33 billion in annual sales, $1 billion in profits, and nearly 100,000 employees in 35 countries.

MELCO is composed of a wide range of operations. The business units with sales of $1 billion or more are listed below.

Mitsubishi Electric
Diversified Electrical and Electronics Manufacturer

Public Utility Systems Government Systems, Transportation	(Kobe, Itami)
Energy & Industrial Systems Power Equipment, Plant Control	(Kobe, Nagasaki)
Building Systems Elevators, Escalators, Building Monitoring	(Inazawa)
Electronic Systems Satellites, Radar, Military Systems	(Kamakura, Itami)
Communication Systems Wired & Wireless Communications, Cell Phones	(Itami)
Living Environment & Digital Media Equipment Air Conditioners, Refrigerators, TVs, DVDs, LCD Projectors	(Shizuoka, Kyoto)
Factory Automation Systems Programmable Logic Controllers, Industrial Machine Tools	(Nagoya)
Automotive Equipment Alternators, Engine Controllers, Car Stereos, Car Navigation	(Himeji, Sanda)
Information Systems & Network Services IT Systems, Servers	(Tokyo, Kamakura)
Semiconductor & Device Optical and Radio Frequency Semiconductors, Power Devices	(Kita Itami, Fukuoka)

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 UFJ Financial Group, the Mitsubishi Corporation, Mitsubishi Heavy Industries, Mitsubishi Chemical Holdings, and Mitsubishi Motors, (all five of which are also on the Fortune Global 500 list—Numbers 118, 146, 257, 317, & 382 respectively). They have shared roots in 19th century Japan; however, these companies have been separate for many years and Mitsubishi Electric Corporation has been separate from all of them since Mitsubishi Electric's founding in 1921.

Mitsubishi Electric US Operations

Approximately 10% of MELCO's sales are in North America and many of MELCO's business units have North American subsidiaries. MERL seeks to work directly with these subsidiaries, particularly when they have substantial local design and manufacturing as well as sales.

The US operations with sales of $100 million or more are listed in order of revenue.

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

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

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

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

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

Mitsubishi Electric Corporate R&D

Number 31 on Intellectual Property Today's most recent list of the top US patent winners, 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 and is by itself number 238 on the list of top patent winners.

Corporate R&D
Headquarters: Dr. K. Kyuma (Director), Mr. A. Morita (GM), 18 people (Tokyo)
Managing Mitsubishi Electric's R&D

- **Advanced Technology R&D Center (ATC)**
  Dr. S. Yamamoto (GM), 900 people (Itami)
  Materials, Semiconductor Devices, Electrical & Mechanical Engineering
- **Information Technology R&D Center (ITC)**
  Mr. R. Nishii (GM), 800 People (Ofuna)
  Information Systems, Communications, Opto-Electronics
- **Industrial Design Center (IDC)**
  Mr. M. Hara (GM), 100 people (Ofuna)
  Industrial Design, Usability Studies
- **Mitsubishi Electric Research Laboratories, Inc. (MERL)**
  Dr. R. Waters (CEO), 82 people (MA)
  Computer Vision, Communications, Video coding, Sensor networks
- **Mitsubishi Electric Information Technology Centre Europe, B.V. (ITE)**
  Dr. K. Oshima (CEO), 60 people (France & England)
  Wireless Communications, Digital Audio & Video

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, 2006 through June 30, 2007.

Current Technical Society Fellows and Senior Members

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

A further thirteen (23%) of MERL's technical staff are Senior Members of the Institute of Electrical and Electronic Engineers (IEEE).

Best Paper Awards

Seltzer, M.L., Raj, B., and Stern, R.M., "Likelihood-Maximizing Beamforming for Robust Hands-Free Speech Recognition", *IEEE TSAP.*, Vol. 12, no. 5. (2006). (Best Young Author paper award from the IEEE Signal Processing Society.)

Otsuka, I., Suginohara, H., Kusunoki, Y.,and Divakaran, A., "Audio-Visual Features for a Personal Video Recorder", *IEEE Transactions on Consumer Electronics*, Vol. 53, Issue 1 (2007).

Tuzel, O., Porikli, F., and Meer, P., "Human Detection via Classification on Riemannian Manifolds", *Proceedings of the 2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'07)*, June 2007. (2nd Best Paper Award out of 1300+ papers.)

Technology Awards from Outside Organizations

In January 2006, Andreas Molisch was elected an IEEE Distinguished Lecturer.

In August 2006, Paul Beardsley and Fatih Porikli won the R&D100 award for the Heli-Tele project led by Maehara-san of ITC.

In December 2006, Bill Yerazunis was selected for inclusion in NetworkWorld's 50 Most Powerful People in Computer Networking, listed in the "Net Security" category.

In May 2007, Huifang Sun received the award of outstanding performance as Associate Editor of the IEEE Transactions on Circuits and Systems for Video Technology in ISCAS2007.

Awards from Mitsubishi Electric

In June 2006, MERL staff received two Corporate R&D Valuable Invention awards for patents on video summarization.

In June 2006, MERL staff in collaboration with staff from MELCO's ATC research laboratory received an R&D award for their work on "Video Summarization Technology".

In May 2007, MERL staff in collaboration with staff from MELCO's ITC research laboratory received a Corporate R&D award for their "contribution to mobile communication business via international standardization" on WiMAX and 3GPP.

In June 2007, MERL staff received a Corporate R&D award for their "outstanding contributions to global standardization activities on mobile communication technologies".

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.

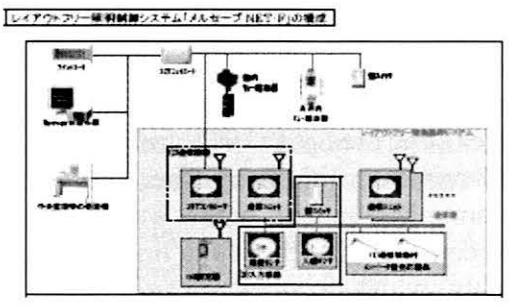
A large part of MELCO's business is in the form of large custom systems for business or government. MERL has contributed components to a number of such systems. For these components, we take the date at which the first system using it was delivered to the customer as the date of impact.

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.

ZigBee Stack for Layout-Free Light Control System



On June 22, 2006, The President of Mitsubishi Electric Corporation, announced its Layout-Free Lighting Control System on the market for sale starting in August 2006. The system is called MELSAVE NET.F and it consists of inverter lighting instruments with communication capability, area controller, communication units, lighting sensors and motion sensors, switches, and lighting set-up device. The system uses ZigBee technology to configure lighting devices freely without any limitation from the layout. The lighting devices can be turned on or off and adjusted one by one in a whole area. In addition, this system will save energy and control a whole system more effectively. For a building with 20,000 square

meters (200,000 square feet) the cost of the system starts at $300K. MERL developed the ZigBee technology and its stack, especially the network layer protocol (NWK), application sublayer (APS) and ZigBee device object (ZDO), which is the enabling technology for this world leading layout free lighting control system.

ZigBee has many other potential applications, such as home automation, building automation, automatic meter reading, equipment condition monitoring, and environment sensing and control. MERL is working with MELCO to apply ZigBee technology to several other business areas. Some new system products including ZigBee are expected to come in the near future.

Details: In October 2002, MELCO became one of the founding members of the ZigBee Alliance, which was established to develop the first standard for low power, low cost, short-range multi-hop wireless ad hoc wireless network for variety industry applications. Shortly after that, MERL was requested by MELCO to support their ZigBee activity and make technical contributions towards the world first ad hoc wireless network standard.

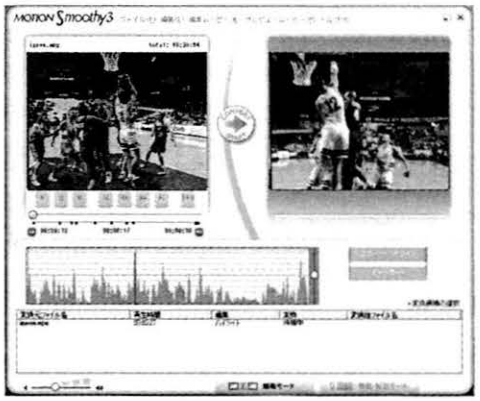
From December 2002 to December 2004, MERL focused on the ad hoc network layer protocol development and become one of pioneer members to make significant contributions to the standard. For example, due to its energy efficiency, MERL's channel quality indicator based routing protocol was included in the specification. Using this protocol, the routing path is selected not just based on number of hops or some other merits, but based on the channel quality so that the overall energy consumption of the network is minimized. In December 2004, version 1.0 the of ZigBee specification was published. Since then, MERL was involved in the development of the enhanced ZigBee specification version 2.0 to promote ZigBee applications in HVAC systems, industry plant monitoring, and other areas.

In 2004 MERL pioneered the development of a ZigBee stack in collaboration with Renesas. Within one year, MERL completed the development of the ZigBee network layer protocol software and delivered to MELCO and Renesas. In 2005, MERL continued the stack development and finished APS and ZDO as well as some basic application functions. By the end of 2005, MERL's ZigBee stack software successfully integrated with Renesas hardware platform, passed all necessary interoperability tests and got ZigBee v1.0 certified. Meantime, MERL and MELCO developed various ZigBee demo systems to demonstrate promising applications. All of this prepared necessary technology for introducing new products such as Layout-Free Lighting Control System (LF-LCS).

The most important features of ZigBee technology include self-organizing, self configuration network, built-in discovery of devices and services, self-healing, and reliable communication. After the LF-LCS network is installed, all those communication units (ZigBee nodes) will be automatically connected and form a wireless network. Although the communication range between any two nodes is short (10m-100m), with multi-hop routing, the network work can cover large area with large number of nodes. Therefore, in this LF-LCS, the setting node can communicate with many other nodes one by one without moving close to one particular communication unit. In this way, all the lighting devices in the network can be freely controlled and configured without any limitation from the physical allocation. Thanks to its self-healing capability, when one of the nodes is out of order, for example, it runs out of the battery, the rest of network will continue to operate normally. After it is recharged, it can re-join the network automatically. This is why ZigBee is a perfect solution for LF-LCS and many other applications.

According to the press release, MELCO is planning to ship 15 LF-LCS systems in near future. Since 2006, MERL has been working with MELCO to develop ZigBee middleware and enhance the ZigBee stack. With the mass production of the ZigBee chip, the cost of the ZigBee will dramatically drop. Many more LF-LCS systems and other application systems are expected on the market in 2008 and beyond.

Sports Highlights Playback for D903i Cell Phone



In September 2006, MELCO began shipping a new cell phone for NTT-DoCoMo (model D903i). This cell phone has companion software that contains an interface concept provided by MERL. For sports programming in particular, MERL's "intelligent fast forwarding" makes it easy to scan through recorded content, skipping from one key play to the next.

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

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

Building on this foundation, MERL did speculative work on automatic video summarization starting in 2000 and continuing through 2001 and into 2002. The general goal of video summarization is to locate a small subset of a video that can serve as a summary of the rest.

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

MERL demonstrated its summarization work to MELCO people on a number of occasions staring in mid-2001. In mid-2001, MERL's prototype was demonstrated to a group of people at MELCO who were designing DVD recorders and generated considerable interest from them.

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

Starting in the last half of 2001, the DVD recorder group began to push the process forward to the look and feel of the prototype. In late 2002, they presented MERL's demo to the home entertainment products business group and started to get buy-in from them.

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

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

MERL's on-site collaboration with Japanese engineers revealed that the percentage of a characteristic mixture of the commentator's excited speech and cheering, in a sliding window centered at the current point in time, is a good indicator of the interest level of the program around the point. The summarization mechanism consists of playing back only the parts that exceed a certain specified interest level. The final algorithm therefore consists of audio classification carried out on the incoming audio stream in real-time followed by a percentage calculation that yields the "interest level." The graph of "interestingness" vs. time enables the user to modify the length of the summary by moving the aforementioned threshold up or down. The technique works well across a wide variety of sports content, as noted by several reviews in the Japanese press.

Close collaboration continued through 2004 and into 2005, leading to the joint creation of product-ready summarization code. This was included in the product released in late 2005.

The product has received critical acclaim in the Japanese press. The prestigious magazine HiVi rated it as the best buy in its category. It is the world's first mainstream PVR with such highlights playback capability.

In 2006, MELCO ported this functionality to the companion software of NTT DoCoMo's cell phone model 903i. The software extracts sports highlights from content recorded on the PC. The extracted highlights are in mp4 format and are transferred to the cell phone through a USB link. The cell phone is equipped to play the sports highlights with essentially the same interface as that of the DVD recorder.

In addition the new work on cell phones, MERL's contribution to MELCO's DVD recorder product continues. In September 2006, MELCO began shipping a new DVD recorder (model DVR-DV635), which includes MERL's "intelligent fast forwarding" implemented using purely the host embedded CPU unlike the previous model that employed an additional DSP. MERL succeeded by both adapting the core algorithm to reduce the computation and memory footprint, as well as by optimizing the software, thus achieving the functionality at no additional hardware cost.

In the future, we plan to extend our techniques to other genres of content to enable similar intelligent "trick play."

Contributions to the ZigBee Cluster Library

The ZigBee standard for short range communications includes a so-called "ZigBee Cluster Library" which defines a set of reusable device descriptions and attributes that can be used to develop advanced application profiles. MERL technology for thermostat and fan control was included in the Cluster Library (Document 075123r01ZB) at the end of 2006.

Since 2003, MERL has been involved in the creation of ZigBee Alliance Specifications. We initiated early attempts to draft Application Profiles in the domains of Industrial Plant Monitoring (IPM) and Heating, Ventilation, and Air Conditioning (HVAC), with MERL researchers acting as chairmen of these two groups. The IPM specification was balloted and approved by the ZigBee Alliance in 2005. However, subsequent releases of the underlying ZigBee protocol stack caused a fundamental change in the way that ZigBee handles specific devices. In particular, the approach was changed from profiles to a library of concepts that could be used by devices. Aspects of MERL's work (particularly on the HVAC profile) were subsequently combined into this library.

Face Recognition for Biometric Terminal

In January 2007, MELCO made the first installation of a new biometric access control system (ACS) for the Japanese domestic market, the "Integrated Face/Fingerprint Biometric Terminal OPG-FACE." The system consists of a client device (or devices) for user interaction—the "terminal"—and a server that does the heavy computation and stores the various biometric data bases. The terminals are capable of capturing both face images and finger print images. They also contain a keypad for user entry of PIN codes if desired. When installed, a terminal is mounted next to a door that requires controlled access. To gain access, a person must be recognized based on one (or both) of the biometrics. The novelty of the system is the inclusion of MERL's face recognition to allow faster and more convenient user interaction with the ACS.

Details: MERL's work on face recognition started in 2002. Face recognition was a natural follow-on to the face detection work done by MERL staff Paul Viola and Mike Jones and now widely imitated throughout the world. Although the face detection work was started outside MERL, we went on to make an improved algorithm for detecting faces in images, and then to the problem of recognizing those faces. All this work is distinguished because of its speed and its light weight implementation. The detection speed derives from two things: first, faces are detected by sequentially testing and rejecting image patches. Most patches are rejected right away, so the analysis can go on to consider the next patch. Second, each test is extremely simple and fast to compute because it is based on simple sums and differences of rectangular pixel areas in the patch. Exactly which tests ("rectangle filters") produce good results is determined by a machine learning selection algorithm called Adaboost. The same rectangle filters and learning

engine were applied to the face recognition problem by pairing face images into "same person" pairs and "different person" pairs. The learning engine selected filters that best classified pairs as same or different. Of course, before faces could be compared for recognition, they had to be first detected, and then aligned so that features were in more-or-less constant positions. This alignment was also accomplished using similar detectors trained to locate eyes, nose, corners of the mouth, etc.

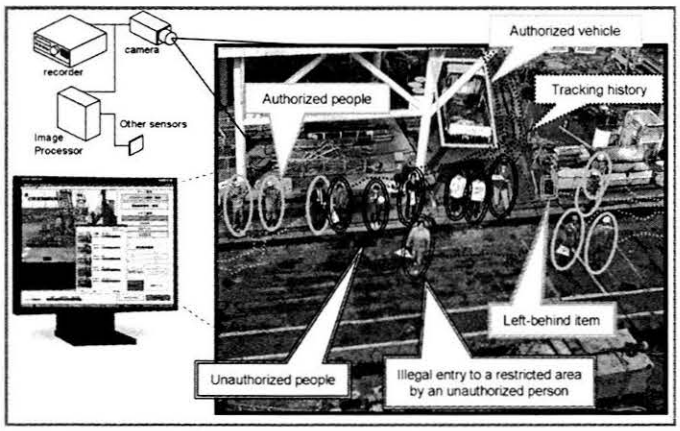
There are many applications of face recognition in MELCO businesses, so MERL was encouraged to benchmark and improve the performance of the method over several years. This effort became a close collaboration with MELCO domestic laboratories and ultimately resulted in error rate reductions by more than an order of magnitude. The process of continual improvement continues today as we strive to extend the range of successful face recognition to less controlled lighting and pose situations.

Human Tracking System Using RFID and Face Recognition

In the spring of 2007, Mitsubishi Electric began selling a human tracking system that uses both face recognition and RFID tags to keep track of the location of people in a secure facility. RFID is good for positive ID at specific checkpoints. Face recognition allows less formal verification between checkpoints that a person passing by is who the system expected and is indeed authorized to be in the location. MERL provided the face recognition algorithm used in the system.

Since 2005, various parts of Mitsubishi Electric have become increasingly interested in systems employing face recognition that build upon existing products and sales channels. The human tracking application grew out of existing business with some large, secure facilities, and it was natural to propose a system to raise the general security level. Development went on in 2005 and 2006, culminating in the product release in the spring of 2007. The face recognition algorithm is the same MERL algorithm that is used in the biometric terminal described above.

Object Tracking for Harbor Surveillance Systems



In February 2007, MELCO began installing the new harbor surveillance system that is built using high-end security DVR as a core (model TL-5000U). The integrated system contains other components such as PTZ cameras, information fusion and storage servers, video encoders, cable detectors, etc. to provide a complete solution for a wide spectrum of applications, ranging from high security areas to small businesses and residential houses. MERL's advanced detection ad tracking technologies make it possible to automatically find humans in variable frame rate video, track humans and vehicles, analyze their actions using motion histories, detect left-behind objects, and effectively control PTZ cameras.

MERL has been an award winning leader in computer vision area in the last decade and MELCO has a long history of video surveillance products from the world's first time-lapse video recorder to smart cameras. In 2000, MERL focused its related research on the understanding of people's and vehicles' movements in video with the goal of developing technology for surveillance and traffic products. This led to a wide range of technologies in the areas of motion detection, object tracking, and human detection.

Details: In 2001, MERL started automatic video object segmentation for consumer video to enable efficient encoding. This work naturally evolved into the background generation by motion detection for static camera setups in 2002, culminating in the development of a robust and accurate online learning method for static cameras that outperforms the conventional approaches proposed by MIT and other companies. MERL's initial work in this area has been later extended for larger surveillance systems that consist of multiple non-overlapping cameras in 2003.

Building on this foundation, MERL did speculative work on the low-frame-rate object tracking to achieve processing of multiple input videos on the same processor in 2004. The goal of object tracking is to extract discriminative information of moving objects that can empower automatic event detection.

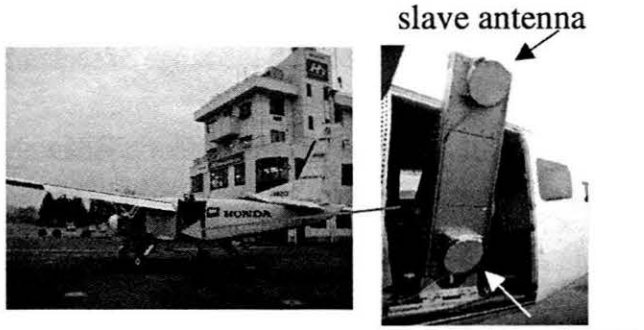
Everything began to come together in 2005, with a strong push toward productization. By the diligent efforts of MELCO team, MERL's object detection and tracking solutions found their way into DVR, access control and human recognition system, wide areas security system, and intelligent camera terminal. In 2006, a parking lot monitoring system was another application that wanted to feature MERL's detection technology in its upcoming versions.

Driven by the needs of the harbor surveillance system and by constantly improving on the core technologies, MERL has won one of the best paper awards at CVPR out of 1300 papers for human detection on manifold technique in 2007.

Technologies developed for this system also benefit other projects. For instance, the human detection using the classification on Riemannian manifolds is now a core solution for pedestrian detection for on-vehicle cameras. The background generation based people counting algorithm is being used in Misubishi Electric's large visual system. The shadow removal and tracking is integrated in the DSSS system for identification of illegal passes at toll gates on highways. The object boundary and pose tracking methods are being revised for tracking of organs in medical data, for extraction of player statistics in sports videos for a potential collaborator in US, for controlling of the airborne camera in a next generation Heli-tele.2 system, and for correlating the visual information in MERL's own ambient intelligence system.

Our detection and tracking methods are shown to be superior to our competitors in terms of accuracy as they employ competent image descriptors and precise learning techniques. We are currently developing ultra-fast and economical implementations that make use of the of-the-shelf components.

Phase Unwrapping for Interferometric SAR



In early 2007, MELCO delivered to a customer Digital Elevation Maps (DEMs) produced using a new Ku-band airborne Interferometric Synthetic Aperture Radar (InSAR) system. A DEM is an elevation (height) map of terrain, providing a height for sampled locations on the ground (in our case on a 50 cm x 50 cm grid) and can be used for a variety of purposes. An InSAR system uses two radar receivers, one of which is also a transmitter, separated in height and traveling on a moving platform such as an airplane, to create DEMs of terrain from long range and in any weather conditions. The new system includes a crucial computational software component from MERL for 2-D phase unwrapping. The system and its resulting DEMs demonstrated less than 50 cm mean absolute height error, which is believed to be unprecedented in InSAR imaging and substantially due to MERL's phase unwrapping algorithm.

MERL plans to continue development of the software, primarily to improve the speed of the algorithm, which may become important to MELCO as more and larger datasets are processed in the near future.

Details: Two-dimensional phase unwrapping is the critical computational step in InSAR and has received a number of recent treatments. In 1998, Constantini published a method based on network programming using a Minimum Cost Flow (MCF) algorithm. In 2000, MELCO's previous phase unwrapping algorithm was developed with MIT and was based on a Weighted Least Squares (WLS) algorithm. In 2001, Frey, Koetter, and Petrovic published an algorithm for phase unwrapping using "loopy" Belief Propagation (BP). In 2002, Dias and Leitao published the Z-Pi-M algorithm.

In the same time period, MERL had been actively involved in research on Belief Propagation and in 2002-2003 undertook a basic research project to accelerate the very time-consuming operations of 2-D BP networks by using a Graphical Processing Unit (GPU). MERL chose the Frey phase unwrapping algorithm as a test case and demonstrated a 30x acceleration of the BP-based phase-unwrapping computation using a GPU. Also during the same time period, MERL was assisting MELCO on a separate project for satellite-based 3-D modeling, including the use of satellite-based SAR.

In 2004, inspired by both the MERL projects, MELCO asked MERL to pursue research on phase unwrapping itself.

By March 2005, MERL had implemented in-house software for the existing WLS-based and BP-based algorithms and was able independently to verify that the Frey BP-based algorithm was superior in many ways to the WLS-based algorithm. Although, despite better results in most cases, the BP algorithm had one important flaw in that it did not guarantee a residue-free solution. "Residues" in 2-D phase unwrapping are points of inconsistency that disallow the

forming of a consistent phase solution throughout the pixel map. Removing all residues is one important requirement of a successful 2-D phase-unwrapping solution.

During this testing period, MERL undertook a research effort to understand the nature of existing algorithms and to design a better one. All existing algorithms had fundamental theoretical flaws. The MCF algorithm always produces a residue-free solution, but does not support a noise model. The WLS algorithm produces a residue-free solution and rejects noise statistically, but at the expense of large systematic local distortions caused the wide-scale smoothing inherent in the formulation. The Frey BP algorithm performs well locally, but does not guarantee a residue-free solution and does not use a noise model. The Z-Pi-M algorithm does not guarantee a residue-free solution.

By June 2005, MERL had developed a solution using a combination of BP and MCF. The BP part was based on a well-formulated factor graph that supports both a noise model and a terrain model. The MCF part guaranteed a residue-free solution at each step. The MERL BP-MCF algorithm was developed and tested over the Summer 2005 showing superior results to existing algorithms including a guaranteed residue-free solution, statistical noise and terrain models, and good performance in both low-noise and high-noise regimes.

The main drawback to the system was that the BP part of the implementation was large and slow. For a large realistic 100 MPixel dataset, the algorithm would require 400 GB memory and take 184 hours (8 days) to compute, which is probably impractical. This drawback led to the design of a more efficient Iterated Conditional Modes (ICM) approximation to the BP part of the algorithm and led to creation of the MERL ICM-MCF algorithm. It was believed that the results would be similar, but the same 100 MPixel dataset would require only 5 MB and 6.5 hours.

Between September 2005 and March 2006, a software engineering effort was made to convert the original research prototype into useable software that was efficient and could be maintained and tested more easily. This effort included converting certain parts of the software from floating-point to integer and replacing BP with ICM.

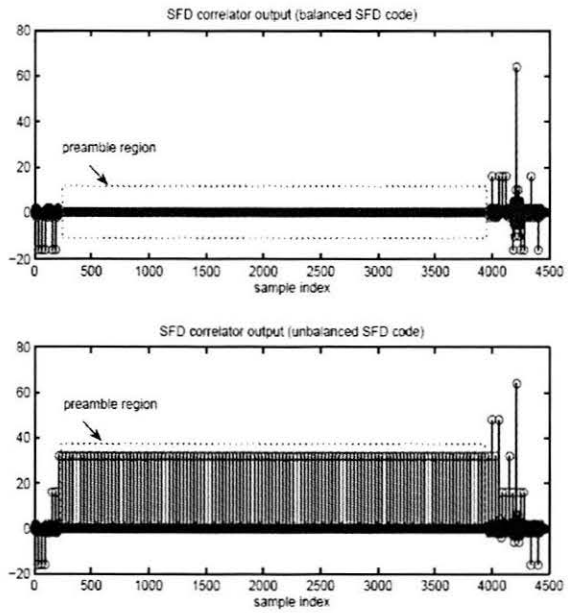
In April 2006, MERL internally released MERL ICM-MCF algorithm to be used for internal data processing. MERL received SAR data from a flight test made by MELCO earlier in the year, processed the results and delivered them to MELCO for analysis. The analysis found the results to be significantly superior to the existing phase unwrapping results. MERL made subsequent improvements in the accuracy of the algorithm and the performance of the software.

In March 2007, MELCO reported results of evaluating the latest software on new flight data collected on behalf of a customer. The results were compared against laser profiler data of the same area and were found to produce height errors under 50 cm, a very good result, believed by MELCO experts to be unprecedented and meeting the requirements of the customer.

In April 2007, researchers from MERL and MELCO collaborated on a report to be published in Japanese at IEICE and in English at the IGARSS 2007 conference in July in Barcelona.

MELCO anticipates new contracts in the near future which will require processing many more and larger datasets. To make this feasible, MERL intends to continue work in FY2007 primarily to make the software faster. Other planned improvements include refining the noise model of the algorithm and improving memory management and diagnostic reporting of the software.

Contributions to IEEE 802.15.4a UWB Personal Area Networks



In March 2007, the IEEE Standards Association officially approved the IEEE 802.15.4a standard, which provides an ultrawideband (UWB) – based physical layer for communications and ranging in Personal Area Networks (PANs). PANs have obtained great importance in recent years especially for sensor network applications and home automation; for example, the widely used ZigBee standard defines a good networking layer for such networks. The IEEE 802.15.4a standard is the first to provide explicit support for ranging, and also provides a modulation scheme that provides extremely energy-efficient communications.

MERL has been a key player in the development of the IEEE 802.15.4a standard, and has provided a number of important technologies. MERL technologies that were added during the past year are the "Start Frame Delimiter" (SFD) and a scaling factor for the range quality indication.

Details: The IEEE 802.15.4a standard defines an alternate physical layer for the IEEE 802.15.4 standard, also commonly known as low layers of ZigBee. While the original 15.4 standard used narrowband signaling, the 15.4a standard employs ultrawideband (UWB) signals, i.e., signals with a bandwidth of 500 MHz. This large bandwidth enables very precise ranging between devices; the range information between several devices can be converted into geolocation information. The use of UWB signaling also allows the design of low-complexity, energy-efficient communications approaches.

MERL has been a key player in the 802.15.4a standard ever since its inception in 2003. Members of MERL have played leading roles as subgroup chairmen and technical editors, over the years. Furthermore, a number of key technologies from MERL were included in the standard. After the official approval of the standard by the IEEE in March 2007, the first products from different manufacturers have been deployed.

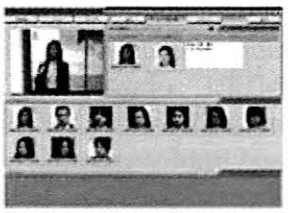
In order to achieve precision ranging, the 15.4a standard defines a preamble that is sent at the beginning of every packet. This preamble, which also serves for acquisition of the packet, has variable length, and repeats a certain pre-determined sequence many times. Since the length of the sequence is not known a priori, and since a receiver does not know which repetition of the sequence it is observing after having acquired the signal, it is necessary to clearly demark the end of the preamble, and the beginning of the actual payload data of a communications packet. This demarcation is achieved by means of the "Start Frame Delimiter" (SFD). The SFD is a unique sequence of bits that allows easy detection. With a well-designed SFD, the receiver correlates output during preamble part of the received packet would not have side lobes as illustrated in the figure, hence a better SFD detection performance.

Since both coherent and noncoherent receivers are envisioned in IEEE 802.15.4a systems, the SFD has to be designed in such a way that good correlation properties are retained with both types of receivers. The IEEE 802.15.4a supports a mandatory short SFD (8 symbols) for the default mode (1Mbps) and an optional long SFD (64 symbols) for the nominal low data rate (106Kbps). Both short and long SFD sequences in the IEEE 802.15.4a standard are provided by MERL.

A further contribution to the standard was the design of the feedback for ranging reliability information. In order to provide precise geolocation, a node has to combine the ranging information from different neighbors. The best location estimate is achieved when the node knows the reliability (probable error) of each ranging estimate; this information is communicated in a quantized form in a feedback packet. But for different applications, different orders of magnitude are most probable for the error – for example, for line-of-sight short-range communications, errors on the order of one millimeter are likely, while for long-range non-line-of-sight situations, one meter is a typical error. In order to accommodate those different ranges in a quantized feedback packet, we have introduced a scaling factor into the packet.

Both the SFD and the ranging scaling factor are now established part of the 802.15.4a standard. Research at MERL and MELCO is ongoing for exploiting the ranging capabilities of 15.4a in a variety of future MELCO products.

"Watch-List" DVR Software DX-PC55EXP



In the first quarter of 2007, MELCO released to customers of its Digital Video Recorders (e.g. the TL5000U) a new software suite, the DX-PC55EXP, that runs on a PC and allows automatic scanning of the stored video for people on a specified "watch-list."

Details: The DX-PC55EXP is the latest version of software that started with the PC55PRO. The goal of the PC55 series is to allow more convenient access to video stored on the DVR. This first version was introduced in the February 2006 and included video search based on face detection and on tracking—another MERL technology (see the MERL 2006 annual report for more information about the the PC55PRO.)

Detection and counting of faces can narrow down the possible video an operator might have to search through, but face recognition can be even more specific about exactly what should be retrieved: someone who "looks like this." In 2004 MERL made a system for video retrieval based on face recognition that we called "Face Browser" (see MERL 2004 annual report.) The idea of Face Browser was that, as faces were detected and recognized in real time, meta-data characterizing that person was recorded along with the video. At a later time, a user could select a person by clicking on a gallery of faces and then automatically jump to points in the video containing that person. Again our MELCO colleagues took over the development that resulted in the PC55EXP release this spring. Although Face Browser demonstrated the possibility, face recognition for identifying people in video had to substantially mature before it was ready to become a MELCO product. Recognition in free video is much more challenging than recognition for access control where a person is cooperating with the system in order to gain entry. In the case of identification in free video, lighting, pose, and expression all vary in complex ways. This complexity promises to keep MERL and MELCO researchers challenged in the near future and is the principle focus of our current face recognition work.

Network Key Management for ZigBee Chip

ZigBee is a wireless communications standard intended for home automation and sensor network applications. ZigBee builds upon the IEEE 802.15.4 standard which defines low-rate (<250kbps) physical and MAC (Medium Access Control) layers. This chip is based on a microprocessor running software written by MERL. MERL provided the network layer (NWK), application support sub-layer (APS), and ZigBee Device Object layer (ZDO). During 2006, MERL completed the ZigBee stack by implementing APL Security. Renesas first shipped a chip implementing the ZigBee communications protocol in 2005. APL security was incorporated into the chip in the first quarter of 2007. Renesas also sells a ZigBee Demo Kit (RZB-ZMD16C-ZDK) for both 900-MHz and 2.4GHz.

Besides completing the last piece of ZigBee stack with the implementation of APL Security, MERL has worked on and continues to work on ZigBee applications with projects like Image Over Zigbee, Voice Over ZigBee and ZigBee Middleware.

ZigBee work at MERL started in 2002 with participation in the ZigBee Alliance standard. MERL started implementing the NWK layer in 2004 and by 2005 had completed the basic stack implementation of NWK, APS and ZDO, which was delivered to Renesas and MELCO. In 2006 MERL completed the ZigBee stack by implementing APL Security. ZigBee APL Security builds upon the security of IEEE 802.15.4, providing an architecture for management of keys and implementation of security policies. The implementation of APL Security for the application layer includes methods for key establishment, key transport, frame protection, and device management.

Key Establishment: ZigBee uses 128-bit symmetric keys. A key can be associated with either the network (NWK and MAC layers) or with a link (application layer). Establishment of link keys is based on a master key which controls link key correspondence between two devices.

Key Transport: Key distribution is an important function in a secure network. A secure ZigBee network will designate one device, called the trust center, that all other devices trust for the distribution of security keys. APL Security includes services for transport of keys between devices and the trust center.

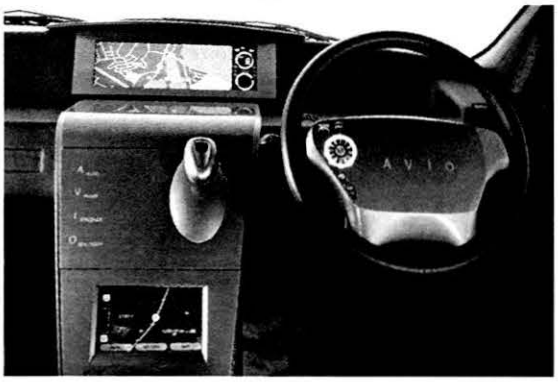
Frame Protection: The implementation of security for ZigBee includes protection of outgoing frames with encryption using the master, network and link keys which have been established and transported between devices. Frame protection is implemented at the APS layer and also includes decrypting of incoming frames.

Device Management: The APL security is responsible for maintaining the security policies of a device and in the network. This includes the addition and removal of devices as well as the propagation of changes in devices or the trust center itself.

The implementation of APL Security completes MERL development of the ZigBee stack producing a two-fold business impact. While Renesas sells MERL's ZigBee stack

implementation with their MCU, MERL is also working to help MELCO business units apply ZigBee technology to their customer applications for a technically competitive advantage.

Spoken Interface Testing & Support for US & EU Car Navigation



In the second quarter of 2007, MELCO started shipping production ready voice enabled navigation head units. This is the first production product to come out of MERL's wide ranging support for the development of voice enabled automotive electronics for markets outside of Japan. MERL has provided testing and evaluation of speech recognition and synthesis engines, voice user interface translation and design, and customer support among other services.

The MERL speech team's primary focus has been in developing easier to use and safer voice interfaces for mobile applications. SpokenQuery (SQ) enhances the usability of commercial, automatic speech recognition (ASR) engines. Using processing of ASR output, SQ allows for natural, free-form spoken requests for information. Using SpokenQuery is like using Google to locate something you need, but you say the descriptive words instead of typing them.

Details: This project was started in response to a request in 2001 to support development of voice interfaces for automotive navigation systems outside of Japan. The MERL speech team's experience in developing products for these markets provided key expertise. The project started with the testing of commercial speech recognition and synthesis engines and giving voice user interface design advice.

In September 2004 MERL sent a team member to Europe for several months to help with product specification and support communication with the automobile manufacturer. MERL has provided ongoing advice on the voice interface specification in several languages.

Technical Staff

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

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

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

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

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

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

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

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

John C. Barnwell III
Research Assistant

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

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

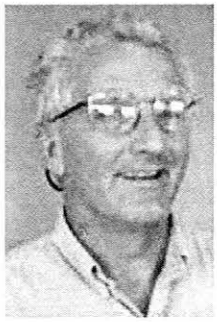
Beardsley's thesis work was on applications of projective geometry to 3D recovery from images. His current focus is on 3D scanning, stereo vision for surveillance particularly looking at depth and 3D connectivity cues to aid segmentation of individuals in a crowd, and hand-held projectors together with 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

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

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, recovery of non-rigid 3D shape from ordinary video, and an entropy optimization framework for learning. He has received best paper awards in computer vision (CVPR2001) and scheduling (ICAPS2003).

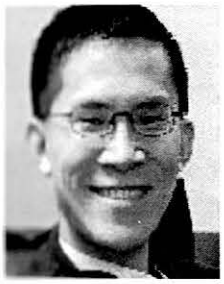


Dirk Brinkman *J.D., Suffolk University Law School, 1990*
Patent Counsel

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

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

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



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

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

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

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

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

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*
Group Manager Data & Sensor Systems

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

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

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

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

Esenther enjoys human-computer interaction (HCI) design, distributed software development, graphical user interfaces and Internet technologies. His recent work has focused on touch applications that support multiple concurrent users (think multiple mice), rapid image presentation for video browsing, and instant co-browsing (lightweight real-time distributed collaboration using unmodified web browsers).

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

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.

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

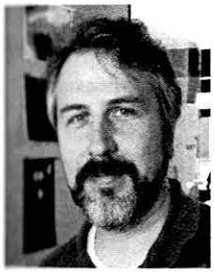
Guo worked at Waterloo Maple for a year and a half as a software developer before joining MERL in 1998. He primary research interests include home networks, digital broadcasting, and wireless computing.

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

Harsham joined MERL in 2001 to pursue interests in speech interfaces and speech-centric devices. Previously, 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

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

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*
Senior Principal Technical Staff

Jones joined MERL in 2001 after 4 years at the Digital/Compaq Cambridge Research Laboratory. His main area of interest is computer vision. He is particularly interested in using machine-learning approaches for solving computer vision problems. He has focused on algorithms for detecting and analyzing people in images and video such as face detection, skin detection and facial analysis using morph able models.

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

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.

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

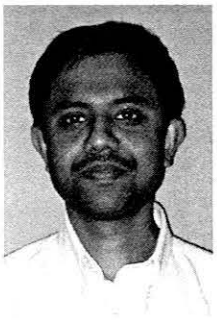
Leigh's research interests range from electronic hardware and embedded systems to signal processing, RF and communications. Before coming to MERL, he worked on the Harvard University/Planetary Society Billion-channel ExtraTerrestrial Assay (Project BETA), a search for microwave signals from extraterrestrial civilizations (SETI). His current research includes DiamondTouch multi-user touch technology and sensor networks.

Shan Liu *Ph.D., University of Southern California, 2002*
Member Technical Staff

Shan Liu joined MERL in 2006 after spending three years at Sony US Research Labs, where she participated in the development of PlayStation3. Shan has broad interests on multimedia data compression, processing, networking, analysis and management, with emphases on advanced video/image coding and processing technologies.

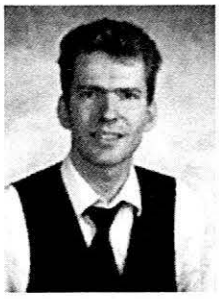
Janet McAndless
Technical Associate

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



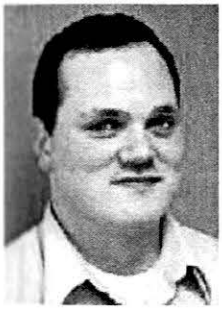
Neelesh B. Mehta *Ph.D., California Institute of Technology, 2001*
Principal Technical Staff

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.



Andreas F. Molisch *Ph.D., Technical University Vienna, 1994*
Distinguished Technical Staff, Chief Architect Wireless Standards

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

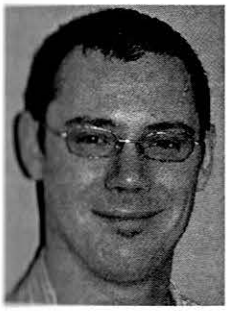


Clifton D. Mueller *J.D., Boston College Law School, 2005*
Patent Attorney

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

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

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 and works on interference cancellation and 3G CDMA systems. His current interests include speech processing and wireless communications.



Barton Nicholls *Northeastern University*
Systems & Network Administrator

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*
Team Leader Analytics

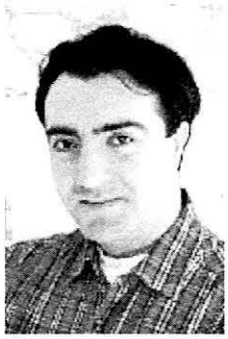
Nikovski's research is focused on algorithms for reasoning, planning, and learning with probabilistic models. His current work is on the application of such algorithms to hard transportation problems such as group elevator control and traffic prediction. He also has varied interests in the field of data mining.

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

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.

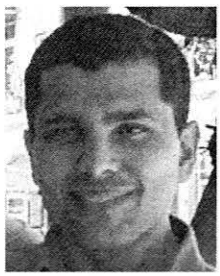
Ronald N. Perry *B.Sc., Bucknell University, 1981*
Distinguished Research Scientist

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, 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.



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

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).

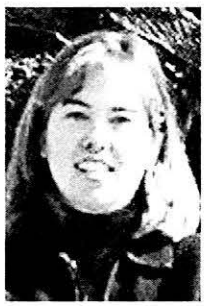


Bhiksha Raj *Ph.D., Carnegie Mellon University, 2000*
Research Scientist

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

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.



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

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

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.

Masashi Saito *Ph.D., Osaka University, 2006*
Senior Principal Technical Staff

Masashi Saito received his Ph.D. in Computer Science specializing in Distributed Systems. Before joining MERL in 2006, he worked as a Senior Software Engineer at MELCO's Information Technology R&D Center doing research on operating systems, Internet protocols and distributed systems. His interests include wireless networking, algorithms, software development and Internet services.

Bent K. Schmidt-Nielsen *B.S. University of CA at San Diego, 1971*
Team Leader, Speech

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

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

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

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

Kuntal Sengupta *Ph.D., Ohio State University, 1996*
Principal Technical Staff

Before joining MERL in 2006, Sengupta worked at AuthenTec, a leading fingerprint sensor company on the design and implementation of indexing, image reconstruction, navigation, template compression and anti-spoofing functionalities. His software implementations made itin to several million cell phones and PCs that AuthenTec shipped.

Chia Shen *Ph.D., University of Massachusetts, 1992*
Senior Research Scientist

Shen's current research focuses on shared interactive surfaces. She led the MidART project, which MELCO has incorporated into several distributed industrial plant control systems. MidART is a real-time middle-ware for applications where humans need to interact, control and monitor instruments and devices in a network environment through computer interfaces.

Samuel E. Shipman *M.Sc., Carnegie Mellon University, 1985*
Principal Technical Staff

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.

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

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.

Zhifeng (Jeffrey) Tao *Ph.D., Polytechnic University, 2006*
Member Technical Staff

Zhifeng Tao joined MERL in September 2006. His research interests include wireless networking, medium access control, quality of service, cooperative communications and analytical modeling. He has been an active participant in IEEE 802.11n and 802.11s standardization since 2004, and is currently involved in developing the IEEE 802.16j and 802.16m standards.

Koon Hoo Teo *Ph.D., University of Alberta 1990*
Team Leader BB Mobile Systems

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

Jay E. Thornton *Ph.D., University of Michigan, 1982*
Group Manager Computer Vision Applications

Thornton worked at Polaroid Corporation for many years, first in the Vision Research Laboratory and then as manager of the Image Science Laboratory. There, he worked on problems in color reproduction, image quality and image processing. He joined MERL in 2002 as Manager of the Computer Human Observation project, focusing on the computer vision problems that arise when computers analyze, measure, count, detect, and recognize people.

Jeroen van Baar *M.Sc., Delft University of Technology, 1998*
Principal Technical Staff

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*
Group Manager Digital Video Technology & Systems

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 products, including MPEG-2/4 transcoding for surveillance, post-filtering for artifact reduction and video down-decoding for a low-cost DTV receiver chip.

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

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

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

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

Jonathan Westhues *B.A., Sc., University of Waterloo, 2005*
Research Associate

Jonathan Westhues has experience in large-scale software, electronics, and signal processing for wireless communications. He was previously employed by Research In Motion, where he wrote software for their GSM products. In addition to this, he has worked on several projects relating to the control of industrial systems and machine tools. His interests at MERL include practical hardware for sensor networks, and radio electronics.

Kevin W. Wilson *Ph.D., Massachusetts Institute of Technology, 2006*
Member Technical Staff

For his doctoral thesis, Kevin Wilson incorporated aspects of the psychoacoustics of the precedence effect into an algorithm for computerized audio source localization. He is currently working on applications of signal processing and machine learning to audio processing, video processing, and equipment condition monitoring.

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

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.

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

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

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

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

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

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

Jinyun Zhang *Ph.D., University of Ottawa, 1991*
Group Manager Digital Communication & Networking

Zhang manages MTL's digital communication and networking group. Before joining MERL in 2001, She worked for Nortel Networks for 10 years where she held engineering and management positions in the areas of VLSI design, advanced wireless technology development and wireless & optical networks. She has a broad technical background, specializing in system design and real-time embedded software for wireless communications.

Recent Major Publications

The following lists the 106 major publications by members of the MERL staff over the past year. (This is an average of more than 1.7 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 37 (35.6%) publications that were subject to highly stringent selection criteria where they were published. Some venues (such as major journals and certain key conferences) are very selective in what they publish and some (such as workshops and many conferences) are not. There are good reasons to publish something in a non-selective venue, the most important of which being that a given workshop or conference may be the best place at which to expose a particular piece of work to the scientific community. However, the mere appearance of a piece of work in a non-selective venue says little if anything about the quality of the work in the eyes of the scientific community.

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

Draper, S.C.; Yedidia, J.S.; Wang, Y., "ML Decoding via Mixed-Integer Adaptive Linear Programming", *IEEE International Symposium on Information Theory (ISIT)*, June 2007 (TR2007-022)

Wu, J.; Mehta, N.B.; Molisch, A.F.; Zhang, J.Y., "Spectral Efficiency of Channel-Aware Schedulers in Non-identical Composite Links with Interference", *IEEE International Conference on Communications (ICC)*, June 2007 (TR2007-029)

Nikovski, D.; Jain, A., "Memory-Based Algorithms for Abrupt Change Detection in Sensor Data Streams", *IEEE International Conference on Industrial Informatics,* ISSN: 1935-4576, Vol. 1, pp. 547-552, June 2007

Zhang, H.; Mehta, N.B.; Molisch, A.F.; Zhang, J.Y.; Dai, H., "On the Fundamentally Asynchronous Nature of Interference in Cooperative Base Station Systems", *IEEE International Conference on Communications (ICC)*, June 2007 (TR2007-030)

Draper, S.C.; Martinian, E., "Compound Conditional Source Coding, Slepian-Wolf List Decoding and Applications to Media Coding", *IEEE International Symposium on Information Theory (ISIT)*, June 2007 (TR2007-023)

* Agrawal, A.; Raskar, R., "Resolving Objects at Higher Resolution from a Single Motion-blurred Image", *IEEE Computer Society Conference on Computer Vision & Pattern Recognition (CVPR)*, June 2007 (TR2007-036)

* Tuzel, O.; Porikli, F.M.; Meer, P., "Human Detection via Classification on Riemannian Manifolds", *IEEE Computer Society Conference on Computer Vision & Pattern Recognition (CVPR)*, June 2007 (TR2007-046)

* Morris, N.; Avidan, S.; Matusik, W.; Pfister, H., "Statistics of Infrared Images", *IEEE Computer Society Conference on Computer Vision & Pattern Recognition (CVPR)*, June 2007 (TR2007-020)

* Palanki, R.; Fossorier, M.P.C.; Yedidia, J.S., "Iterative Decoding of Multiple-Step Majority Logic Decodable Codes", *IEEE Transactions on Communications*, ISSN: 0090-6778, Vol. 55, Issue 6, pp. 1099-1102, June 2007 (TR2007-038)

Xin, J.; Li, J.; Vetro, A.; Sun, H.; Sekiguchi, S., "Motion Mapping for MPEG-2 to H.264/AVC Transcoding", *IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1991-1994, May 2007 (TR2007-004)

Dagtas, S.; Pekhteryev, G.; Sahinoglu, Z., "Multi-Stage Real Time Health Monitoring via ZigBee in Smart Homes", *IEEE International Conference on Advanced Information Networking and Applications Workshops (AINA Workshops)*, pp. 782-786, May 2007 (TR2007-056)

Ince, S.; Martinian, E.; Yea, S.; Vetro, A., "Depth Estimation for View Synthesis in Multiview Video Coding", *3DTV Conference (3DTV-CON)*, May 2007 (TR2007-025)

Wigdor, D.; Shen, C.; Forlines, C.L.; Balakrishnan, R., "Perception of Elementary Graphica Elements in Tabletop and Multi-Surface Environments", *Conference on Human Factors In Computing Systems (SIGCHI)*, ISBN: 978-1-59593-593-9, April 2007 (TR2007-054)

Forlines, C.L.; Wigdor, D.; Shen, C.; Balakrishnan, R., "Direct-touch vs. Mouse Input for Tabletop Displays", *ACM Conference on Human Factors in Computer Systems*, ISBN: 978-1-50503-503-9, pp. 647-656, April 2007 (TR2007-053)

Tse, E.; Shen, C.; Greenberg, S.; Forlines, C.L., "How Pairs Interact Over a Multimodal Digital Table", *Conference on Human Factors In Computing Systems (SIGCHI)*, ISBN: 978-1-59593-593-9, pp. 215-218, April 2007 (TR2007-055)

Draper, S.C.; Khisti, A.; Martinian, E.; Vetro, A.; Yedidia, J.S., "Using Distributed Source Coding to Secure Fingerprint Biometrics", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-6149, Vol. 2, pp. II-129--II-132, April 2007 (TR2007-005)

Shashanka, M.V.S.; Raj, B.; Smaragdis, P., "Sparse Overcomplete Decomposition for Single Channel Speaker Separation", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-6149, Vol. 2, pp. 11-641 - II-644, April 2007 (TR2007-031)

Smaragdis, P.; Shashanka, M., "A Framework for Secure Speech Recognition", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ISSN: 1520-6149, Vol. 4, pp. IV-696 - IV-972, April 2007 (TR2007-057)

Raj, B.; Singh, R.; Shashanka, M.; Smaragdis, P., "Bandwidth Expansion with a Polya URN Model", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vol. 4, pp. IV 597 - IV 600, April 2007 (TR2007-058)

* Li, Y.; Mehta, N.B.; Molisch, A.F., Zhang, J.Y., "Optimal Signaling and Selection Verification for Single Transmit-Antenna Selection", *IEEE Transactions on Communications*, ISSN 0090-8778, Vol. 55, Issue 4, pp. 778-789, April 2007 (TR2006-095)

* Cassioli, D.; Win, M.Z.; Vatalaro, F.; Molisch, A.F, "Low Complexity Rake Receivers in Ultra-Wideband Channels", *IEEE Transactions on Wireless Communications*, ISSN: 1536-1276, Vol. 6, Issue 4, pp. 1265-1275, April 2007 (TR2007-049)

Schonberg, D.; Yeo, C.; Draper, S.C.; Ramchandran, K., "On Compression of Encrypted Video", *Data Compression Conference (DCC)*, pp. 173-182, March 2007 (TR2007-059)

Tao, Z.; Teo, K.H.; Zhang, J.Y., "Aggregation and Concatenation in IEEE 802.16j Mobile Multihop Relay (MMR) Networks", *IEEE Mobile WiMAX Symposium*, pp. 85-90, March 2007 (TR2007-019)

Korakis, T.; Tao, Z.; Slutskiy, Y.; Panwar, S., "A Cooperative MAC Protocol for Ad Hoc Wireless Networks", *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 532-536, March 2007 (TR2007-017)

* Du, J.; Li, Y.; Gu, D.; Molisch, A.F.; Zhang, J.Y., "Statistical Rate Allocation for Layered Space - Time Structure", *IEEE Transactions on Communications*, ISSN: 0090-6778, Vol. 55, Issue 3, pp. 489-496, March 2007 (TR2007-050)

* Calcev, G.; Chizhik, D.; Goransson, B.; Howard,S.; Huang, H.; Kogiantis, A.; Molisch, A.F.; Moustakas, A.L.; Reed, D.; Hao, X., "A Wideband Spatial Channel Model for System-Wide Simulations", *IEEE Transactions on Vehicular Technology*, ISSN: 0018-9545, Vol. 56, Issue 2, pp. 389-403, March 2007 (TR2007-051)

Joshi, N.; Matusik, W.; Avidan, S.; Pfister, H.; Freeman, W.T., "Exploring Defocus Matting: Nonparametric Acceleration, Super-Resolution and Off-Center Matting", *IEEE Computer Graphics and Applications*, ISSN: 0272-1716, Vol. 27, Issue 2, pp. 43-52, March 2007 (TR2007-052)

* 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, awarded Best Young Author, March 2007 (TR2004-088)

Porikli, F.M.; Kocak, T., "Fast Distance Transform Computation Using Dual Scan Line Propagation", *SPIE Conference Real-Time Image Processing*, ISBN: 9780819466099, Vol. 6496, February 2007 (TR2007-013)

Sun, H.; Vetro, A.; Xin, J., "An Overview of Scalable Video Streaming: Research Articles", *Wireless Communications & Mobile Computing*, ISSN: 1530-8669, Vol. 7, Issue 2, pp. 159-172, February 2007 (TR2007-007)

* Liu, P.; Tao, Z.; Narayanan, S.; Korakis, T.; Panwar, S.S., "CoopMAC: A cooperative MAC for Wireless LANs", *IEEE Journal on Selected Areas in Communications*, ISSN: 0733-8716, Vol. 25, Issue 2, pp. 340-354, February 2007 (TR2007-018)

* Otsuka, I.; Suginohara, H.; Kusunoki, Y.; Divarakan, A., "Detection of Music Segment Boundaries using Audio-visual Features for a Personal Video Recorder", *IEEE Transactions on Consumer Electronics*, ISSN: 0098-3063, Vol. 53, Issue 1, pp. 150-154, February 2007 (TR2007-037)

* Avidan, S., "Ensemble Tracking", *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, ISSN 0162-8828, Vol. 29, Issue 2, pp. 261-271, February 2007 (TR2005-065)

Ivanov, Y.A.; Sorokin, A.; Wren, C.R.; Kaur, I., "Tracking People in Mixed Modality Systems", *SPIE Conference on Visual Communications and Image Processing (VCIP)*, Volume 6508, January 2007 (TR2007-011)

Draper, S.C.; Khisti, A.; Martinian, E.; Vetro, A.; Yedidia, J.S., "Secure Storage of Fingerprint Biometrics Using Slepian-Wolf Codes", *Information Theory and Applications Workshop (ITA)*, January 2007 (TR2007-006)

Shubin, H.; Divakaran, A.; wittenburg, K.; Peker, K.A; Radhakrishnan, R., "Assessment of End-User Response to Sports Highlights Extraction for Personal Video Recorders", *SPIE Conference Multimedia Content Access Alorithms and sytems*, Vol. 6506, January 2007

Hata, T.; Kuwahara, N.; Schwenke, D.L.; Vetro, A., "Surveillance System with Mega-Pixel Scalable Transcoder", *SPIE Conference Visual Communications and Image Processing*, Vol. 6508, January 2007 (TR2007-008)

Bashir, F.; Porikli, F.M, "Collaborative Tracking of Objects in EPTZ Cameras", *SPIE Conference Visual Communications and Image Processing* , ISBN: 978089466211, Vol. 6508, January 2007 (TR2006-088)

Smaragdis, P.; Boufounos, P., "Position and Trajectory Learning for Microphone Arrays", *IEEE Transaction on Audio, Speech and Language Processing*, ISSN: 1558-7916, Vol. 15, Issue 1, pp. 358-368, January 2007 (TR2007-001)

Smaragdis, P., "Convolutive Speech Bases and their Application to Supervised Speech Separation", *IEEE Transaction on Audio, Speech and Language Processing*, ISSN: 1558-7916, Vol. 15, Issue 1, pp. 1-12, January 2007 (TR2007-002)

Shipman, S.E.; Divakaran, A.; Flynn, M., "Highlight Scene Detection and Video Summarization for PVR-Enabled High-Definition Television Systems", *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 1-2, January 2007 (TR2007-060)

* Wang, B.; Wang, Y.; Selesnik, I.; Vetro, A., "Video Coding Using 3D Dual-Tree Wavelet Transform", *EURASIP Journal on Image and Video Processing*, Vol. 2007, Article ID 42761, 15 pages, January 2007 (TR2007-015)

Vetro, A.; Schwenke, D.L.; Hata, T.; Kuwahara, N., "Scalable Video Streaming Based on JPEG2000 Transcoding with Adaptive Rate Control", *Advances in Multimedia*, Volume 2007, Article ID 62094, 7 pages, December 2006 (TR2007-016)

Sahinoglu, Z.; Gezici, S., "Ranging in the IEEE 802.15.4a Standard", *IEEE Wireless and Microwave Technology Conference (WAMICON)*, December 2006 (TR2006-097)

* Sudarshan, P.; Mehta, N.B.; Molisch, A.F.; Zhang, J.Y., "Channel Statistics-Based RF Pre-Processing with Antenna Selection", *IEEE Transactions on Wireless Communications*, ISSN: 1536-1276, Vol. 5, Issue 12, pp. 3501-3511, December 2006 (TR2006-101)

* Almers, P.; Tufvesson, F.; Molisch, A.F., "Keyhold Effect in MIMO Wireless Channels: Measurements and Theory", *IEEE Transactions on Wireless Communications*, ISSN: 1536-1276, Vol. 5, Issue 12, pp. 3596-3604, December 2006 (TR2003-136)

* Mehta, N.B.; Molisch, A.F.; Greenstein, L.J., "Macrocell-Wide Behavior of the Orthogonality Factor in WCDMA Downlinks", *IEEE Transactions on Wireless Communications*, ISSN: 1536-1276, Vol. 5, Issue 12, pp. 3394-3399, December 2006 (TR2006-102)

* Molisch, A.F.; Asplund, H.; Heddergott, R.; Steinbauer, M.; Zwick, T., "The COST259 Directional Channel Model - Part I: Overview and Methodology", *IEEE Transactions on Wireless Communications*, ISSN: 1536-1276, Vol. 5, Issue 12, pp. 3421-3433, December 2006 (TR2006-111)

* Asplund, H.; Glazunov, A.A.; Molisch, A.F.; Pedersen, K.I.; Steinbauer, M., "The COST 259 Directional Channel Model - Part 11: Macrocells", *IEEE Transactions on Wireless Communications*, ISSN: 1536-1276, Vol. 5, Issue 12, pp. 3434-3450, December 2006 (TR2006-112)

* Zhang, H.; Molisch, A.F.; Zhang, J.Y., "Applying Antenna Selection in WLANs for Achieving Broadband Multimedia Communications", *IEEE Transactions on Broadcasting*, ISSN: 0018-9316, Vol. 52, Issue 4, pp. 475-482, December 2006 (TR2006-113)

Yim, R.; Mehta, N.B.; Molisch, A.F.; Zhang, J.Y., "Progressive Accumulative Routing in Wireless Networks", *IEEE Global Telecommunications Conference (GLOBECOM)*, November 2006 (TR2006-091)

Madan, R.; Mehta, N.B.; Molisch, A.F.; Zhang, J.Y., "Energy-Efficient Cooperative Relaying over Fading Channels with Simple Relay Selection", *IEEE Global Telecommunications Conference (GLOBECOM)*, November 2006 (TR2006-075)

Molisch, A.F.; Mehta, N.B.; Yedidia, J.S.; Zhang, J.Y., "WLC41-6: Cooperative Relay Networks Using Fountain Codes", *IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 1-6, November 2006 (TR2006-092)

Eriksson, G.; Tufvesson, F.; Molisch, A.F., "WEC11-2: Propagation Channel Characteristics for Peer-to-Peer Multiple Antenna Systems at 300 MHz", *IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 106, November 2006 (TR2006-093)

Porikli, F.M; Kocak, T., "Robust License Plate Detection Using Covariance Descriptor in a Neural Network Framework", *IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 107, November 2006 (TR2006-040)

Radhakrishnan, R; Nikovski, D.N.; Peker, K.A.; Divakaran, A., "A Comparison between Polynomial and Locally Weighted Regression for Fault Detection and Diagnosis of HVAC Equipment", *Conference on IEEE Industrial Electronics (IECON)*, ISSN: 1553-572X, pp. 3668-3673, November 2006 (TR2006-094)

Lee, C.; Leigh, D.L.; Ryall, K.; Miyashita, H.; Hirata, K., "Very Fast Subarray Position Calculation for Minimizing Sidelobes in Sparse Linear Phased Arrays", *European Conference on Antennas and Propagation (EuCAP)*, November 2006 (TR2006-022)

Forlines, C.L.; Shen, C.; Wigdor, D.; Balakrishnan, R., "Exploring the Effects of Group Size and Display Configuration on Visual Search", *ACM Conference on Computer Supported Cooperative Work (CSCW)*, ISBN: 1-59593-249-6, pp. 11-20, November 2006 (TR2006-082)

* Tse, E.; Greenberg, S.; Shen, C., "GSI Demo: Multiuser Gesture/Speech Interaction over Digital Tables by Wrapping Single User Applications", *International Conference on Multimodal Interfaces (ICMI)*, ISBN: 1-59593-541-X, pp 76-83, November 2006 (TR2006-085)

Gezici, S.; Sahinoglu, Z.; Poor, H.V., "On the Optimality of Equal Gain Combining for Energy Detection of Unknown Signals", *IEEE Communications Letters*, ISSN: 1089-7798, Vol. 10, Issue 11, pp. 772-774, November 2006 (TR2006-081)

* Ding, G.; Sahinoglu, Z.; Orlik, P.V.; Zhang, J.Y.; Bhargava, B., "Tree-Based Data Broadcast in IEEE 802.15.4 and ZigBee Networks", *IEEE Transactions on Mobile Computing*, ISSN: 1536-1233, Vol. 5, Issue 11, pp. 1561-1574, November 2006 (TR2006-089)

* Tufvesson, F.; Gezici, S.; Molisch, A.F., "Ultra-Wideband Communications using Hybrid-Matched Filter Correlation Receivers", *IEEE Transactions on Wireless Communications*, ISSN: 1536-1276, Vol. 5, Issue 11, pp. 3119-3129, November 2006 (TR2004-016)

* Molisch, A.F.; Cassioli, D.; Chong, C-C.; Emami, S.; Fort, A.; Kannan, B.; Karedal, J.; Kunisch, J.; Schantz, H.G.; Siwiak, K.; Win, M.Z., "A Comprehensive Standardized Model for Ultra Wideband Propagation Channels", *IEEE Transactions on Antennas and Propagation*, ISSN: 0018-926X, Vol. 54, Issue 11, Part 1, pp. 3151-3166, November 2006 (TR2006-114)

* Lau, B.K.; Andersen, J.B.; Kristensson, G.; Molisch, A.F., "Impact of Matching Network on Bandwidth of Compact Antenna Arrays", *IEEE Transactions on Antennas and Propagation*, ISSN: 0018-926X, Vol. 54, Issue 11, Part 1, pp. 3225-3238, November 2006 (TR2006-115)

Wren, C.R.; Minnen, D.C.; Rao, S.G., "Similarity-based Analysis for Large Networks of Ultra-Low Resolution Sensors", *Pattern Recognition*, ISSN: 0031-3203, Vol. 39, Issue 10, pp. 1918-1931, October 2006 (TR2005-003)

* Weinberg, G.L.; Raj, B.; Kalgaonkar, K., "Two New Techniques for Natural Spoken User Interfaces", *ACM Symposium on User Interface Software and Technology (UIST)*, October 2006 (TR2006-098)

* Forlines, CL..; Esenther, A.W.; Shen, C.; Wigdor, D.; Ryall, K., "Multi-User, Multi-Display Interaction with a Single-User, Single-Display Geospatial Application", *ACM Symposium on User Interface Software and Technology (UIST)*, ISBN: 1-59593-313-1, pp. 273-276, October 2006 (TR2006-083)

* Wigdor, D.; Leigh, D.L.; Forlines, C.L.; Shipman, S.E.; Barnwell, J.C.; Balakrishnan, R.; Shen, C., "Under the Table Interaction", *ACM Symposium on User Interface Software and Technology (UIST)*, ISBN: 1-59593-313-1, pp. 259-268, October 2006 (TR2006-076)

* Forlines, C.L.; Vogel, D.; Balakrishnan, R., "HybridPointing: Fluid Switching between Absolute and Relative Pointing with a Direct Input Device", *ACM Symposium on User Interface Software and Technology (UIST)*, ISBN: 1-59593-313-1, pp. 211-220, October 2006 (TR2006-090)

Shao, J.; Porikli, F.M.; Chellappa, R., "A Cascading Framework of Contour Motion and Deformation Estimation for Non-Rigid Object Tracking", *International Conference on Image Processing (ICIP)*, October 2006 (TR2006-044)

Porikli, F.M.; Tuzel, O., "Fast Construction of Covariance Matrices for Arbitrary Size Image Windows", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, pp.1581-1584, October 2006 (TR2006-043)

Martinian, E.; Behrens, A.; Xin, J.; Vetro, A.; Sun, H., "Extensions of H.264/AVC for Multiview Video Compression", *IEEE International Conference on Image Processing (ICIP)*, ISSN: 1522-4880, pp. 2981-2984, October 2006 (TR2006-048)

Feris, R.; Tuk, M.; Raskar, R., "Dealing with Multi-Scale Depth Changes and Motion in Depth Edge Detection", *Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI)*, pp. 3-10, October 2006 (TR2006-110)

Martinian, E.; Vetro, A.; Yedidia, J.S.; Ascenso, J.; Khisti, A.; Malioutov, D., "Hybrid Distributed Video Coding Using SCA Codes", *IEEE Workshop on Multimedia Signal Processing (MMSP)*, pp. 258-261, October 2006 (TR2006-069)

Molisch, A.F.; Orlik, P.V.; Sahinoglu, Z.; Zhang, J.Y., "UWB-based Sensor Networks and the IEEE 802.15.4a Standard", *International Conference Communications and Networking in China (ChinaCom)*, pp. 1-6, October 2006 (TR2006-117)

Molisch, A.F.; Mehta, N.B.; Zhang, H.; Almers, P.; Zhang, J.Y., "Implementation Aspects of Antenna Selection for MIMO Systems", *International Conference Communications and Networking in China (ChinaCom)*, pp. 1-7, October 2006 (TR2006-116)

Radhakrishnan, R.; Divakaran, A., "Merging Segmentations of Low-level and Mid-level Time Series for Audio Class Discovery", *Asilomar Conference on Signals, Systems and Computers (ACSSC)*, ISSN: 1058-6393, pp. 64-68, October - November 2006 (TR2006-121)

* Porikli, F.M., "Achieving Real-Time Object Detection and Tracking Under Extreme Conditions", *Journal of Real-time Image Processing*, ISSN: 1861 8200, Vol. 1, Number 1, pp. 33-40, October 2006 (TR2006-039)

Mei, X.; Porikli, F.M., "Fast Image Registration via Joint Gradient maximization: Application to Multimodal Data", *SPIE Conference Electro-Optical and Infrared Systems*, ISBN: 9780819464934, Vol. 6395, September 2006 (TR2006-109)

Ivanov, Y.A.; Hamid, R., "Weighted Ensemble Boosting for Robust Activity Recognition in Video", *International Conference on Computer vision and Graphics (ICCVG)*, September 2006 (TR2006-080)

Sahinoglu, Z.; Guvenc, I.; Orlik, P.V.; Molisch, A.F., "Interference Suppression in Non-coherent Time-hopping IR-UWB Ranging", *IEEE International Conference on Ultra-Wideband (ICUWB)*, pp. 507-511, September 2006 (TR2006-120)

Gezici, S.; Poor, H.V.; Kobayashi, H.; Molisch, A.F., "Optimal and Suboptimal Linear Receivers for Impulse Radio UWB Systems", *IEEE International Confrence on Ultra-Wideband (ICUWB)*, September 2006 (TR2006-067)

Ryall, K.; Esenther, A.W.; Forlines, C.L.; Shen, C.; Shipman, S.E.; Morris, M.R.; Everitt, K.; Vernier, F.D., "Identity-Differentiating Widgets for Multiuser Interactive Surfaces", *IEEE Computer Graphics and Applications*, ISSN: 0272-1716, Vol. 26, Issue 5, pp. 56-64, September 2006 (TR2006-100)

Li, Y.G.; Molisch, A.F.; Zhang, J.Y., "Practical Approaches to Channel Estimation and Interference Suppression for OFDM-based UWB Communications", *IEEE Transactions on Wireless Communications*, ISSN: 1536-1276, Vol. 5, Issue 9, pp. 2317-2320, September 2006 (TR2004-095)

Shen, C.; Ryall, K.; Forlines, C.L.; Esenther, A.W.; Vernier, F.D.; Everitt, K.; Wu, M.; Wigdor, D.; Morris, M.R.; Hancock, M.; Tse, E., "Informing the Design of Direct-Touch Tabletops", *IEEE Computer Graphics and Applications*, ISSN: 0272-1716, Vol. 26, Issue 5, pp. 36-46, September 2006 (TR2006-086)

Almers, P.; Santos, T.; Tufvesson, F.; Molisch, A.F.; Karedal, J.; Johasson, A.J., "Measured Diversity Gains from MIMO Antenna Selection", *IEEE Vehicular Technology Conference (VTC-2006* Fall), ISBN: 1-4244-0063-5, pp. 1-6, September 2006

Raskar, R.; Ziegler, R.; Willwacher, T., "Cartoon Dioramas in Motion", *ACM SIGGRAPH*, ISBN: 1-59593-364-6, Article 6, July 2006 (TR2006-017)

Bimber, O.; Raskar, R., "Modern Approaches to Augmented Reality", *ACM SIGGRAPH*, ISBN: 1-59593-364-6, Article 1, July 2006 (TR2006-105)

Frisken, S.F.; Perry, R.N., "Designing with Distance Fields", *ACM SIGGRAPH*, ISBN: 1-59593-364-6, pp. 60-66, July 2006 (TR2006-054)

Raskar, R.; van Baar, J.; Beardsley, P.A.; Willwacher, T.; Rao, S.; Forlines, C.L., "iLamps: Geometrically Aware and Self-configuring Projectors", *ACM SIGGRAPH*, ISBN: 1-59593-364-6, Article 7, July 2006 (TR2006-103)

Raskar, R.; Feris, R.; Yu, J.; Turk, M., "Non-photorealistic Camera: Depth Edge Detection and Stylized Rendering Using Multi-flash Imaging", *ACM SIGGRAPH*, ISBN: 1-59593-364-6, Article 2, July 2006 (TR2006-107)

Raskar, R.; Beardsley, P.A.; van Baar, J.; Wang, Y.; Dietz, P.; Lee, J.; Leigh, D.L.; Willwacher, T., "RFIG Lamps: Interacting with a Self-Describing World via Photosensing Wireless Tags and Projectors", *ACM SIGGRAPH*, ISBN: 1-59593-364-6, Article 11, July 2006 (TR2006-108)

- Golovinskiy, A.; Matusik, W.; Pfister, H.; Rusinkiewicz, S.; Funkhouser, T., "A Statistical Model for Synthesis of Detailed Facial Geometry", *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 25, Issue 3, pp. 1025-1034, July 2006 (TR2006-072)

- Weyrich, T.; Matusik, W.; Pfister, H.; Bickel, B.; Donner, C.; Tu, C.; McAndless, J.; Lee, J.; Ngan, A.; Jensen, H.; Gross, M., "Analysis of Human Faces using a measurement-Based Skin Reflectance Model", *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 25, Issue 3, pp. 1013-1024, July 2006 (TR2006-071)

Jakubiak, E.J.; Perry, R.N.; Frisken, S.F., "An Improved Representation for Stroke-based Fonts", *ACM SIGGRAPH*, ISBN: 1-59593-364-6, Article 137, July 2006 (TR2006-119)

* Nayar, S.K.; Krishnan, G.; Grossberg, M.D.; Raskar, R., "Fast Separation of Direct and Global Components of a Scene Using High Frequency Illumination", *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 25, Issue 3, pp. 93-944, July 2006 (TR2006-087)

* Joshi, N.; Matusik, W.; Avidan, S., "Natural Video Matting Using Camera Arrays", *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 25, Issue 3, pp, 779-786, July 2006 (TR2006-106)

* Raskar, R.; Agrawal, A.; Tumblin, J., "Coded Exposure Photography: Motion Deblurring Using Fluttered Shutter", *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 25, Issue 3, pp. 795-804, July 2006 (TR2006-124)

* Lawrence, J.; Ben-Artzi, A.; DeCoro, C.; Matusik, W.; Pfister, H.; Ramamoorthi, R.; Rusinkiewicz, S., "Inverse Shade Trees for Non-Parametric Material Representation and Editing", *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 25, Issue 3, pp. 735-745 (TR2006-104)

* Peers, P.; Berge, K.; Matusik, W.; Ramamoorthi, R.; Lawrence, J.; Rusinkiewicz, S.; Dutre, P., "A Compact Factored Representation of Heterogeneous Subsurface Scattering", *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 25, Issue 3, pp. 746-753, July 2006

* Gu, J.; Tu, C.; Ramamoorthi, R.; Belhumeur, P.; Matusik, W.; Nayar, S., "Time-Varying Surface Appearance: Acquisition, Modeling and Rendering", *ACM Transactions on Graphics (TOG)*, ISSN: 0730-0301, Vol. 25, Issue 3, pp. 762-771, July 2006

Esenther, A.W.; Ryall, K., "RemoteDT: Support for Multi-Site Table Collaboration", *International Conference on Collaboration Technologies (CollabTech)*, July 2006 (TR2006-084)

Peker, K.A.; Otsuka, I.; Divakaran, A., "Broadcast Video Program Summarization using Face Tracks", *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1053-1056, July 2006 (TR2006-123)

Schwenke, D.L.; Vetro, A.; Hata, T.; Kuwahara, N., "Dynamic Rate Control for JPEG 2000 Transcoding", *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1437-1440, July 2006 (TR2006-073)

Xin, J.; Vetro, A.; Sekiguchi, S.; Sugimoto, K., "Motion and Mode Mapping for MPEG-2 to H.264/AVC Transcoding", *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 313-316, July 2006 (TR2006-074)

Radhakrishnan, R.; Divakaran, A.; Otsuka, I., "Sports Program Boundary Detection", *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1621-1624, July 2006 (TR2006-122)

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 six topic areas.

- Digital Communications
- Multimedia
- Data & Sensor Systems
- Imaging
- Mechatronics
- Algorithms

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 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.

Digital Communications

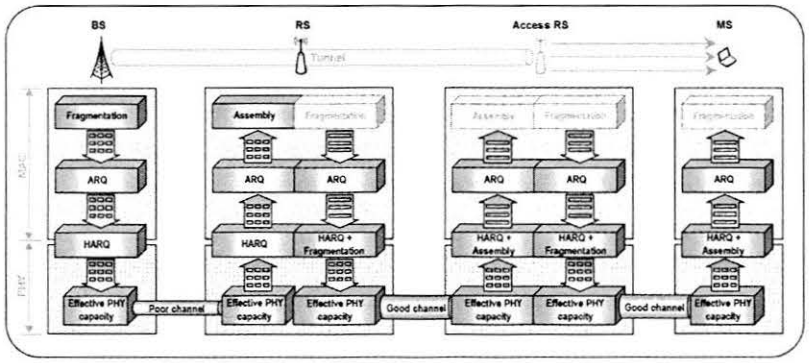
Digital communications and networking are pervasive in today's society. With advanced technologies at physical layer, medium access control layer and network layer, it will provide high speed communication capability for transmissions of voice, data as well multimedia information with quality of service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

At MERL, our goal is to seek new business and technology trends in the area of digital communications and networking. We are not only conducting fundamental researches on communication theories and developing new core technologies, but also apply these technologies to international standards and emerging products.

For broadband mobile communications, MERL continues to develop MIMO (Multi-Input-Multi-Output) technologies, such as antenna selection and control signaling design. MERL is developing cooperative communications, cognitive radio and interference management, frame structures and efficient medium access control (MAC) and contributes to 3GPP LTE and Advanced WiMAX.

Project Descriptions

Hybrid ARQ for Future WiMAX System



As a mandatory feature of WiMAX system, hybrid ARQ (HARQ) plays a crucial role in ensuring transmission reliability needed on the wireless link. We evaluated the performance of current HARQ protocol in the context of 802.16j multihop relay network and 802.16m next generation WiMAX system. Based upon the insights gained therein, we further designed an adaptive HARQ scheme, which performs intelligent fragmentation to adapt to the channel condition, and provides indispensable support to traffic aggregation/tunneling.

Background and Objectives: HARQ is a key technology in next generation wireless systems that spans both MAC and PHY layers, and exploits time/frequency diversity and coding gain. However, the complicated interaction between HARQ, ARQ at MAC layer and the capacity at physical layer has not been completely understood, let alone been optimized. The objective of this work is to establish a full understanding of the cross-layer protocol interaction and address the performance issues identified thereof.

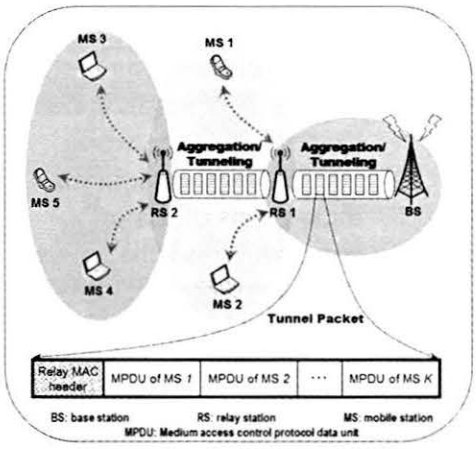
Technical Discussion: We first developed a more profound understanding of the interaction between MAC layer ARQ, HARQ and physical layer capacity, the confirmed that no window lock effect would occur at ARQ or HARQ level, given the current setting of IEEE 802.16e PHY. Nevertheless, the study reveals that HARQ performance can severely deteriorate in IEEE 802.16j relay network or IEEE 802.16m next generation WiMAX system, due to potential mismatch of various operating parameters. As a solution, we designed an adaptive HARQ protocol, which fragments upper layer packet at the MAC layer instead of in HARQ, if the physical channel condition degrades below certain threshold. Moreover, we identified the operating region of the number of parallel channels (i.e., ACID) in HARQ that can optimize the system level performance for the 802.16j and 802.16m network. Last but not least, we proposed a set of new features for HARQ so that it can support the aggregation/tunnel operation on relay links.

Future Direction: The proposed HARQ scheme that provides support to tunneling/aggregation operation has been accepted into IEEE 802.16j draft standard. New contributions based upon the research results of this project will be made to IEEE 802.16m.

Contact: Zhifeng (Jeffrey) Tao
http://www.merl.com/projects/mmrharq/

Project Type: Research

Aggregation in IEEE 802.16j Network



The mobile multihop relay-based (MMR) network architecture of IEEE 802.16j imposes a demanding performance requirement on relay stations. We proposed a concatenation and enhanced packing scheme, which incarnates the inherent notion of "aggregation", alleviates the dismal efficiency degradation on relay links, and significantly simplifies the traffic handling at intermediate relay stations.

Background and Objectives: The IEEE 802.16j mobile multi-hop relay-based (MMR) network architecture imposes a demanding performance requirement on relay stations. These relays will functionally serve as an aggregating point on behalf of the base station (BS) for traffic collection from and distribution to multiple mobile stations (MSs) associated with the relays, and thus naturally incorporate a notion of "traffic aggregation". However, the legacy concept of connection and the associated packet construction mechanism defined in IEEE 802.16d/16e standard, if applied directly on relay link, may render a potential bottleneck and preponderantly limit the overall network capacity. In order to provide high capacity demanded by relay links, therefore, design of highly efficient connection protocols becomes imperative.

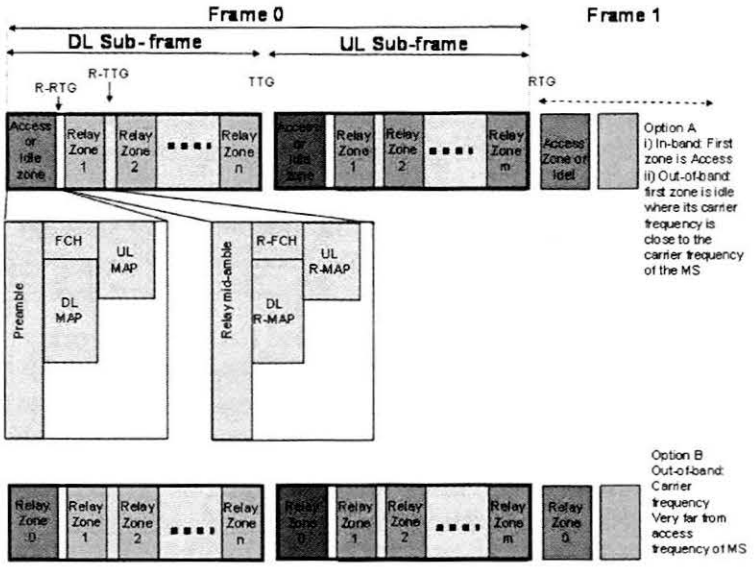
Technical Discussion: As a solution, we proposed a concatenation and enhanced packing scheme, wherein a logical "mega-pipe" (a.k.a. tunnel) is established between access relay station and base station to transport traffic belonging to multiple individual connections. These individual connections to be aggregated can originate from different mobile stations, but shall share some common characteristics (e.g., quality of service (QoS) requirement). The creation, maintenance and identification of such a "mega-pipe", among others, have been optimized so that the efficiency at data plane is substantially improved while the associated overhead in the control plane is minimized, thereby enabling IEEE 802.16j MMR network to deliver a superior performance. Given its nature of aggregation, the proposed scheme also helps simplify the QoS and path management at intermediate relay stations significantly. Moreover, we designed a packet and header format for relay link, which not only enables the operation of connection aggregation, but also supports co-existence of "mega-pipe" and legacy 802.16e connection.

Future Direction: The proposed aggregation protocol and the associated high level packet format design have been accepted into IEEE 802.16j draft standard. Continued participation in future standardization activities is needed. In addition, new contributions will be made to IEEE 802.16j to specify detailed realization and format design.

Contact: Zhifeng (Jeffrey) Tao
http://www.merl.com/projects/mmraggregation/

Project Type: Research

Adaptive Frame Structure for Broadband Mobile Multihop Wireless System



The frame structure design is more challenging in the new mobile multihop relay based (MMR) network architecture, as numerous dimensions of design constraints and challenges have been introduced therein. In this project, we proposed an adaptive frame structure to meet the challenges of a mobile multihop relay based network.

Background and Objectives: The new IEEE 802.16j standard was created to help to resolve problems like limited coverage and terminal transmit power of the IEEE 802.16e standard. In this project, our objective is to develop an adaptive frame structure that is simple yet flexible and one which enables multihop operation while still maintaining the backward compatibility with the legacy mobile stations of IEEE802.16e.

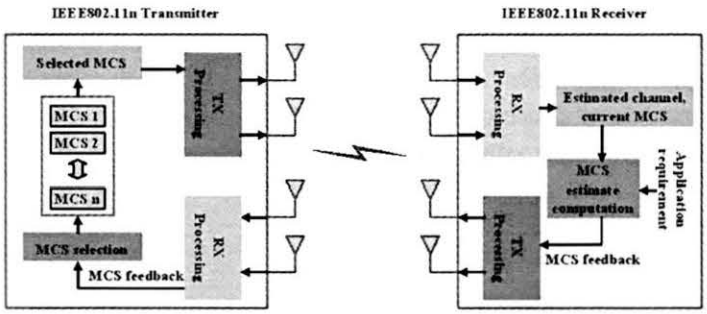
Technical Discussion: Similar to the legacy design, the new frame structure for MMR network is also composed of a DL (Down Link) and an UL (Up Link) portion. However, in order to enable multihop communication, the DL and UL subframe is further divided into multiple zones in the time domain. As depicted in the Figure, the first zone in both the DL and UL subframe is dedicated for communication that directly engages MSs (Mobile Stations), and thus is naturally called the access zone. The access zone in both DL and UL may be followed by one or multiple relay zones. In each relay zone, BS (Base Station) and RS (Relay Station) can stay in the mode of transmission, reception or being idle. The number of relay zone in the DL sub-frame "n" may or may not be equal to the number of relay zone in the UL sub-frame "m". When "n" and "m" equal to zero, it implies that the frame structure becomes the frame structure of the 802.16e standards. The duration of both the access and relay zone are flexible as long as they are confined within the duration of the sub-frame.

Future Direction: To develop an adaptive frame structure that can more effectively deal with real time signal and data simultaneously.

Contact: Koon Hoo Teo, Jinyun Zhang
http://www.merl.com/projects/framestructure/

Project Type: Research

Adaptive Modulation and Coding for IEEE802.11n



The link adaptation is critical to achieve high throughput and sustain reliable communications for MIMO-OFDM based IEEE802.11n WLAN. We developed an intelligent link adaptation technique for IEEE802.11n MIMO-OFDM WLAN with convolutional coding. Our technique dynamically selects the best modulation and coding scheme based on channel condition and MAC layer performance (coded BER or PER) requested by applications. We have shown that the coded performance at decoder output can be modeled as a function of decoder-input uncoded BER and the code rate. For MIMO-OFDM systems with different modulation types for different spatial streams at each subcarrier, we developed a simple and closed form solution to compute the BER and PER for each spatial stream mixed from different subcarrier.

Background and Objectives: IEEE802.11n standard employs multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) transmission technology to enable high throughput communication for up to 600 Mbps in a 40 MHz channel bandwidth. Since the wireless fading channel varies with time, link adaptation must be employed to sustain reliable communications and maximize throughput. To fully exploit MIMO channel variations and transmit beacomforming on a MIMO link, IEEE802.11n standard defines a large set of modulation and coding schemes (MCS) to facilitate this goal. Therefore, the efficient and practical link adaptation techniques are needed to be developed for IEEE802.11n WLAN. Our objective is to develop a simple and intelligent link adaptation technique to dynamically select the best link for IEEE802.11n WLAN.

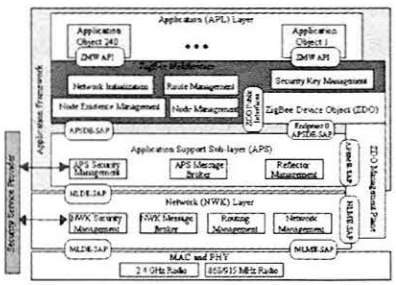
Technical Discussion: There are extensive techniques proposed on adaptive modulation and coding. However, these techniques do not benefit IEEE802.11n WLAN well since they are either too complicated for implementation or do not fully exploit characteristics of MIMO fading channel. We have developed a dynamic link adaptation technique that maximizes the throughput while satisfying the application performance requirement: code bit-error rate (BER) or MAC packet-error rate (PER). We have demonstrated that our dynamic link adaptation technique is a simple and practical solution for IEEE802.11n MIMO-OFDM WLAN with or without beamforming.

Future Direction: The proposed link adaptation technique needs to be incorporated into IEEE802.11n WLAN to achieve a better performance.

Contact: Jianlin Guo, Jinyun Zhang
http://www.merl.com/projects/adaptation/

Project Type: Initial Investigation

ZigBee Middleware Development



ZigBee Middleware – Stack View

ZigBee Middleware is a control software that offers an effective control over ZigBee networks for dynamic load balancing, better resources utilization, and increased reliability. It provides a global view of entire network that allows real-time monitoring of network traffic for avoiding hot spots and reporting vital node information. In addition, it implements certain security related functions such as calculation of device keys and distribution of field key. ZigBee control software resides between ZigBee stack and application profiles and acts as an interface between the two layers while offering added functionality of resource management and operational control.

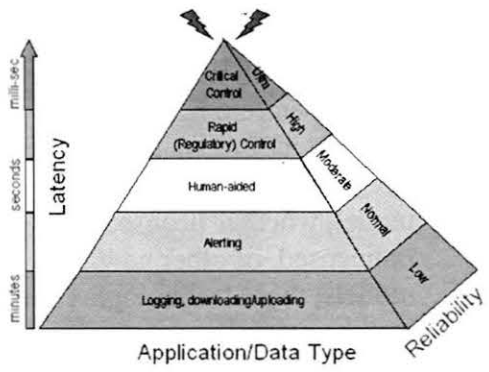
Background and Objectives: ZigBee is short-range multi-hop wireless mesh networking in diverse environments such as industrial, commercial, home, environmental monitoring, and military applications. Distributed algorithmic approach of ZigBee, however, introduces important limitations on its performance. Embedded address assignment and its automatic route discovery, despite being useful features, can result in uneven use of network resources. Moreover, a lack of real-time status reporting can cause significantly higher packet drop rate if some nodes start encountering operational problems. Useful node information, such as current battery level, packet loss rate, and average traffic delay, is not reported or leveraged in ZigBee specification. Our ZigBee middleware aims at addressing these limitations. While it can be deployed in diverse applications, the initial target systems are home automation, commercial building automation, and AMR.

Technical Discussion: ZigBee uses IEEE 802.15.4 PHY/MAC and defined its network layer, APS/ZDO sub-layers, and several profiles. Our middleware resides above ZDO but below application profiles. It follows a modular approach for a robust and fault tolerant control. It operates in two modes, one for host device, which acts as a controller, and another for regular ZigBee nodes. The host acts as a hub for object associations and message reflections. It creates virtual nodes and objects corresponding to the real ones in the network. All application commands pass through the host and all state changes in real objects are notified to it as well. The host allows primitives for model translation that may be performed while forwarding commands across device classes.

Future Direction: MERL will continue working on ZigBee based value added systems.

Contact: Ghulam Bhatti, Jinyun Zhang
http://www.merl.com/projects/zbmiddleware/ **Project Type:** Advanced Development

Ultra Reliable Wireless



To achieve ultra reliable wireless network, this project first aims to address radio redesign including cross-layer optimizations for routing and channel access with frequency hopping, and then focuses on developing cooperation algorithms for mitigating interference and exploiting path diversities.

Background and Objectives: Wireless networks are currently being used for remote sensing and industrial plant monitoring. However, for wireless to be adopted in industrial automation, requirements for ultra-reliability and end-to-end latency in the order of several milliseconds should be met. ZigBee and emerging ISA SP100.11a standards are away from provisioning these stringent QoS levels. Clearly, there will be a need to improve performance levels of wireless communication technologies within the next several years.

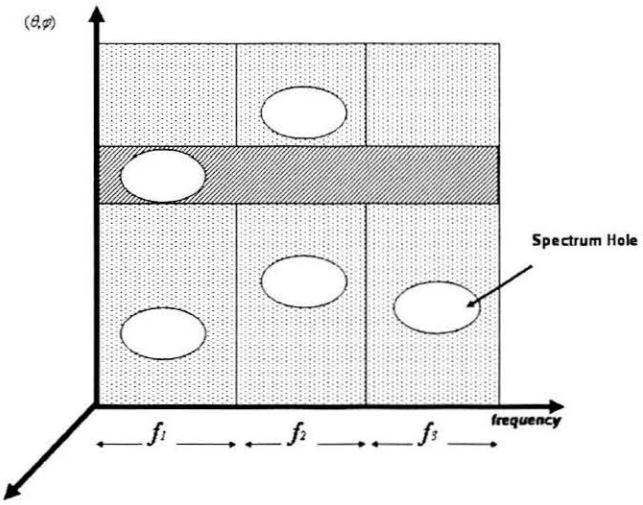
Delay and reliability requirements of industrial process monitoring applications are achievable with state-of-the-art wireless technologies such as wireless HART, SP100.11a and ZigBee-pro. Although, customers already desire to switch to wireless industrial automation platforms, wireless solutions are not mature yet to be adopted into industrial automation and milliseconds level delay critical applications. On the other hand, the emerging ISA SP100.11a standard provides a major milestone in this direction. It is not very far from today that we will start seeing wireless factories. Our objective is to conduct research on ultra-reliable and low-latency wireless networks. The scope includes development of cross-layer design techniques and improvements to existing standard based wireless solutions.

Technical Discussion: In order to improve communication reliability and lower latency to required levels, certain issues need to be addressed at PHY, MAC and networking layers. Adaptive modulation and channel coding together with frequency hopping provide robustness. Furthermore, interference aware channel access, QoS based prioritization of message transmissions and deploying hybrid ARQ schemes improve both MAC efficiency and reliability. Developing reservation based dynamic route discovery and maintenance protocols and exploiting machine learning algorithms in making network routing more intelligent and adaptive to changes will also contribute to overall network performance.

Future Direction: We will be developing cooperative communication and cross-layer optimization techniques to improve reliability of wireless communications and to minimize end-to-end latency to the levels of several milliseconds.

Contact: Zafer Sahinoglu
http://www.merl.com/projects/urwireless/ **Project Type:** Research

Cognitive Radio with Space-Time-Frequency Spectrum Sensing



To date, cognitive radio (CR) has been proposed as the means to promote efficient utilization of the spectrum by exploiting the existence of spectrum holes. Irregardless of the regulatory model, whether licensed, unlicensed, or other new models, these technologies will increasingly allow more intensive and efficient access to, and use of, radio spectrum than possible with traditional, hardware-based radio systems.

Background and Objectives: With the rapid growth of wireless communication in the past two decades, there is a vast and growing demand for radio spectrum, and in particular for those frequencies under 3GHz. However, in a recent FCC task force finding, it confirmed that in many bands, spectrum access is a more significant problem than physical scarcity of spectrum. Given the limited spectrum resource and the increasing demand for radio spectrum by the wireless industries, there is an urgent need to improve both the spectrum access and the efficiency of spectrum utilization and one objective of cognitive radio is to fulfill this function.

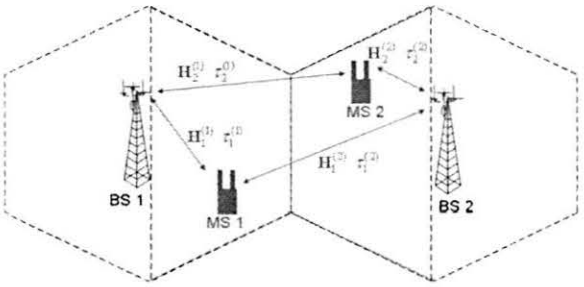
Technical Discussion: Figure depicted a method known as Space-time-frequency sensing of RF spectrum in cognitive radio. Conventionally, cognitive radio techniques conduct one-dimensional spectrum sensing by periodically scanning only the frequency domain to locate unused frequency bands (spectrum holes). The unused frequency bands can then be used for CR to transmit signals that do not interfere with the signals of other radios. In cooperative spectrum sensing, associated CRs can exchange local sensing results, so that a cognitive network obtains an accurate estimate of unused frequency band, or even, locations of the other radios.

Future Direction: Common goals of a cognitive radio include optimizing throughput to the clients, maximizing reliability, minimizing interference and latency, or doing a combination of these along with other network performance measures.

Contact: Koon Hoo Teo, Jinyun Zhang
http://www.merl.com/projects/cognitiveradio/

Project Type: Research

Base Station Cooperation



This project considers cooperative transmission by base stations in multiuser, multi-cell multiple input multiple output (MIMO) systems. A key result is that in such systems, the multiuser interference is asynchronous by nature. We establish an accurate mathematical model for the asynchronity and show that it leads to a significant performance degradation of existing designs, which were derived based on the idealized assumption of synchronous interference. We therefore propose three new linear pre-coding algorithms that are shown to be better in mitigating the impact of the asynchronous interference and improve overall system spectral efficiency.

Background and Objectives: In the downlink of conventional cellular systems, base stations send out signals to the mobile stations in their cells without regard to the mobile stations in neighboring cells. The resulting interference can be mitigated by the use of different frequency bands or different codes in the cells, but such an approach decreases the spectral efficiency of the overall system. A more efficient method is based on a cooperation between neighboring base stations that reduces interference without the use of different transmission frequencies. Each base station knows, and transmits - with linear precoding - the signals to be transmitted to the mobile stations not only in its own cell, but also its neighbors. The precoding is designed to optimize the signal-to-noise and interference ratio at all the mobile stations. The major aspect of our work is to take into account the fact that the interference is inherently asynchronous to the desired signals.

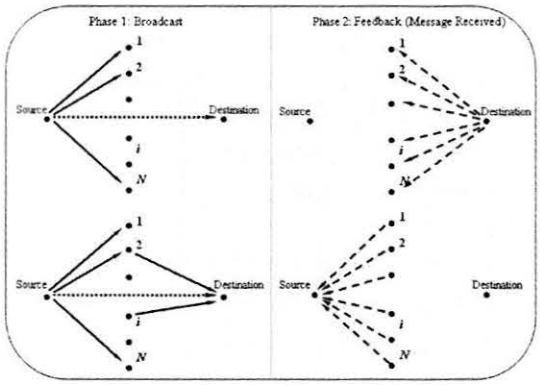
Technical Discussion: In recent years, several advanced techniques have been proposed to better mitigate the effect of inter-cell CCI in cellular MIMO systems. all the schemes proposed for BS cooperation invariably assume that both the desired and the interfering signals from different BSs arrive at each of the MSs synchronously (at the same time). The BSs can align their transmissions so that the signals intended for any MS arrive at that MS synchronously. However, even under the assumption of perfect BS cooperation, the BSs cannot simultaneously control when these signals are received as interference by other MSs. This severely degrades the performance of the previously proposed designs, which do not account for it, especially, in high data rate regimes. In our work, we develop a mathematical formulation of the asynchronous interference problem, and use it to derive three new algorithms for linear precoding design that provide different trade-offs between complexity and performance in different SNR regimes.

Collaboration: This work was done in collaboration with Dr. Hongyuan Zhang and Prof. Huaiyu Dai from North Carolina State University. Dr. H. Zhang is now with Marvell Semiconductors.

Contact: Andreas F. Molisch
Confidential

Technology Area: Digital Communications
Project Type: Research

Cooperative Communication Using Fountain Codes



We consider the transmission of information via a number of relay nodes. By using so-called rateless codes, we allow the receiver to accumulate information in a fast and energy-efficient way. During the past year, we have considered a number of theoretical as well as practical issues that are important for migration to products based on our technology. In particular, we have analyzed (i) the impact of nonorthogonality of the data streams transmitted by different nodes, (ii) the performance of the scheme in wireless channels with shadow fading, and (iii) analytical bounds on the performance of our previously suggested asynchronous transmission scheme.

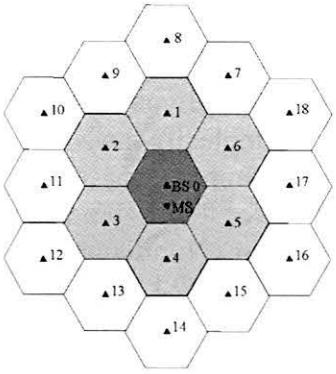
Background and Objectives: In cooperative communications with information accumulation, messages are transmitted from one node to another via several, parallel, relay nodes. At the receiver, the information from all the different relay nodes is added up. This is different from the conventional approach, where the energy from the different relay nodes is added up at the receiver. In order for the receiver to be able to distinguish the messages from different relay nodes, they have to be transmitted on orthogonal channels, e.g., using different spreading codes; nonorthogonality would result in inter-stream interference and thus lead to a reduction in performance.

Technical Discussion: As proposed in the previous year's phase of the project, the transmission from the source to the destination via several parallel relays uses rateless codes. They encode and transmit the source information in an infinitely long codestream; a receiver can recover the original information from unordered subsets of the codestream, once the total obtained mutual information from multiple sources marginally exceeds the entropy of the source information. During this year, we analyzed the impact of inter-stream interference, i.e., the case that the receiver cannot perfectly separate the information transmitted from different relay nodes on different CDMA spreading codes. We found that the impact of such interstream interference is very small; the reason being that the stream that contributes most to the total received mutual information is also the one that suffers least from the interstream interference. For our previously developed asynchronous transmission scheme, we developed an upper and lower bounds on the performance, which are related to the cases where the links between relay nodes are either extremely weak or extremely strong. Finally, we analytically derived the energy consumption and latency for our previously developed synchronous scheme in the case that the fading follows a lognormal distribution, which is relevant for the case that shadowing occurs in the channel.

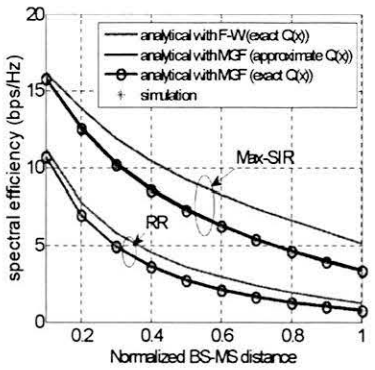
Future Direction: Simulation of the scheme with realistic fountain codes, and implementation on a testbed.

Contact: Andreas F. Molisch
http://www.merl.com/projects/ratlesscodes/ **Project Type:** Research

Spectral Efficiency Analysis with Channel-aware Schedulers



Multi-tier cellular layout



Performance analysis results

Accurate planning and performance evaluation of current and next generation cellular wireless communication systems requires understanding the joint impact of scheduling, interference, and fading. This project strives to provide accurate analytical expressions for the spectral efficiency in bits/sec/Hz of cellular systems that use either channel-aware of channel-unaware schedulers.

Background and Objectives: While considerable amount of work has been done in the literature on the spectral efficiency of cellular systems, it either requires costly numerical simulations or makes several simplifying assumptions that limit its applicability. We develop a general analysis that allows for non-identical co-channel interference from neighboring cells, and in which all links undergo both small-scale Rayleigh fading and large-scale lognormal shadowing. Doing so avoids the loose bounding approaches of the literature that assume that interferers are located in the worst-case or best-case locations. The effect of finite modulation constellations on spectral efficiency is also modeled. Both the fair, but channel-unaware, round-robin scheduler and the greedy, but unfair, max-throughput or max-SIR scheduler are analyzed.

Technical Discussion: Our analysis incorporates two additional tweaks, which are essential in making it more accurate than previous works: (i) the use of a novel moment generating function based lognormal approximation method instead of the conventional Fenton-Wilkinson approximation approach, and (ii) an accurate approximation of the Gaussian-Q function instead of a looser bound. Consequently, our analytical results match the reference simulation results very well. Our results demonstrate that multi-tier interference has a greater impact on the performance of the round-robin scheduler than the max-throughput scheduler. On the other hand, the limited constellation size affects the max-throughput scheduler much more than the round-robin scheduler, especially when the users are close to the serving base station. Our analysis thus provides a useful benchmark for calibrating system-level simulators, which are required to evaluate the overall multi-user, multi-cell performance of cellular systems.

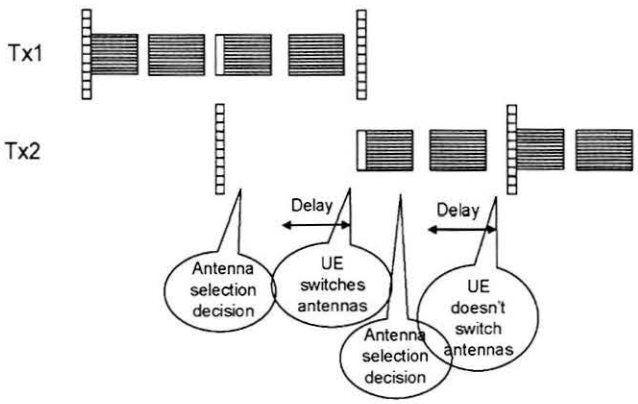
Collaboration: Prof. Jingxian Wu, Sonoma State University, Sonoma, CA.

Future Direction: Future work includes generalizing the analysis to proportional fair schedulers, which effectively balance throughput and fairness.

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

Project Type: Research

Antenna Selection for Broadband Mobile Communications



Antenna selection is a promising technique that realizes the diversity benefits of multiple-antenna systems, while keeping the potentially prohibitive hardware cost and complexity low. We are investigating transmit antenna selection at handsets for fourth-generation OFDMA-based cellular communications systems, in particular, 3GPP LTE (long-term evolution) systems. We consider training for transmit antenna selection on the uplink, and propose a method that is well suited for LTE.

Background and Objectives: In antenna selection (AS), only the signals from a limited-sized subset of the available antenna elements are adaptively chosen and processed by radio frequency (RF) and baseband circuitry. This greatly reduces the cost of a transceiver, since antenna elements are typically cheap, while RF chains that consist of mixers, low-noise amplifiers, and oscillators, are expensive.

Considerable work has appeared in the literature on AS, its capacity, diversity benefits, and selection criteria. However, one critical aspect that has received little attention is training, through which the channel states of all antennas are determined in order to select the best antenna(s). While AS requires fewer RF chains, it also implies that only a subset of the antennas can be activated at any instant to transmit OFDMA pilot signals for training.

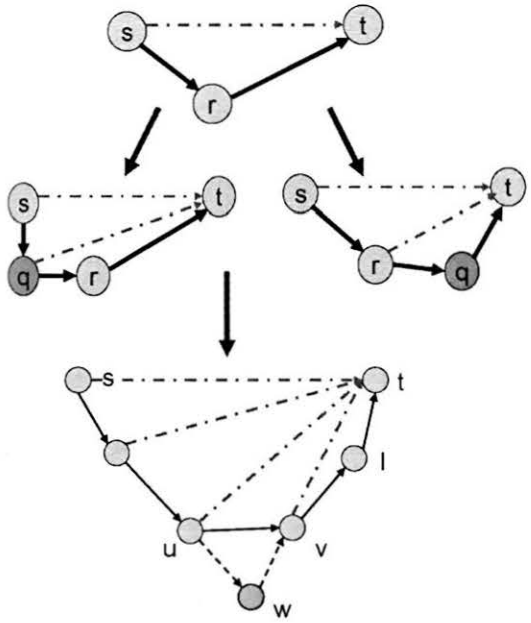
Technical Discussion: The LTE standard defines two pilots for the uplink: (i) the demodulation pilot, which is used for accurate channel estimation for coherent demodulation and is transmitted in the subcarriers assigned to a user, and (ii) the wideband sounding pilot, which occupies the entire system bandwidth (e.g., 5 MHz). Through it, the base station determines the full frequency response of each user's current channel and performs frequency-domain scheduling to get high cell throughput.

We have studied the use of the demodulation pilot and the sounding pilot for AS training. Using the spare sounding pilot transmitted at regular intervals for AS training enables the base station to also do joint frequency-domain scheduling and antenna selection. Our results show that antenna selection is extremely robust to the higher levels of interference the wideband sounding reference signal is expected to encounter.

Future Direction: Future work involves methods to reduce the wideband pilot overhead incurred by antenna selection.

Contact: Neelesh Mehta
http://www.merl.com/projects/AntennaSelection/ **Project Type:** Advanced Development

Progressive Accumulative Routing



We design a distributed and progressive protocol called Progressive Accumulative Routing (PAR) that enables energy accumulation at the destination, and enables the use of conventional relays that need not accumulate energy.

Background and Objectives: Energy accumulative routing is a novel concept that improves the energy efficiency of wireless relay networks. In it, relay and destination nodes store and do not discard the received signal of a packet that is too weak for decoding and combine it with copies of the same packet that arrive later. Destination energy accumulative networks are more energy-efficient than traditional multi-hop networks that do not accumulate energy. At the same time, they are also significantly less complex that complete energy accumulative routes that require each relay to store every copy of every packet en route in the network.

Technical Discussion: We consider energy-efficient unicast networks in which the destination accumulates energy, but the relay nodes do not. We propose, develop the fundamentals of, and analyze the Progressive Accumulative Routing (PAR) protocol that determines the best multi-hop route and exploits energy accumulation by the destination. PAR has a number of key properties that make it suitable for practical implementation in ad-hoc networks: (i) Progressive addition of nodes in an incremental fashion, which avoids having to tear down established routes every time, (ii) Distributed computation of route nodes and powers, which requires only local channel knowledge available at each relay node, (iii) Simple protocol structure, which involves the transmission of simple request for cooperation packets by nodes that can themselves determine whether they can reduce the current route's total energy consumption (iv) Large energy savings, which can often be almost as large as those achieved by routes determined by a centralized controller. Thus, the route discovery in PAR has a very low complexity, requires very limited feedback from nodes that can be added to the route as relays, and results in large energy savings.

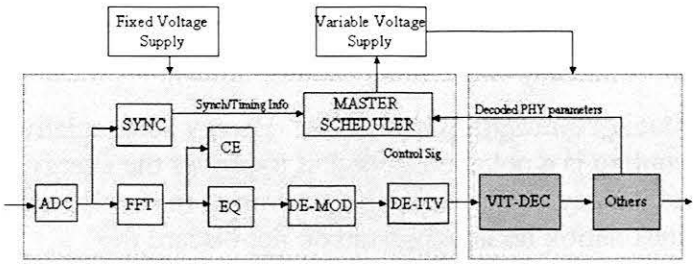
Collaboration: Prof. Raymond Yim, Olin College of Engineering, Needham, MA.

Future Direction: A multi-layer adaptation and optimization of PAR that accounts for not energy consumed for radio transmission but also transmitter inefficiencies and receive circuitry energy consumption, which can be significant for high node densities.

Contact: Neelesh Mehta
http://www.merl.com/projects/pa-routing/

Project Type: Research

Low Power MBOFDM-UWB Transceiver Design



Dynamic Voltage Scaling in UWB Transceiver Design

We investigated the power profile of the ECMA Multiband-OFDM standard compliant transceiver circuit. The research focused on the low-power, high speed FEC decoder. Our objective is to incorporate our research results into ASIC development.

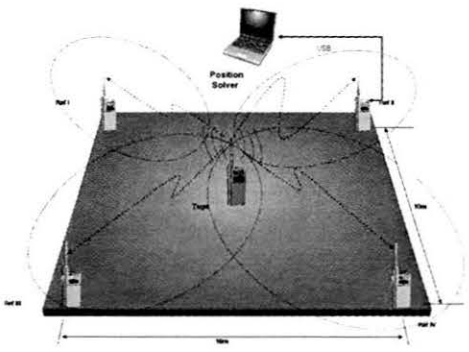
Background and Objectives: Multiband-OFDM has become the de facto technology for High data rate UWB and has been adopted as an ECMA standard. Due to the large bandwidth and high data rate, the average power consumption of current commercial UWB transceivers typically exceeds 400 mW. This is far from its potential ultra-low power consumption that UWB technology offers. Such high power consumption also limits the UWB technology from being used in certain applications such as in battery powered mobile nodes. The objective of this project is to provide the optimization in different design levels (network, algorithm, hardware architecture and circuit) and ultimately achieve a reduced total power consumption.

Technical Discussion: Our initial study generated the power profile for the transceiver baseband circuit. Our investigation shows that to achieve the maximum data rate of 480 Mbps, the power consumption and the size of the Viterbi decoder is expected to be more than 30% of the entire chip. We proposed two schemes that can lower the power adaptively based on the data rate of the received packet: a) Dynamic voltage scaling (DVS) in UWB receiver and b) Adaptive sliding block Viterbi decoder (ASDV). Both schemes exploit the variable data rate and the data packet structure in UWB standard. Using our proposed DVS design, we were able to achieve over 3X power saving at 55Mbps compared to a conventional design. We also compared the power and speed performance of different implementation options in individual components (e.g. Add-Compare-Select-Unit and Trace-Back-Unit) in the Viterbi decoder. The ASDV is an improve version of the sliding block Viterbi decoder that allows us to disable some portion of the circuit (and therefore save power) while still meet the throughput requirement.

Future Direction: We are working on low complexity/energy synchronization algorithm/architecture and cross-layer optimization of memory structure and other functional blocks.

Contact: Chunjie Duan, Jinyun Zhang
http://www.merl.com/projects/low_power_uwb/ **Project Type:** Research

Development of Location Centric Networks



Ultra-Wideband (UWB) Impulse radio communicate by using a sequence of short pulses. UWB can potentially provide a solution for low cost, mid/high data rate communication link. More importantly it has the advantage of being able to include ranging/location capability in each node. Due to the large signal bandwidth, UWB can potentially achieve high accuracy ranging and when incorporated with network functionality the ability to achieve accurate geo-location.

Background and Objectives: MERL has successfully built and demonstrated an UWB Impulse radio network that communicates data and produces ranging and location estimates between transceivers. The range estimates serve as data for a network positioning algorithm that then computes the location of the radios. The radios are based on the published IEEE 802.15.4a standard which MERL helped develop and contributed several essential patents. One of the challenges for impulse radio UWB is to implement such a system with low cost and high performance. The objective of this project is to demonstrate low cost radio location techniques.

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

Project Type: Research

Multimedia

Multimedia enriches our everyday lives through rich forms of audio-visual communication and entertainment. We experience multimedia through television broadcasts, radio, DVD and mobile devices, while surfing the Internet, and in our automobiles. Multimedia is also used extensively for security and factory automation systems.

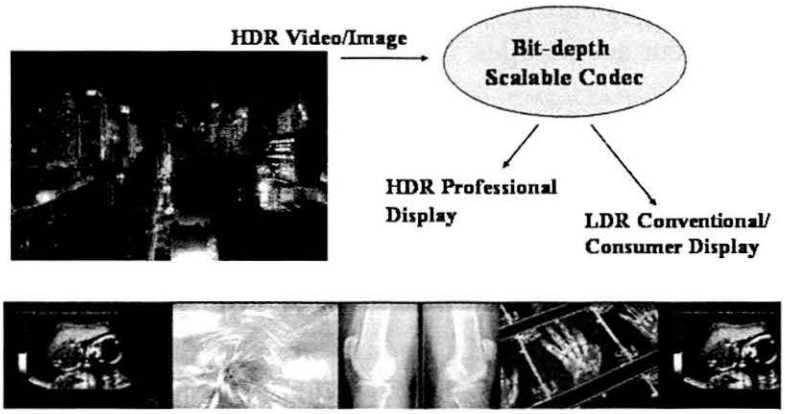
At MERL, our multimedia work is focused in several main areas of concentration, which are described further below. Efficient representation, transmission, processing and interaction of multimedia, are some of the underlying technical themes.

- Video Compression: Key activities target compression of rich video formats, e.g., multiple views, higher resolution, better color, and greater bit-depth. Our research results are applied to international standards and across a wide range of AV products.
- Information Coding: Current efforts are focused on application of distributed source coding principles to secure storage of fingerprint data. New applications related to robust media transmission, multimedia authentication, and low cost image sensors, are being considered.
- Speech/Audio Processing: The work in this area emphasizes spoken-language interfaces for automotive and handheld devices. There is also ongoing research on acoustic doppler-sensing for speech denoising and human gait recognition.
- Multimedia Platform: The primary aim is to enhance the capabilities of various multimedia platforms, including DTV platforms (noise reduction and AV networking), factory automation (JPEG 2000 transcoding), and automotive platforms (improved interaction).

In the pages that follow, we provide a brief overview of related projects in the above areas that we have been working on within the past year and have substantial results to report on.

Project Descriptions

Bit-depth Scalable Coding for High Dynamic Range Video



See Color Figure 1

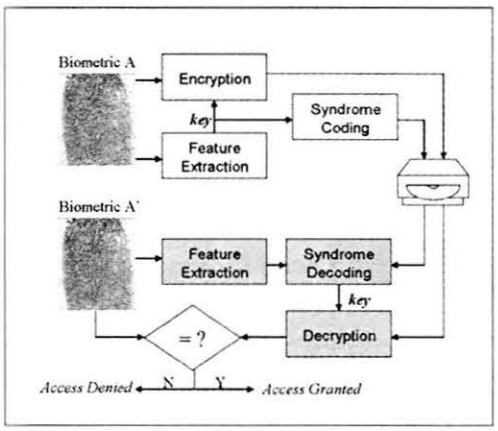
Conventional video sources are represented in 8bits per pixel (bpp) component, while many professional applications require video sources in higher bit-depths, i.e. 10, 12, 14, etc. bits per pixel component. Given a higher bit-depth input video, the compressed bit stream should consist of two (or more) layers, where the base layer contains 8 bpp content, which can be displayed on conventional 8 bpp consumer display devices; while the enhancement layer carries higher bit-depth content, which can fulfill professional application requirements. This multi-layer bit stream can be easily achieved by simulcast; however, the compression efficiency is very low.

Background and Objectives: The key issue in bit-depth scalable coding is to predict higher bit-depth pixels (e.g. 10 bpp) from base layer (e.g. 8 bpp) pixels, which is referred as inter-layer prediction. The inter-layer prediction can be considered as an inverse tone mapping procedure, whose performance is strongly related to the forward tone mapping method used to convert pixels from higher to lower bit-depth. However, the tone mapping method is not always known and it may have variations in many ways. For example, different tone mapping methods, such as linear or non-linear, can be used for various applications; it may also vary in different frames, scenes, or in different regions within one picture; it may vary in different color channel as well; and the dynamic range can also be big or small.

Technical Discussion: In order to cope with bit-depth scalability, several contributions have recently been proposed in JVT standard meetings. With extensive study of the prior arts, we developed a novel inter-layer prediction scheme and proposed it to JVT standard group. Our proposed solution is localized (block-based), which is more efficient for coding videos generated by local tone maps in different scenes, pictures, and regions of interest. The proposed solution is also independent to each color channel; thus, it provides the flexibility and efficiency to compress videos in various color spaces. Further more, our approach is suitable for a wide variety of tone mapping techniques, and can cover a wide bit-depth dynamic range. Experimental results show that that our proposed scheme achieves significant coding gain over other localized (block-based) approaches. Coding gains over global approaches on locally tone mapped videos are expected.

Contact: Shan Liu, Anthony Vetro
http://www.merl.com/projects/hdr-compress/ **Project Type:** Research

Secure Biometrics



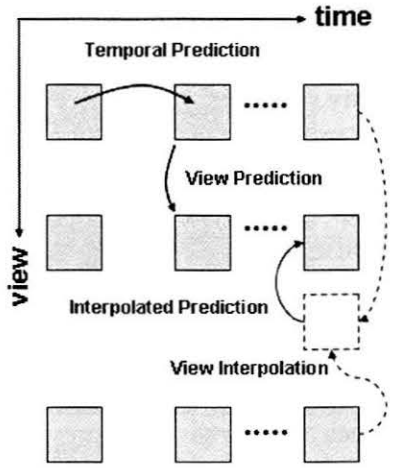
Current methods of using biometrics are often insecure since they store the biometric in the clear, compromising security and enabling identity theft. Our method obtains security by transforming the biometric into a syndrome (a compressed and scrambled bit stream that contains less information than the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

Background and Objectives: Biometrics such as fingerprints, irises, and faces are increasingly prevalent in authentication, encryption and access control. Biometrics are slightly different each time they are measured. Therefore they cannot be stored in encrypted form as passwords are because the encrypted form of the original biometric and the encrypted form of a later measurement of the same biometric would not match. Consequently, most systems store biometrics in the clear. For biometrics to be broadly accepted, we need a way to store biometrics in a secure form that cannot be used by an attacker to impersonate a valid user. At the same time the authentication method needs to be robust to the natural measurement variation of the biometric.

Technical Discussion: Our method obtains security by transforming the biometric into a binary vector which is then multiplied by the parity check matrix of a publicly known low density parity check code. The output is the biometric's syndrome, a compressed and scrambled version of the original biometric with two essential features. First, the syndrome contains less information than the original. If only the syndrome is stored, and not the original biometric itself, an attacker that learns the syndrome cannot recover the original biometric. Second, when the syndrome is combined with another measurement of the same biometric, it is possible to correct the measurement noise and exactly recover the original biometric through belief propagation decoding. The original biometric can therefore serve as a shared secret. The original biometric can be used, e.g., as a secret password or a cryptographic key, using standard techniques.

Contact: Anthony Vetro, Jonathan Yedidia
http://www.merl.com/projects/secure-biometrics/ **Project Type:** Research

Multiview Video Coding



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

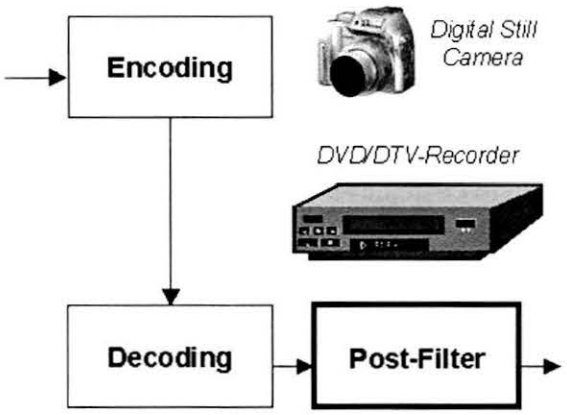
Background and Objectives: The need for multiview video coding is driven by two recent technological developments: new 3D display technologies and the growing use of multi-camera arrays. A variety of companies are starting to produce 3D display technologies that do not require glasses and can be viewed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

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

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

Contact: Sehoon Yea
http://www.merl.com/projects/multiview-coding/ **Project Type:** Research

Noise Reduction for Image/Video Data



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. Noise reduction 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.

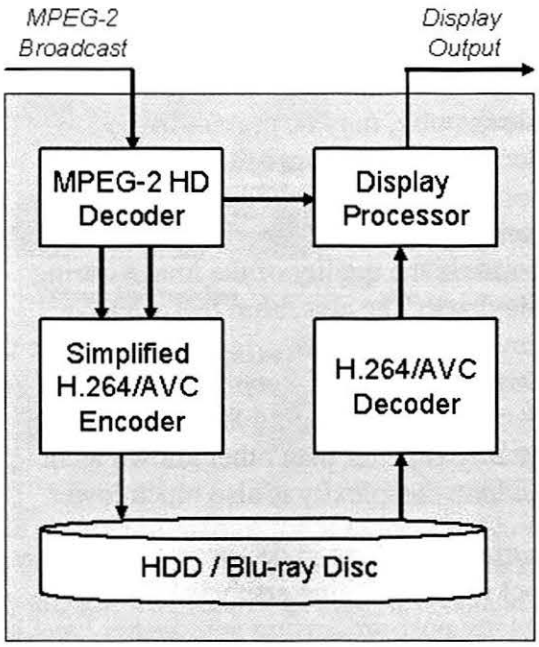
Technical Discussion: We have proposed a new adaptive approach for both blocking and ringing artifacts removal that applies pattern classification techniques to first identify different type of artifacts and then performs the filtering accordingly. Our strategy is as follows: 1) form an edge map based 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.

Future Direction: Continue to research new methods for image/video noise reduction.

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

Project Type: Advanced Development

Efficient Storage of Broadcast Video



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

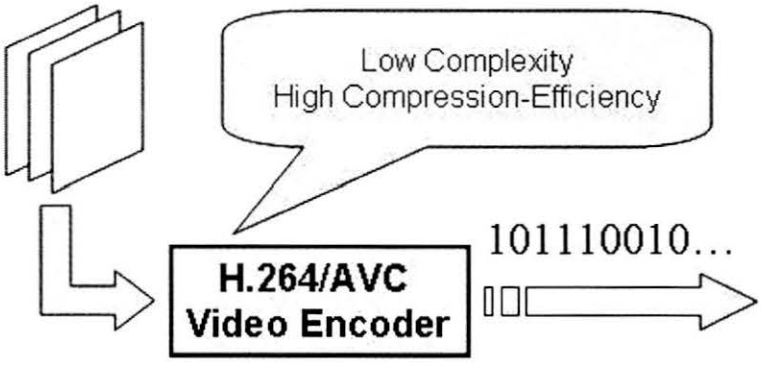
Background and Objectives: MPEG-2 is the primary format for broadcast video, where the data rate for high-definition video is approximately 20Mbps. The latest video coding standard, referred to as H.264/AVC, is able to achieve the same quality as MPEG-2 with about half the data rate. Since the H.264/AVC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

Technical Discussion: In this project, we aim to transcode the incoming MPEG-2 bitstream to an H.264/AVC bitstream. The key to reducing the complexity is efficient motion re-estimation and mode decision, which would typically account for more than 80% of a full H.264/AVC encoder. Our transcoder takes incoming motion vectors and coding modes from MPEG-2 bitstream and maps them to the motion and modes of H.264/AVC using a novel distance weighted algorithm. After mode and motion mapping, a small-range motion refinement around the mapped motion vector is performed. We also include support for transcoding to half-horizontal resolution output. Extensive simulations show that the our transcoder performs very well compared to the reference (cascaded) approach with a fraction of the complexity.

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

Project Type: Research

H.264/AVC Encoder Optimization



H.264/AVC is the state-of-the-art 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 being widely adopted for video conferencing, mobile TV broadcasting and high-definition optical disc storage. However, its encoding complexity is extremely high. The goal of this project is to develop low-complexity H.264/AVC coding techniques that can still achieve state-of-the-art video compression.

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

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

Contact: Anthony Vetro, Huifang Sun
http://www.merl.com/projects/avc-optimization/ **Project Type:** Research

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. In our past work, we have built a complete end-to-end 3D TV system that performs real-time acquisition, transmission, and 3D display of dynamic scenes. More recently, we have been participating in standardization on multiview video coding, and we have also developed an anti-aliasing method for 3D displays.

Background and Objectives: Today, digital broadcast networks carry hundreds of channels and will presumably be capable of carrying even more as the most advanced video codecs become deployed. 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. Our initial system demonstrated that 3D TV offers a richer, more immersive experience than regular TV. It increases entertainment value and realism without the encumbrance of special glasses. Recent work has centered on improving the quality of the experience further.

Technical Discussion: In our initial 3D TV system, image acquisition consisted of an array of hardware-synchronized cameras that captured multiple views of the scene. To deal with the high processing and bandwidth requirements, the system used a fully distributed architecture with clusters of PCs. A multi-projector 3D display with horizontal parallax was used on the output side. The system was scalable in the number of acquired, transmitted, and displayed realtime video streams. More recently, we have concentrated research efforts on anti-aliasing techniques that aim to improve the rendering of multiview video on 3D displays. View interpolation techniques are utilized to achieve an oversampling of the multiview signal in the view dimension. The oversampled signal is then filtered to suppress high frequency portions of the signal that contribute to aliasing, and finally sub-sampled to match the display characteristics.

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

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

Contact: Anthony Vetro
http://www.merl.com/projects/3dtv/

Project Type: Research

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 - For User Safety: By using DT Controls, Navigation system functions may be safely enabled for the passenger while the vehicle is in motion.

(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.

Contact: Bret Harsham, Jonathan Westhues
http://www.merl.com/projects/dtcontrols/

Project Type: Advanced Development

Safe and Easy to Use Voice Interfaces for Automotive

Simplicity is the key to safe, intuitive user interfaces for the automotive environment, and speech is a logical input modality for a telematics system because it allows for hands-on-the-wheel, eyes-on-the-road interaction. However, many state-of-the art automotive speech interfaces may be unsafe because their complex, multi-state, often system-paced spoken dialogs impose a high cognitive load on the user, and can therefore interfere with the primary task of driving. Effective speech interfaces should instead employ flatter interaction models with a minimum of states to remember, and should use physical controls rather than speech to carry out simple commands.

Background and Objectives: The skyrocketing amount of information and content available on both standalone and Internet-enabled automotive devices presents a challenge for MELCO's design and implementation teams. That is: how to provide users with the information and content they desire in a seamless, intuitive fashion, while at the same time minimizing driver distraction. MERL plans to develop and test new interfaces that will meet this challenge.

Technical Discussion: One such interface paradigm is "Speech In, List Out" (SILO), wherein a user searches a content database (for example, points of interest in a navigation system or music on a connected portable player) using simple, single-shot spoken queries, rather than entering into a confusing, turn-taking dialog with the system. SILO retrieval may be combined with concise, consistent speech commands and/or well-placed physical controls to create a driving experience that affords access to the latest information and content, while at the same time keeping both driver and passengers safe.

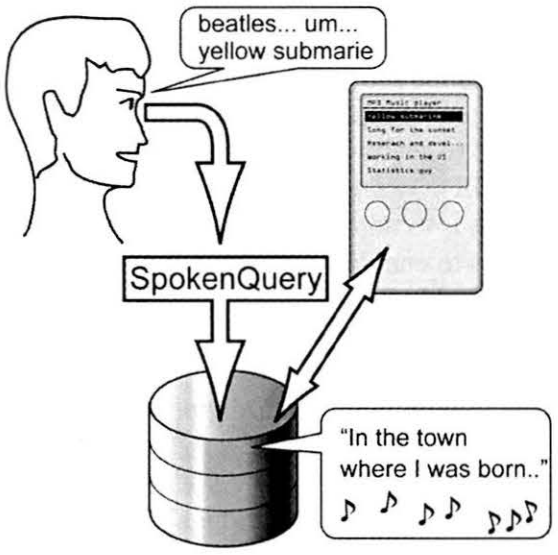
Other important work includes the extension of SILO's enabling technology, SpokenQuery, to handle the combination of query and command components within a single utterance. Also, several members of the group collaborated on a chapter entitled "Speech Based UI Design for the Automobile" for the volume Handbook of Research on User Interface Design and Evaluation for Mobile Technology (ed. Joanna Lumsden), to be published in late 2007 by Idea Group Inc.

Collaboration: UMass-Amherst HPL (Don Fisher et al.)

Future Direction: Future work may involve prototyping complete automotive interfaces, incorporating both visual, voice, and haptic input and/or output. Additional usability and driver safety studies should be carried out in concert with this.

Contact: Garrett Weinberg
http://www.merl.com/projects/vi4auto/ **Project Type:** Research

SpokenQuery



This year the Speech group at MERL has made numerous additions and improvements to the embeddable SpokenQuery (SQ) core and the R&D toolkit.

Background and Objectives: End-users often need access to information that they know is available on a device they're using, but their manual access is slowed due to interface limitations or because they are otherwise occupied. For example, it's tedious to scroll to a particular song on an iPod, and it's dangerous to do so while driving. With SpokenQuery technology, users access such information simply by speaking. The SQ project supports the delivery of easy-to-use, speech-based search-and-selection capabilities in embedded application environments.

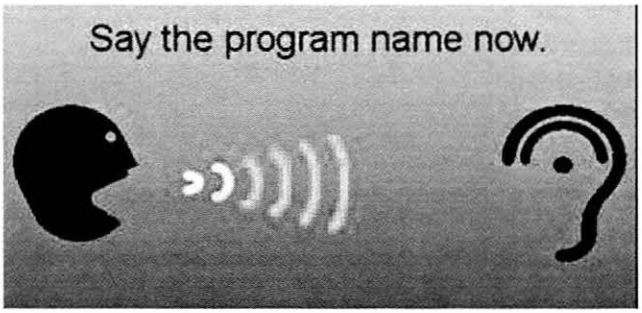
Technical Discussion: SpokenQuery enhances the usability of commercial, automatic speech recognition (ASR) engines. Using patented processing of ASR output, SQ allows for natural, free-form spoken requests for information. Using SpokenQuery is like using Google to locate something you need, but you say the descriptive words instead of typing them. Examples include iPod music selection, GPS destination entry and field service manuals. Our work is focused on the research, development, and delivery of core ASR and SQ technology, including fundamental research on the representation of documents and speech queries. We also collect and analyze speech data from realistic settings to support the development of interfaces with reduced distraction and improved usability.

We have deployed the SpeakPod prototype using the core C++ library on three embedded platforms, instrumented the library for performance measures, developed a performance and regression testing harness, and added new algorithms to the core library. Research efforts include deployment of the SpeakPod application on a real-world platforms so that we may analyze usage data from MERL employees.

Future Direction: Long-term research goals include improved task-completion and user-satisfaction rates, and safer and more effective information-retrieval interfaces. Our ongoing analysis of unscripted speech data will improve fundamental SpokenQuery modeling and search technology.

Contact: Bent Schmidt-Nielsen
http://www.merl.com/projects/SpokenQuery/ **Project Type:** Advanced Development

TV Content Search Using Voice



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

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

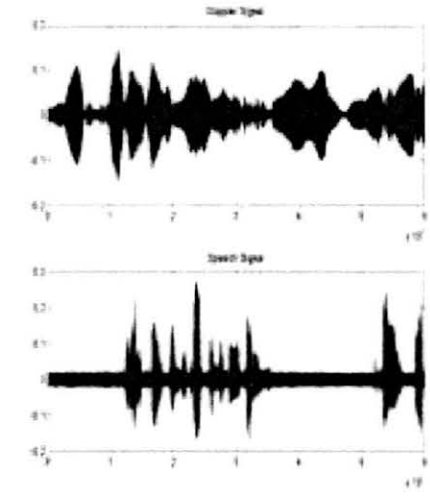
Technical Discussion: Our solution adds a microphone to remote controls that allows the user to quickly find programs or browse the nearest matches using voice. The system utilizes MERL's SpokenQuery technology in a new domain of EPG search. SpokenQuery allows search terms to be entered that are unrestricted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

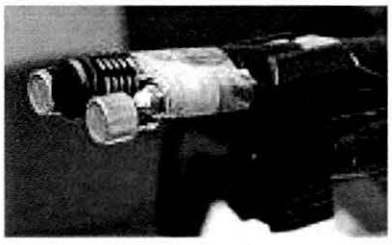
Future Direction: We plan to refine the prototype and conduct further user testing. Advanced voice browsing techniques will also be investigated.

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

Project Type: Research

Acoustic Doppler for Denoising Speech Signals





Acoustic Doppler readings provide measurements of the movements of a talker's face. These measurements will typically not be corrupted by the same noise sources that may corrupt a speech signal. They may hence be utilized to provide secondary evidence that can be used to denoise speech signals.

Background and Objectives: Speech-based devices and applications such as cellphones and kiosks are frequently used in very noisy environments. Denoising techniques that directly work on the speech signal are often ineffective in these environments. Their performance can be greatly enhanced through the use of secondary sensors that measure other characteristics of the speech that do not get affected by the noise. Such secondary sensors are, however, highly expensive. Our goal is to develop an inexpensive, but effective secondary sensing mechanism through Acoustic Doppler to effect the denoising.

Technical Discussion: Our goal is to denoise speech signal for improved coding, transmission, recognition etc. Any speech activity is accompanied by corresponding movement of facial features such as lips, cheeks etc. These movements are correlated with the speech signal. Any measurements of these movements that are not corrupted by any noise that is correlated to the noise that corrupts the speech signal can hence be used to restrict the space of possible values for parameters derived from the speech signal. These restrictions can further be employed for improved denoising of the speech signal. We derive our measurements of the movements of the talker's face through a acoustic Doppler radar. We incident a 40Khz tone on the talker and capture the reflections. These reflections are FM demodulated. Finally thet joint distribution of the Doppler and audio signals is modelled by a time-series model. On noisy data, predictions of the value of audio features are obtained from the Doppler and used to denoise the speech.

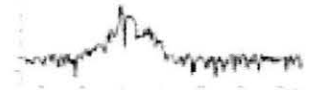
Future Direction: We are in the process of developing more detailed models for the joint distribution of speech and Doppler. Future work will be the completion of this task and the development of a real-time denoising mechanism using the Doppler sensor.

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

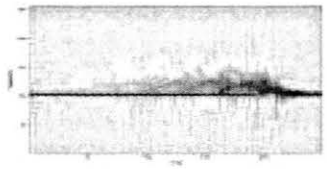
Project Type: Research

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

Acoustic Doppler measurements may be utilized to characterize walkers and to recognize them.

Background and Objectives: Gait is a strong characteristic of individuals, and may be used to recognize them. Traditionally Gait has been characterized through video recordings. Subjects are detected, segmented out, and their image characteristics measured either directly or through stick models etc. These methods work best when the subject is walking at right angles to the vector to the camera. In this project we aim to develop a different acoustic Doppler based mechanism to characterize gait. The Doppler-based sensor will be cheaper than conventional video based sensor, and have complimentary capabilities.

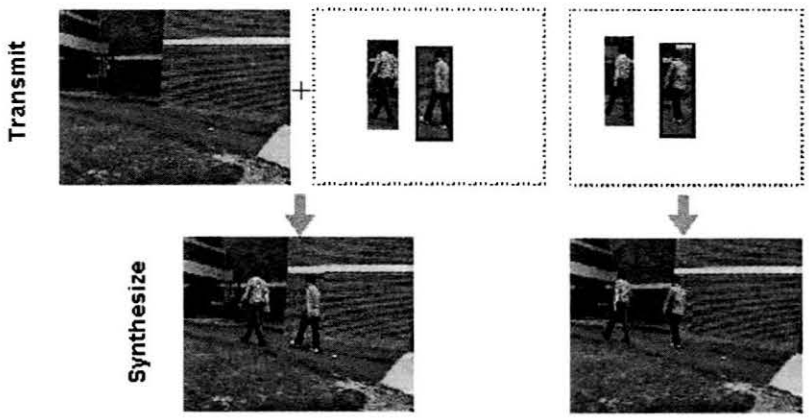
Technical Discussion: Gait consists of a sequence of movements of the walker's limbs. During this movement, various parts of the body, such as the feet, shin, knees, thighs, arms, elbows etc. progress through a cycle of movements in which their velocity also changes cyclically. The ensemble of cyclic variations of these velocities characterizes the gait. If a high-frequency tone is incident on a walker, the frequency of the reflected signal get modulated by the velocities of the various moving body parts. The set and pattern of modulated frequencies in the reflected signal characterize the gait. We capture the reflected signal through a resonant transducer. The signal is frequency demodulated to enhance all frequency components. A series of cepstrum-like feature vectors are derived from it. We learn a Gaussian Mixture model for each subject from a short recording of training data. Walkers are recognized thereafter by regular MAP classification. Our results indicate that we are able to identify subjects with over 90% accuracy using this approach.

Future Direction: Results indicate that we can not only identify subjects, but also classify them according to height, gender etc. Further research is being conducted in this direction. We have also developed a new time-series model to represent temporal patterns in the Doppler signal. These will also be investigated.

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

Project Type: Research

Scalable Streaming of JPEG 2000



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

Background and Objectives: Smart surveillance systems will require the latest techniques most efficient storage and transmission of images. Many new systems employ JPEG 2000 for image compression processing to gain a better understanding of objects in a scene. To satisfy bandwidth constraints and display important parts of the images fast and precisely, the compressed images are efficiently transcoded in the compressed-domain based on the ROI information.

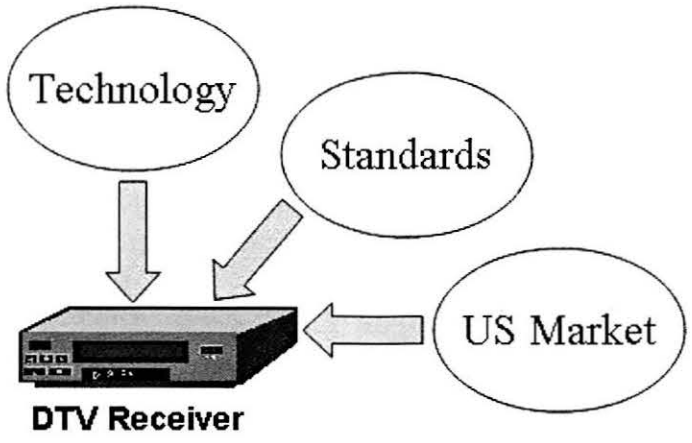
Technical Discussion: In our system, an image sequence is encoded and stored as a JPEG 2000 bitstream, and then the stored images are efficiently transcoded in the compressed-domain, to reorder the image stream and achieve a wide range of effects and progressions. In one effect the regions of interest (ROIs) are preserved while higher quality portions of the background are replaced with empty packets to satisfy network constraints. We have designed a dynamic rate control algorithm that achieves high bandwidth utilization, hence better overall quality, while minimizing the quality variations over time. Another technique allows the order of transmission to differ from the ordering of the stream, thereby achieving flexible progressive transmission. Other possible effects include resolution reduction, component reduction and cropping functions.

Future Direction: We are considering additional scalable progressive transmission techniques.

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

Project Type: Advanced Development

DTV Standards in the US



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

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

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

Contact: Anthony Vetro
http://www.merl.com/projects/dtv-standards/ **Project Type:** Advanced Development

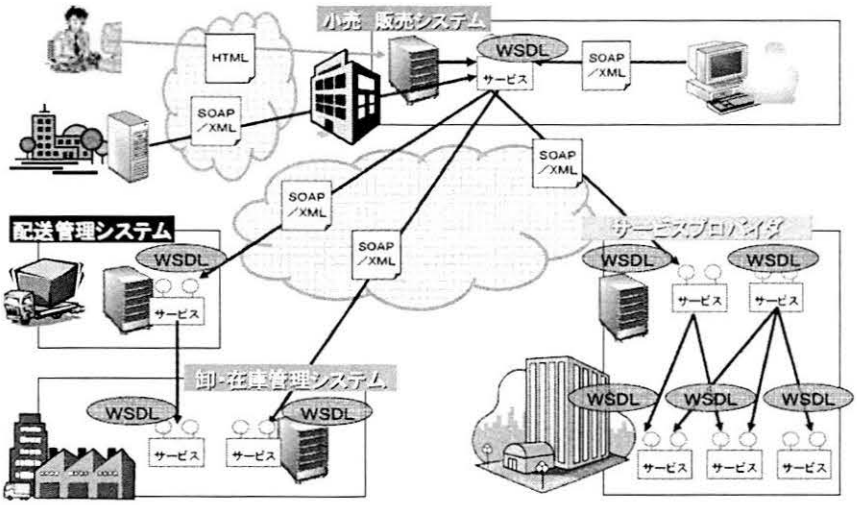
Data & Sensor Systems

Traditional applications using sensor data have largely presented the user with content that requires a great deal of human intervention and interpretation using conventional keyboard and mouse interactions. The gap between when the data was collected and when the system or human could respond was a long one. However, with the advent of ubiquitous, low-cost sensors and the reduction in costs of computing and networking, it is now becoming possible for a computer system to quickly and automatically collect data from the local environment, process that data adaptively, and then respond to changing conditions either automatically or through human intervention with sophisticated visualizations and presentations. These capabilities are driving new applications in sensor networking, data mining, and ubiquitous computing. New human-computer interaction methods are needed to drive many of these applications and the computational devices used.

MERL's work in "Data & Sensor Systems" is creating new technologies for this exciting area, ranging from fundamental ideas to create and support the technology itself to applications aimed at growing new businesses and enhancing current ones. Projects span the gamut from sensor hardware, which can autonomously and inexpensively collect data, to new systems for communicating that data, to algorithms for uncovering subtle underlying interrelations and trends in the data, to new methods for interacting with the data through novel input/output devices.

Project Descriptions

SOA Integration Technology



We have initiated work on a support tool for integrating existing Web applications designed for humans into fully-fledged web services that can be used within standard service oriented architectures (SOA).

Background and Objectives: SOA applications are a new style of applications that extend the client-server model to a massively distributed collection of loosely coupled web services that can also be re-used for many other applications. Although there is a huge number of existing Web applications that serve all kinds of useful purposes, these applications are meant to be used by humans, and are not standards-based web services designed to be used by other computers. It is highly desirable to be able to convert such Web applications into web services by creating a standard wrapper, ideally with none or minimal human intervention.

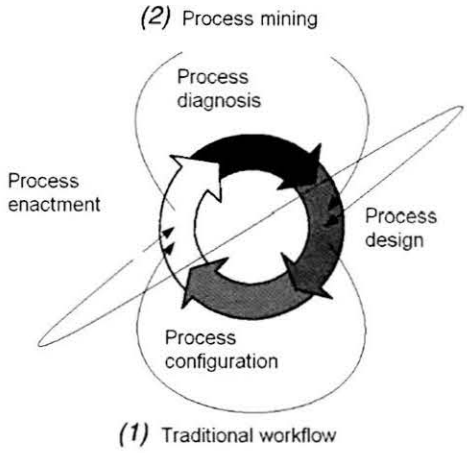
Technical Discussion: The technical challenge is to wrap an existing Web application into a web service, without actually understanding the semantics of the operation of this Web application, or having access to its code. While performing this task in a completely automated manner might not be possible, our goal is to create a semi-automated wrapping tool. Our approach is to analyze the web server (e.g. Apache) log files that document the interaction of humans with a Web application, and extract the protocols and interaction sequences from these logs, so that they can be documented in standard WSDL and UDDI specifications for a web service.

Future Direction: Technically, we are addressing the fundamental problem of state-less vs. state-preserving services and applications. We are also working on mashup demos involving a mixture of original web services and converted Web applications.

Contact: Daniel Nikovski, Alan Esenther
http://www.merl.com/projects/soa/

Project Type: Initial Investigation

Business Process Management Software



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

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

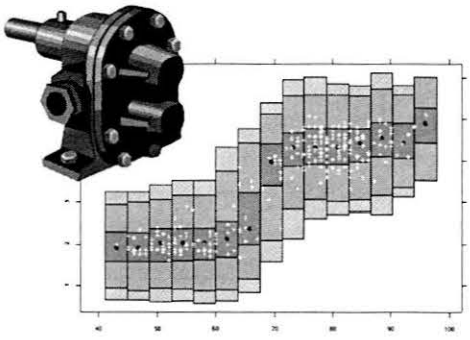
Technical Discussion: To a large degree, the two modules are complementary: while the simulator produces synthetic execution logs from a given BP model, the process mining software produces a synthetic BP model given actual execution logs. In spite of this complementarity, however, the technology and technical challenges behind the two modules are completely different. The discrete event simulator uses efficient random number generation and management of event sequences in priority queues, while the main challenge in process mining is to select the most appropriate BP model among those that successfully explain the execution log. Since there are exponentially many such models, the computational complexity of existing algorithms has been very high. In contrast, MERL's novel algorithm for process mining uses a structured representation of business processes called Workflow Trees, and has complexity only cubic in the number of tasks in the model. To our knowledge, this is the only existing process mining algorithm with (low) polynomial complexity.

Future Direction: The project has completed, and we are working on bringing the two modules to existing customers. The process simulation module has been extended by MELCO to handle time-varying processes, and is currently being deployed at the call center of one MELCO business unit. The process mining algorithm is undergoing comparison with other mining algorithms within the ProM framework from TU-Eindhoven.

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

Project Type: Research

Equipment Condition Monitoring



Recent advances in wireless networks and embedded computational devices have made it possible to monitor cost-efficiently and in real-time the vast majority of the electromechanical devices. This has opened the possibility for timely fault detection, diagnosis, and even prognosis of future abnormalities. Our objective is to develop fully automated machine learning algorithms for building probabilistic models of normal operation from normal data, and classifiers and predictors for fault disambiguation and prognosis.

Background and Objectives: We have employed various machine learning methods for fault detection and diagnosis, and have achieved success with memory-based learning algorithms such a locally-weighted regression. Furthermore, we have developed original algorithms for fast abrupt change detection in sensor data streams based on the same memory-based approach.

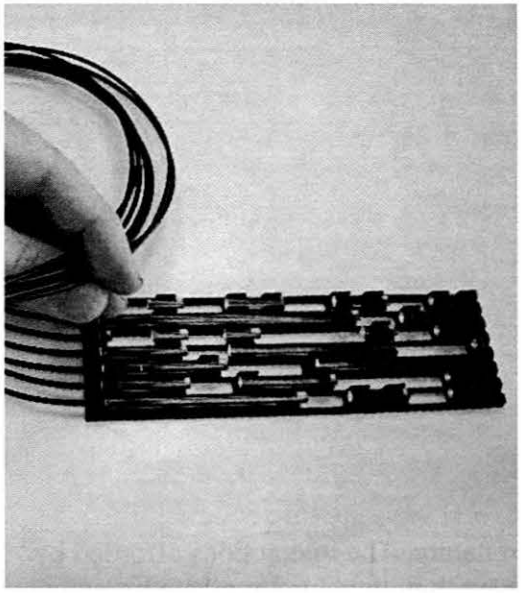
Technical Discussion: The readings produced by sensors attached to electromechanical machines are random variables that fluctuate constantly depending on the operating mode of the equipment. The main technical challenge in equipment condition monitoring is to distinguish those normal variations from deviations due to abnormal behavior and/or faulty operation. The problem reduces to learning probabilistic models of conditional densities, where the conditioning is upon external driving variables, and test robustly for deviations from these densities in real time. We have been pursuing a memory-based learning approach, and have been able to exploit the repetitive computational structure of memory-based density estimates to propose novel algorithms for abrupt anomaly detection in sensor data streams that have excellent computational complexity. Our algorithms MB-GT and MB-CUSUM have complexity only quadratic in the size of the memory buffer of sensor readings, and allow sensor stream monitoring in real time. They have been implemented in C and are available for operation in embedded systems.

Future Direction: We are continuing work on expanding this technology to high-dimensional data streams, typically generated by sensors attached to rotating machinery. If successful, such technology would allow us to address the problem of monitoring of motors and generators and all their industrial uses, such as elevators, compressors, pumps, etc.

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

Project Type: Research

Fluid Level Encoder



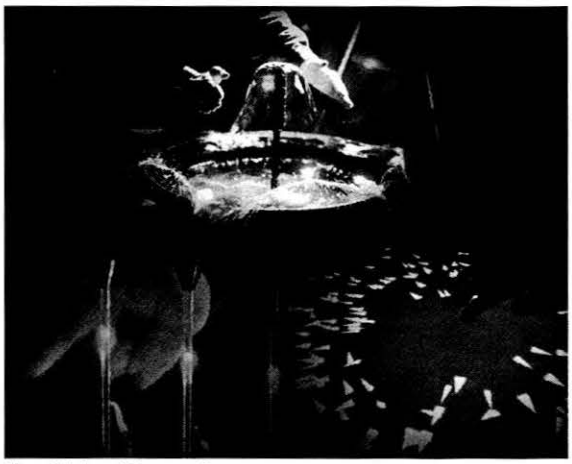
Fluid level measurement is of vital importance in many circumstances. As such, many techniques have been invented to measure fluid level. We have invented a fundamentally new technique which is analogous to rotary optical encoders. In this case, we build an optical structure which blocks the transmission of light only when the fluid is within a certain range of levels. These structures can be serially stacked to create encoder channels which respond to fluid levels in a plurality of ranges. Multiple stacks can then be combined to create incremental, absolute, or any of a variety of standard encoder topologies. This provides a direct, calibration-free, digital measurement of fluid level that can be mass produced at minimal cost.

Background and Objectives: There are many methods of measuring fluid level. One of the most elegant uses a cylindrical lens which comes to a point. In air, light sent through the lens is reflected back by total internal reflection. However, when immersed in fluid, the difference in index in refraction gets smaller, allowing the light to escape, so little is returned. The one draw back of this otherwise elegant technique is that a separate sensor is needed for each detection level. Our goal was to create a method which had similar desirable properties (inexpensive, reliable, explosion proof, no moving parts, etc.) but that only needed log2(n) sensors to detect n levels.

Technical Discussion: The key to the optical fluid level encoder is to create a structure which blocks the transmission of light only when the fluid is within a certain range of levels. If you imagine a flashlight held under water, perpendicular to the surface, you would be able to see much of the light escaping through the surface. As the angle of the flashlight is changed, the amount of light escaping decreases until the critical angle is reached, at which point all of the light is reflected off the air-water interface and none escapes. However, if one were to place a glass rod along the line of the beam, the light would then be able to escape, following the clear path of the rod through the surface. If we instead had sections of rod all lined up along the beam path, light would be able to pass through, unless the water surface lay in between rod segments. This creates an encoder channel.

Contact: Jonathan Westhues
http://www.merl.com/projects/fluid-encoder/ **Project Type:** Research

Submerging Technologies



See Color Figure 3

Submerging Technologies consists of three different interactive water sculptures: the Tantalus Fountain, which withdraws when a hand comes near; the AquaHarp, which is a musical harp with strings made of water; and the TouchPond - a liquid touchscreen. All three of the pieces exploit the electro-optic properties of water to demonstrate innovative sensing techniques.

Background and Objectives: Submerging Technologies was an unusual project for MERL. It began as a small demonstration of new sensing and interaction techniques. However, it quickly became apparent that these displays are a uniquely compelling mix of art and technology with wide appeal beyond the scientific community. Flowing water has long been a symbol of the tranquility of nature. The interactions afforded by our display create a sense of technology empowering us to work with nature to create beauty. This fits well with the image we have of MELCO and we feel that this project could become a vehicle to help express this understanding to the public. Submerging Technologies was first shown at SIGGRAPH 2006, where it was enthusiastically received, and was frequently cited as the best work at this prestigious conference. Since then, we have been invited to display at numerous art shows, museums and technical conferences around the world.

Technical Discussion: All three of the pieces use the water itself as a key element of the sensing system. The Tantalus Fountain uses the water as an electrode of a capacitive proximity sensor, and modulates the pump speed to maintain a distance from an approaching hand. This works well because the laminar flow yields consistent electrical properties. The dynamics of the trapped air in the water bell allow the user to sculpt wildly flared shapes. The other two pieces depend upon the total internal reflection of light inside a waveguide. In the case of the AquaHarp, LEDs shine down the water streams, carrying the light to sensors at the bottom. Breaking or deflecting the stream breaks the optical path which triggers a music synthesizer. The TouchPond is a similar idea, except the light guide is a sheet of water. IR LEDs shine parallel to the surface, and stay contained by total internal reflection until a disruption in the surface causes some light to scatter. A camera detects this, and suitably reacting graphics are rear-projected on the underside of the pond.

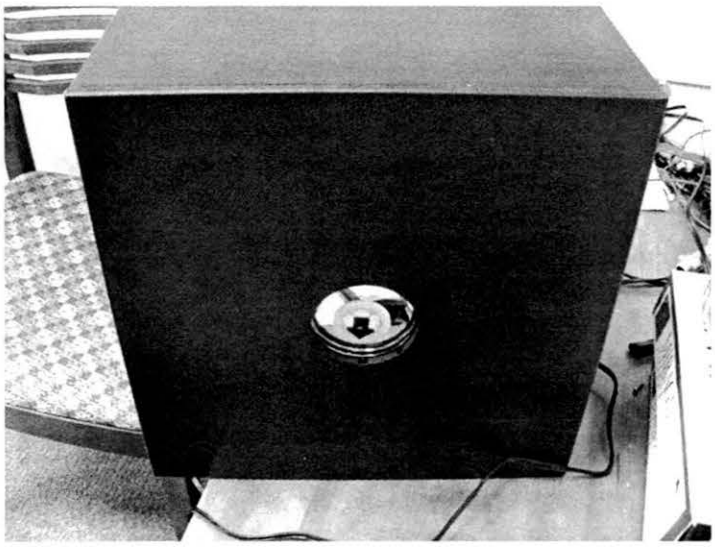
Collaboration: The TouchPond was created by Jeff Han at New York University based on a suggestion from MERL.

Future Direction: It is our hope that our interactive water sculptures will be widely enjoyed at venues around the world and serve to promote the MELCO brand.

Contact: Jonathan Westhues
http://www.merl.com/projects/submerging/

Project Type: Research

Sanitary Interface



Elevator buttons and other public interfaces provide a path for the spread of communicable diseases. MERL has developed a technology which allows for the creation of inherently sanitary interfaces which are intuitive and fun to use. The system creates the 3-D illusion of a button floating in space using a pair of parabolic mirrors. Sensors detect an attempted touch, and the virtual button image is moved to give the illusion that a button is being pressed. In addition, a compressed air supply provides a brief pulse to give tactile feedback.

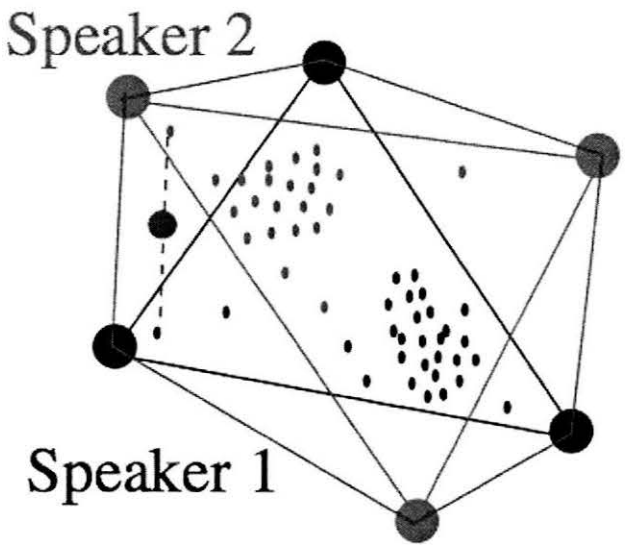
Background and Objectives: MERL has been exploring numerous ways to improve elevators ranging from new scheduling algorithms to advanced sensors that anticipate elevator trips. This project focuses on the user experience of the elevator. Public interfaces that must be touched by large numbers of people provide a path for the spread of communicable diseases such as influenza and SARS. We wish to provide an interface that is as familiar as a standard call button, yet inherently sanitary. By creating the illusion of a button, users intuitively know how to use it. In addition, the illusion itself is fascinating - most people find it magical - and it is likely that the first commercial elevators to sport this feature will attract significant attention from the public.

Technical Discussion: The combination of a floating image which moves with haptic feedback provides a compelling experience. However, to save cost, it may be desirable to remove features, such as haptic feedback. Similarly, the parabolic mirrors provide a very high quality illusion, but they are somewhat large and costly. Holographic and lenticular displays could also be used, albeit with somewhat lower quality.

Contact: Jonathan Westhues
http://www.merl.com/projects/sanitary/

Project Type: Research

Audio Separation



Schematized representations of speaker spectra in a latent variable space

See Color Figure 4

We are developing new approaches and algorithms to solve the problem of source separation. Although we focus on the problem of audio mixtures our work is directly applicable to multiple types of signals ranging from audiovisual, to biomedical/chemical, to vibrations and others. Our focus is to enable processing on non-clean data to not be influenced by interfering sources.

Background and Objectives: As is often the case audio signals are captured with interference from other sources. Since most time-series algorithms are designed to work on a single source signal, this interference results into suboptimal performance. This problem becomes especially prevalent when dealing with systems that perform speech recognition, or when users need to evaluate noisy data (e.g. a noisy cell phone recording). Our objective is to investigate methods with which we can perform various tasks on multiple sound recordings as well as we can with single sound ones.

Technical Discussion: We have recently presented a sequence of papers which describe some new techniques for source separation based on latent model decompositions. We have fine-tuned our techniques to work on source separation problems and our results are very competitive with the state of the art.

Future Direction: There are currently multiple ways to move ahead with our work and we are actively investigating multiple extensions and applications. We expect to make sever more contributions to the field in the coming year.

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

Project Type: Research

Video Summarization for PVRs

Personal Video Recorders have increasingly large storage capacity extending beyond 100 hours of content. Video Summarization is therefore essential to enable the consumer to skim through the content and view the content in differing detail depending on preference. We have developed a suite of summarization algorithms that are based on rapid audio-visual analysis in the compressed domain, and work well across diverse content genres. We are investigating the usability issues in a practical PVR application as well. We define tasks such as browsing a collection of programs, skimming through a program, watching the highlights, etc., and how summarization and video segmentation technologies apply.

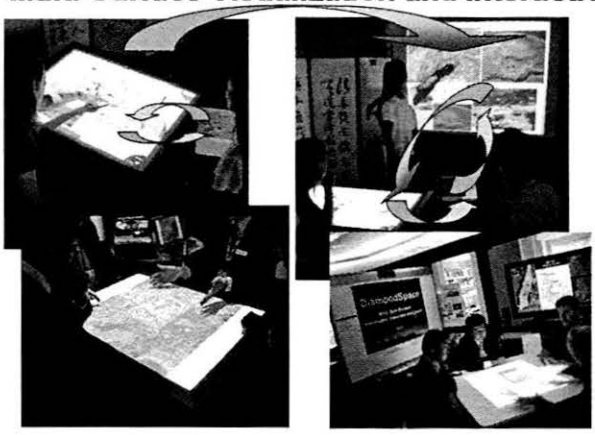
Background and Objectives: In this project we emphasize the Personal Video Recorder application, which provides the user with the content he wants when he wants it by storing a large volume of content recorded from broadcast and then providing effective navigation of the stored content using summarization and indexing. Our summarization algorithms are based on compressed domain analysis of both the audio and the video. Since such analysis is fast, our algorithms have been easy to realize on our target platforms. Our target products include personal video recorders such as DVD recorders. Our sports highlights detection was featured in MELCO's DVR-HE50W DVD Recorder, the world's first DVD recorder with sports highlights playback. We also study the usability issues and integration of summarization and browsing with PVR user interfaces.

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. ;Recent progress on this front includes the development of a genre-independent scene-change detector that uses a support vector machine classifier trained on hand-labeled examples of scene changes from a wide variety of broadcast genres. The second front is the usability of video summarization and browsing in PVR applications. The technology should be seamlessly integrated with the typical tasks that a PVR user have, such as browsing through a 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.

Future Direction: We will refine our genre-independent framework and work to incorporate new types of audio and video features into our scene change classifier. We will develop a holistic approach to application of video analysis, summarization and browsing technologies to PVRs that centers around user tasks.

Contact: Ajay Divakaran
http://www.merl.com/projects/VideoSummarization/ **Project Type:** Advanced Development

DiamondSpace



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

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

Technical Discussion: Our research is situated within a three-pronged agenda. (1) Multi-surface Visualization and Interaction (MVI): developing three distinct conceptual models of data sharing, visualization, manipulation and application interactions across multiple surfaces: independent (e.g., MultiSpace), reflective and coordinated/multi-view. (2) Interaction Techniques: designing and developing novel interaction techniques including freehand touch gestures, as well as interaction techniques that span across multiple displays and interaction surfaces. (3) Evaluation: We conduct evaluations, exploring the efficacy of new techniques, as well as the fundamental perception and tactial properties of touch surfaces.

Collaboration: Harvard University's Initiative in Innovative Computing Center; Ravin Balakrishnan, University of Toronto.

Future Direction: We have started to incorporate very large data walls into the DiamondSpace research investigation. Interacting with large data walls presents new challenges including large display area beyond arms' reach, visual and perceptual distortion, and multi-user collaboration.

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

Project Type: Research

Scalable Activity Recognition for Sensor Networks

This project demonstrates that with an appropriate analysis methodology, motion-sensor networks are capable of providing useful contextual information to building services. The methodology we have developed offers scalability and robustness by adopting a probabilistic and hierarchical framework. We call this framework Scalable Activity Recognition for Sensor Networks (SARSEN).

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

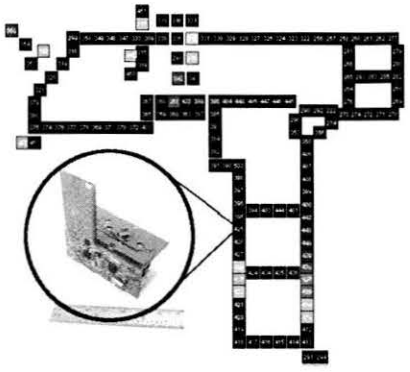
Interacting people generate ambiguities. Denser crowds generate deeper ambiguities. Ambiguities can lead to erroneous interpretations if an attempt is made to resolve them too early, with too little information. The SARSEN methodology instead delays resolution as long as possible, making the best local decision possible, but explicitly representing ambiguity that remains, so that it may be resolved at higher levels of context or larger scales of space, where disambiguating context may be available.

Technical Discussion: Empirical results show the methodology to be sound. Results include robust localization of meeting rooms, gathering points, and building resources such as printers. We also show the ability to recover interconnectedness in a building and then track changes in that social network over time using probabilistic techniques to accumulate evidence of behavioral patterns from collections of ambiguous motion traces. These results prove that the SARSEN framework is capable of correctly analyzing the complex behavior in our laboratory despite significant ambiguities. These results are fed by a database of over 30 million motion events collected from hundreds of sensors at MERL over the course of the year. sensors are the product of the Reduced Operating Cost Sensors (ROCkS) projects.

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

Project Type: Research

Ambient Intelligence for Better Buildings



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

Background and Objectives: The falling cost of sensors, microprocessors, and radios combined with the rising interest in safe, secure, and efficient buildings leads us to believe that holistic building systems are the future of building management. Many groups across the globe are focused on the problem of better packet routing, better sensing modalities, and lower-power computing. At MERL we are leveraging that work by focusing on the perceptual question: what is it possible to sense with these systems, what are the fundamental structures of human activity at the building scale, what can we learn, and what benefit can we gain? We have assembled a host of experiments and demonstrations that illustrate the usefulness of these systems for security, efficiency, safety and situational awareness within buildings.

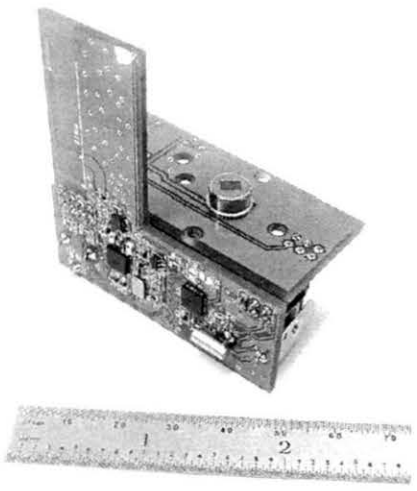
Technical Discussion: These results rest on the technical innovations of several other projects. Ultra-low power, easy to manufacture and easy to install, the Reduced Operating Cost Sensors (ROCkS) platform has enabled a host of experiments by allowing us to collect data from large, multi thousand square meter installations, over year-scale spans of time to assemble a unique dataset about human behavior at the building-scale. The Scalable Activity Recognition for Sensor Networks (SARSEN) project has focused on finding structures in that data, and probabilistic models to detect and analyze those structures. The Integrated Event Recognition project has focused using sensor networks to help solve problems that are hard given only classical sensor modalities.

Future Direction: We are pursuing several technical and business directions for the project. Technically, we are pursuing more detailed models of crowds, further refinements and validation of the ROCkS platform, and deeper interaction between sensor networks and classical modalities.

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

Project Type: Research

ROCkS: Reduced Operating Cost Sensors



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

Background and Objectives: The ROCks project has demonstrated the feasibility of low-cost, ad hoc installation of sensor networks within buildings by utilizing standards-compliant wireless communication and multi-year battery life. The modular architecture and easy to manufacture design allow the ROCkS to serve as a flexible platform for sensor network research at MERL and MIT.

Technical Discussion: The ROCkS provide a great deal of flexibility by replacing a platform licensed from MIT. ROCkS, the second-generation sensor box being designed at MERL, improves on the MIT design in terms of power consumption, radio compatibility, and manufacturability. The ROCkS were designed specifically with power consumption in mind. An MSP430 microcontroller was chosen for its flexible clocking modes, which allow the software to enter various "sleep" states and thus conserve power. This low power consumption profile means that the sensors can operate of small batteries for years at a time, and opens up to possibility of investigations into parasitic power models. The ability to install networks of sensors without the need for wires to deliver power represents a significant reduction in the cost of installation for the networks. Over the last year MERL has installed sensors in several locations, including a home where wired installation was not possible. MERL also worked with MIT to install a 150 sensors network that was used for social awareness research. The new boxes employ an IEEE 802.15.4 radio; this is the physical layer typically used with Zigbee. These new radios are similar in concept to the radios used in the MIT boxes, but they are standards-compliant and therefore capable of interoperating with equipment designed by many different vendors. The board layout is designed to be easy to manufacture and assemble, further reducing total cost of deployment. Modular design also means that experiments can leverage the work on the processor and communication systems for a multitude of other sensor modalities.

Future Direction: Currently we are working to deploy approximately 200 sensors at SIGGRAPH Emerging Technology to gather data in crowded environments.

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

Project Type: Research

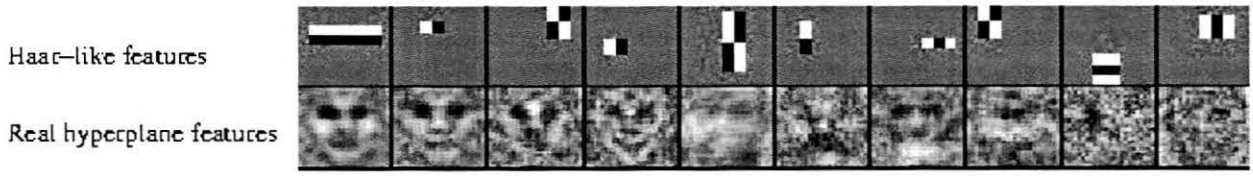
Imaging

The research in the Imaging group at MERL covers all aspects of extracting information from images. For instance, from a picture of a face we could calculate a numerical code for that face that would allow recognizing that person again in another picture. Or we could track a moving object in video to quantify its trajectory. In some cases we might modify the actual image creation process to make subsequent information extraction more effective as when we use multiple flash exposures to identify an object's edges. In other cases we might combine information from cameras with information from other sensors like when we search a historical database from a network of motion sensors to access stored video that documents the motion at the time and place specified.

For several years MERL has contributed the products in the security and surveillance area. MELCO has introduced an access control device based on the face detection and face recognition algorithms that were developed at MERL. Several other MELCO surveillance products use tracking algorithms that originated in MERL. Most of the object detection and tracking research heavily uses machine learning algorithms and image processing. Recently we have been getting more involved in the measurement of shape - a fundamental aspect of problems in robot vision and medical imaging. In the last year we also did research pushing the boundaries of camera image forming by introducing structure in the details of the camera exposure time or aperture, allowing recovery of richer image information.

Project Descriptions

Face Detection using Real-valued Hyperplanes



The goal of this project is to take advantage of special hardware that can quickly compute dot products in order to speed-up a face detector. The current face detector developed by Viola and Jones uses simple Haar-like features as the building block of the classifier. These were chosen because they are fast to compute in software. However, using special hardware, more complex features such as real-valued hyperplanes can be quickly computed. It should be possible to use these more complex features to improve the speed or accuracy of the face detector. Our results show that we can speed up the face detector by about 20% using real hyperplane features while maintaining the same accuracy.

Background and Objectives: Our objective is to take advantage of special hardware to improve the speed or accuracy of MERL's state-of-the-art face detector. The Universal Classifier project is building special hardware that runs the Viola-Jones type detector on an FPGA board. This work raised the possibility of using more complex features in the detector. This project at MERL is exploring the improvements that can be obtained using more complex features.

Technical Discussion: The difficulty with using real hyperplane features in place of Haar-like features in the Viola-Jones detection framework is that the set of possible features grows exponentially. It is no longer possible to use a simple feature selection algorithm that searches over all possible features. Instead we replaced the brute force search algorithm with a gradient descent algorithm that uses hill climbing to search the space of possible real hyperplanes. The disadvantage of this algorithm is that it can get stuck in bad local minima. We found that in practice if we use multiple starting points for the search, we find good hyperplanes in most cases. A frontal face detector that uses 10 real hyperplane features and 1520 Haar-like features achieves the same accuracy as one that uses just Haar-like features, but it runs about 20% faster.

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

Project Type: Research

Fast Human Detection

We have a developed a human detector algorithm that is as accurate as state-of-the-art detectors, while being about 70 times faster.

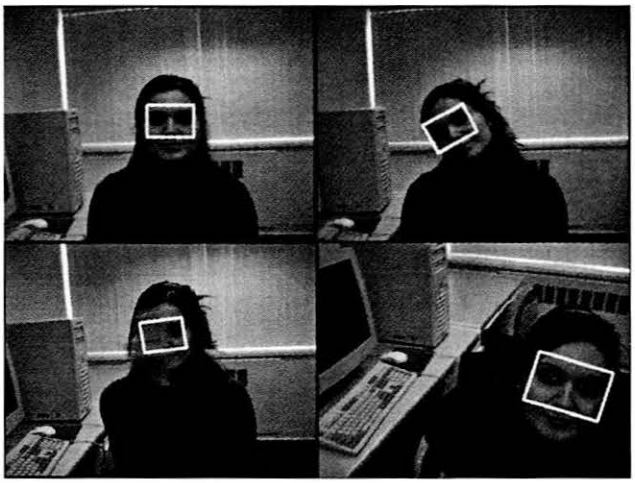
Background and Objectives: The problem of human detection in still images has been extensively investigated within the computer vision community. It is difficult because of the large variations in human appearance due to changes in illumination, camera position, clothing and body pose. We have developed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

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

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

Project Type: Research

Pose Tracking



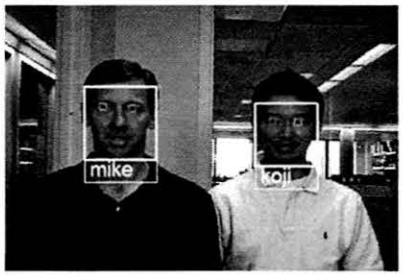
As opposed to the existing approaches that can only predict the location of an object, our goal is to track both pose and deformations in real-time. We aim to accurately estimate the scale changes, in-plane and out-plane rotations, affine and perspective motions, and any parametric deformations in addition to object's translational movements. To our advantage, this method can also be trained to track only specific object types, such as faces, vehicles, etc.

Background and Objectives: This technology is an essential component of many vision tasks in addition mainstream surveillance applications. It improves object identification performance by enabling adaptive selection of the most suitable representation. For example, the accuracy of the face recognition significantly increases as the correct pose of the face is provided as a prior. It is also required in robot cameras that are commonly used for factory automation to monitor production lines, in computer animation and human interfaces to obtain more realistic rendering and feedback, in intelligent vehicle systems to achieve more salient situation awareness, etc. Even more importantly, pose tracking has potential to speed up the human and face detectors.

Technical Discussion: Conventional pose tracking approaches, e.g. active appearance models, either require computationally prohibitive algorithms or make restrictive assumptions on the object appearance and quantize the space of possible motions. Instead, we learn how pose changes reflect in the object features, then use this information in a non-linear regression framework that uses Lie algebra as it is modeled on a manifold. Our pose/deformation tracking method is very fast and runs in real-time since it does not need a search operation or testing of multiple hypotheses. We iteratively estimate the most likely pose by employing simple matrix multiplications.

Contact: Fatih Porikli
http://www.merl.com/projects/posetracking/ **Project Type:** Research

2-D Face Recognition



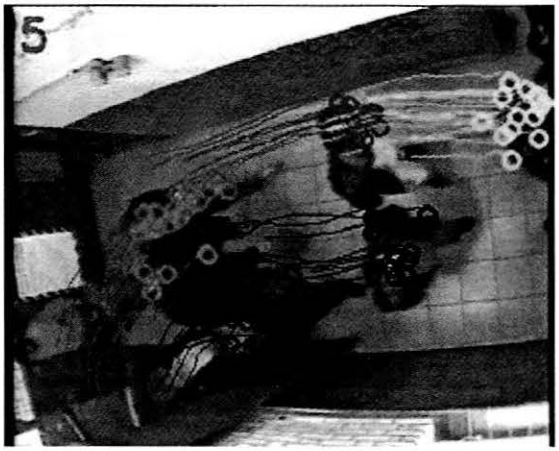
The goal of this project is to have a computer recognize a person from an image of his or her face. There are many applications for face recognition. Some examples are: access control, summarizing surveillance video, browsing image and video databases and user-interfaces. The face recognition problem is broken into two important steps. The first is aligning the face to a standard position, size and rotation. This is done by finding a number of facial feature points (such as the corners of the eyes, the tip of the nose, etc) and mapping those points to standard positions. The next step is to compare two aligned faces to get a similarity score. As the major improvement in the past year, we reduced the memory size by 75% and the error rate by 50%.

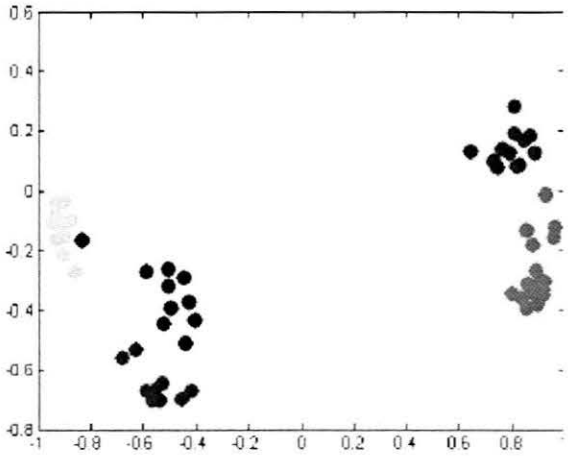
Background and Objectives: Our objective is to develop a state of the art face recognition system. We have been mostly focused on an access control scenario in which the user must cooperate to gain access to a secure room or building. In this scenario lighting can be controlled to a large extent to insure good image quality. We are now beginning to concentrate on less controlled scenarios such as walk-through face recognition or surveillance scenarios in which pose and lighting are much less controlled.

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. The next step of comparing two aligned faces is currently done using features that are simple Haar-like wavelets. These features are selected using the AdaBoost learning algorithm to best separate pairs of face images of different people from pairs of face images of the same person. After the recognition classifier is trained on a large set of example faces, new pairs of faces that have never been seen before can be compared.

Contact: Michael Jones, Jay Thornton
http://www.merl.com/projects/FaceRecognition/ **Project Type:** Advanced Development

People Counting





See Color Figure 7

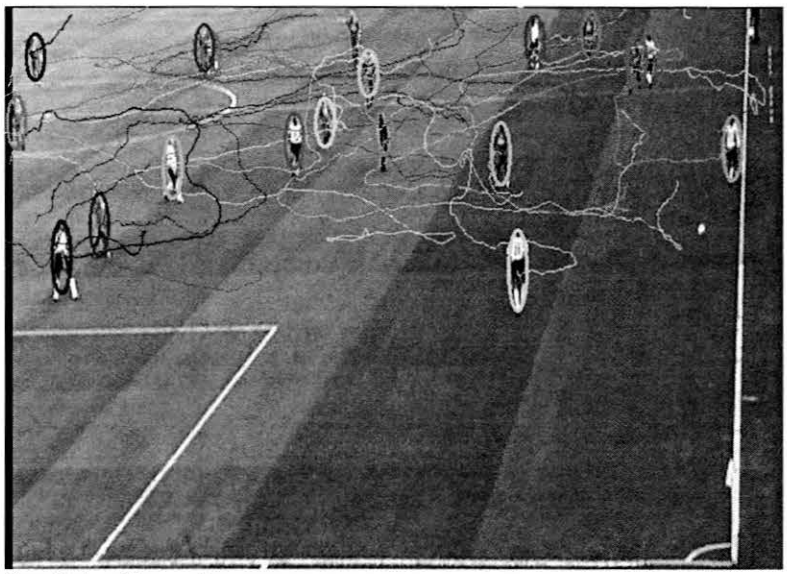
We have developed a new method for people counting in video - a challenging task in computer vision. The method is general and can be applied to other objects, such as vehicles, as well.

Background and Objectives: The problem of counting moving objects has many applications such as: Surveillance applications that need to count the number of people moving in a given field of view, In all cases, the objects in question, move in group which means heavy occlusion. Furthermore, most, if not all, of these systems must operate in real time, if they are to be of any use. Our goal, therefore, is to develop a general ''object counting'' real-time algorithm that can operate in all these scenarios, without much tuning.

Technical Discussion: Our people counting algorithm applies the mean shift algorithm to a non-Euclidean space. We start by tracking feature points in the video and compute a distance metric between these motion feature points. These feature points live in a high dimensional feature space that is not directly amenable to mean-shift clustering. Therefore, we use Multidimensional Scaling to embed the feature points in some Euclidean low-dimensional space in which we perform the mean shift clustering. The algorithm runs at about 10 frames per second.

Contact: Shai Avidan
http://www.merl.com/projects/peoplecounting/ **Project Type:** Research

Integrated View Calibration and Tracking



See Color Figure 5

The goal of this project is to take video from a moving camera and to do view calibration to compute camera motion and create an environment model - in parallel with doing tracking. The benefits are two-fold. Firstly for application use, it is possible to position tracking results within the context of the actual environment model. Secondly in terms of algorithms, the computed camera motion is expected to support and extend tracking capabilities.

Background and Objectives: MERL has advanced tracking technologies but currently has no way to automatically capture an environment model such that the motion of the tracked objects can be shown within the context of the environment. When an environment model is used, it is typically a rudimentary manually-created model, maybe just the ground plane, and is used only with a fixed camera. The objective of this work is to show that view calibration in conjunction with tracking can significantly extend system robustness and can enable new applications.

Technical Discussion: Initial work will treat view calibration and tracking as independent system components. View calibration will involve detection of point and line features from video, matching between frames, computation of camera motion, and computation of 3D features for constructing an environment model. Tracking will be performed on the same images, and a subsequent stage of processing will position the tracking results within the environment model. Initial application will be sports-field analysis, but pedestrian detection and determining pedestrian motion (towards the kerb, parallel to the kerb etc) is also a significant application of interest.

Future Direction: The longer-term goal is to further integrate the view calibration and tracking components, and to demonstrate that they are mutually supporting.

Contact: Paul Beardsley, Fatih Porikli
http://www.merl.com/projects/viewcalib/

Project Type: Initial Investigation

Aerial Terrain Mapping

3D model building for urban and other environments is an increasingly important technology. This project is the first-stage component of a system that builds 3D models from aerial images. The goal is to compute camera position for aerial images followed by 3D features for structures and terrain within the images.

Background and Objectives: 3D urban models support everything from online maps and orientation tools, to live augmentation of video of an urban area, to government planning for urban infrastructure, to disaster recovery. Existing model-creation tools often require manual support which is time-consuming. Our goals are to automate the 3D model construction as far as possible and to support real-time operation.

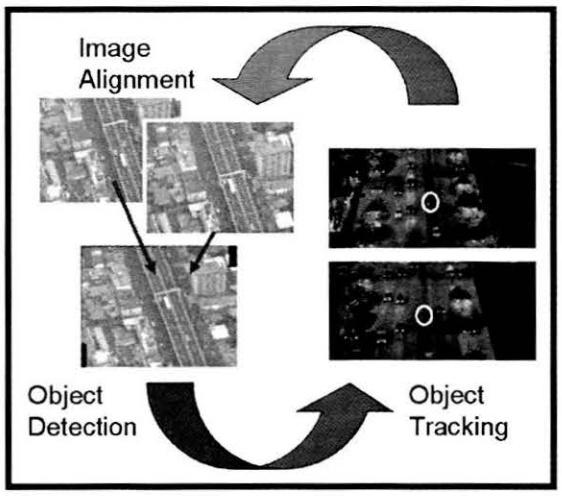
Technical Discussion: This system uses a traditional approach to the problem - computing point features in the aerial images, matching them, computing the camera positions, then reconstructing the 3D point positions. The processing proceeds in an iterative fashion, starting with no initial estimate of the 3D scene, then constraining the matching at each subsequent iteration such that points lie on a terrain-like surface and outlying points are eliminated. The software accepts various starting conditions such as unknown/known camera position, and allows the user to provide hints about the position of the terrain surface if desired.

Future Direction: The initial work was on computing 3D points only. This can be extended by computing 3D lines, which are as common as points in urban environments.

Contact: Paul Beardsley, Jeroen van Baar
http://www.merl.com/projects/aerial3d/

Project Type: Advanced Development

Detection and Tracking for Moving Cameras



The goal of this project is to develop robust algorithms that can automatically detect and track moving objects in moving cameras. Since such algorithms tend to be computationally intensive, we also want to optimize the accuracy while keeping the computational complexity low to process streaming video in real-time. Severe occlusions, poor object representations, and low-quality data are among the major challenges.

Background and Objectives: Detecting moving objects automatically is a key component of many visual content analysis tasks. This technology has many potentially promising applications such as news broadcasting, traffic monitoring, safety and disaster relief maintenance and industrial quality control.

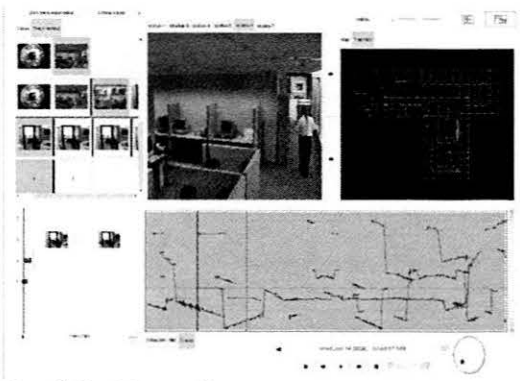
Technical Discussion: Most automatic detection and tracking algorithms are designed for stationary camera setups where priori information and heuristics about the scene and the object motion can be easily incorporated. However, for moving cameras, the object motion is a combined with camera motion. This requires estimation and compensation of camera motion to obtain the true object motion. Along this line, we developed a suite of image registration, feature extraction, background generation and frame-difference based motion detection, evidence accumulation, and inter-frame tracking methods. In low-quality video, the moving objects may be small. We improve the detection performance by integrating backward and forward motion history images, and tracking accuracy by employing an adaptive estimation in multiple likelihood state spaces while selectively applying Markovian prediction schemes such as particle filters.

Future Direction: We are working on general purpose and self-adaptive solutions. We consider extending the possible object classes by incorporating appearance models. We explore various ways of improving the speed by taking advantage of the parallel processors to analyze much higher resolution images.

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

Project Type: Advanced Development

Integrated Event Recognition



See Color Figure 6

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

Background and Objectives: Typically in monitoring spatially extended and geometrically complex areas problems of coverage arise. In order to provide a video coverage of the space many video cameras need to be deployed. In such situations recognition of human activities inevitably begins with attempths to solve smaller irrelevant problems, such as multi-camera tracking, camera handover, object correspondence, multi-camera geometry calibration, etc. This requires massive computational resources and does not address directly the main problem of what is going on in the scene. We solve this problem by using a wide area sensor network to provide the sensor context for the sparse array of video cameras and improve search speeds by several orders of magnitude.

Technical Discussion: We developed algorithms that have the potential to sidestep the video coverage issues by providing a global context of the event. We use the wide area motion sensor network jointly with a sparse array of video cameras to specify and search for events in the entire area without the regard to its observability by a particular camera. We describe these events in terms of space and time, rather than a direct appearance of it in a camera view. This approach allows us to solve problems of large area tracking and event detection efficiently. We use our network of 250+ sensors and 6 PTZ cameras to monitor the entire MERL office with its ~80 occupants for a period of over a year and search it in a matter of seconds. It allows us to judge global office-wide events and activity patterns as well as find and track individuals during the day.

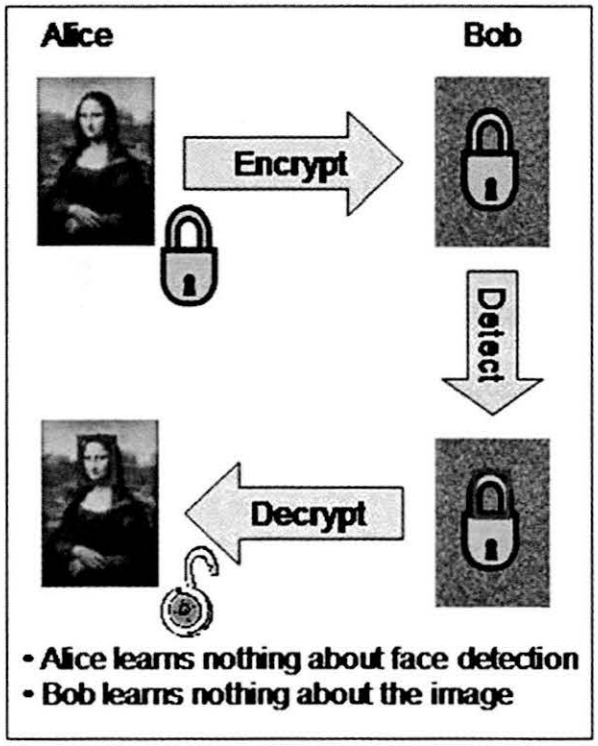
In addition to the effecient search algorithms we have developed large scale visualization techniques and a user interface that allow us to specify events of interest "on-the-fly" and translate simple intuitive gestures into search constraints and query statements. Visualization solutions allow us to display the search results such that complex spatio-temporal relationships are understood pre-attentively.

Future Direction: We plan to integrate the search and forensic tracking system with live event detection module that allows fast autonomous control of cameras, parallel information gathering and best shot detection for recording applications.

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

Project Type: Research

Blind Vision



We have developed a general framework for secure image and video analysis that allows a client to have his data analyzed by a server, privately. For example, the client might submit his images to the server for face detection, without letting the server learn anything about the content of the images. Or, more generally, the client might use a query image to query an image database stored on the server, without revealing the content of the query image to the server. In the last year, we have implemented a secure face detector as a proof-of-concept, presented our work at a scientific conference and extended the method to work with different types of machine learning technologies. We have extended the blind vision framework in two directions: 1) We have explored ways to accelerate the process, using domain specific knowledge; and 2) We have extended blind vision to handle general image matching problems.

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

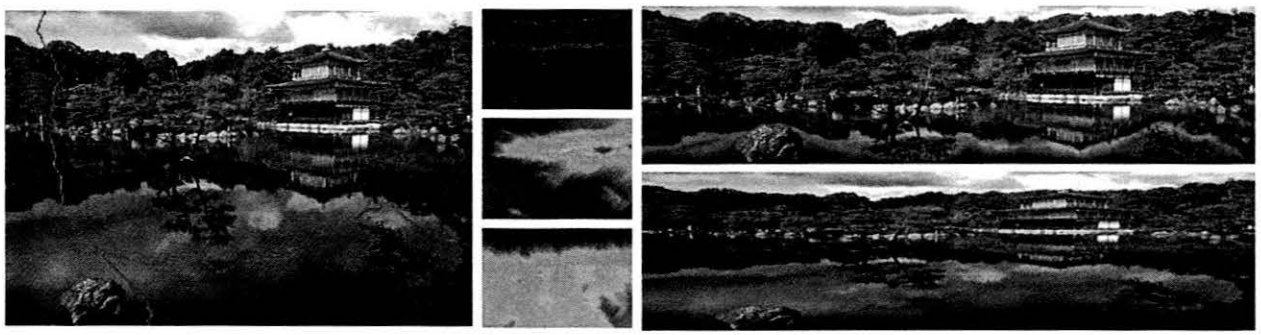
Technical Discussion: Our work combines methods from Computer Vision, Machine Learning and Secure Multi Party Computations. We use cryptographic primitives such as Oblivious Transfer to convert non-secure operations into a secure one. To this end, we have shown (i) how to accelerate a secure face detection algorithms and (ii) how to perform general image matching algorithms in a secure fashion.

Future Direction: We plan to keep exploring other computer vision and machine learning applications that will benefit from a secure framework.

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

Project Type: Research

Media Retargeting



See Color Figure 8

We have developed a new method for image retargeting - changing the size of the image - in a content aware manner. This has wide applications as image today are viewed on a variety of displays with different size.

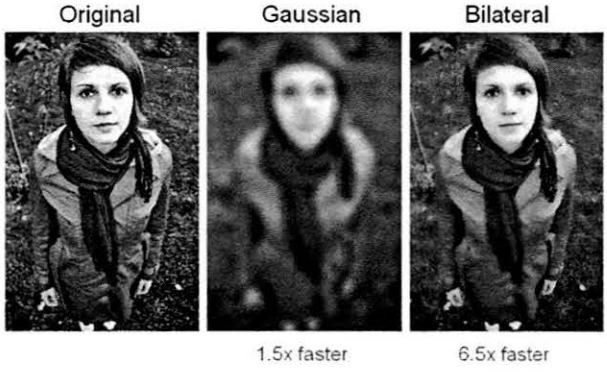
Background and Objectives: We view images on a large number of digital displays such as cell phones, ipod, television, computer screens or high definition television. Yet it each display has its own unique aspect ratio and we need to retarget the image to fit the display. To this end we have developed a novel real-time algorithm that can adapt the aspect ratio of the image by removing low-saliency portions of the image. The technique, termed seam carving, can be used to decrease, or increase the size of an image, perform image manipulations, such as object removal and even serve as an image saliency operator.

Technical Discussion: Effective resizing of images should not only use geometric constraints, but consider the image content as well. We present a simple image operator called seam carving that supports content-aware image resizing for both reduction and expansion. A seam is an optimal path of pixels on a single image from top to bottom, or left to right, where optimality is defined by an image energy function. By repeatedly carving out or inserting seams in one direction we can change the aspect ratio of an image. By applying these operators in both directions we can retarget the image to a new size. The selection and order of seams protect the content of the image, as defined by the energy function. Seam carving can also be used for image content enhancement and object removal. We support various visual saliency measures for defining the energy of an image, and can also include user input to guide the process. By storing the order of seams in an image we create multi-size images, that are able to continuously change in real time to fit a given size.

Future Direction: We plan to extend the method to work on video as well.

Contact: Shai Avidan
http://www.merl.com/projects/mediaretargeting/ **Project Type:** Research

Fast Bilateral Filters



We develop a novel method that accelerates the application of the spatial and edge preserving bilateral filters up to 7 times. In addition to the computational advantage, our method keeps a minimal memory imprint and it is suitable for parallel implementation. Unlike the traditional filters, it can process any arbitrary shaped kernel.

Background and Objectives: Filtering is perhaps the most fundamental operation of image processing and computer vision. Fast realization of spatial and bilateral filters is important for many vision applications from video encoders to consumer cameras to handheld display devices such as cell phones.

In the broadest sense of the term "filtering", the value of the filtered image at a given location is a function of the values of the input image in a small neighborhood of the same location. For example, Gaussian low-pass filtering computes a weighted average of pixel values in the neighborhood, in which the weights decrease with distance from the neighborhood center. However, such an averaging consequently blurs the image. How can we prevent averaging across edges, while still averaging within smooth regions? Bilateral filtering is a simple, non-iterative scheme for edge-preserving smoothing. The basic idea underlying bilateral filtering is to do in the range of an image what traditional filters do in its domain.

Technical Discussion: We show that certain bilateral norms can be expressed as a mixture of the spatial filters without any approximation. We achieve to speed up the other bilateral norms, including Gaussian, using the second and third order approximations without any degradation of the filter response. Our method reshuffles and finds a set of unique filter coefficients, constructs a set of relative links for each coefficient, and then sweeps through the input data by accumulating the responses while applying the unique coefficients using their relative links. It takes advantage of the overlaps between the kernels of the neighboring points to avoid the redundant operations. To further decrease the total number of operations, it quantizes the coefficients while keeping the distortion at minimum.

Future Direction: We are working on parallel implementations of the fast bilateral algorithms to further accelerate their applications.

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

Project Type: Research

Time-Lapse Video Factorization

(a) Original (b) Reconstructed, no shadows (c) Sun illumination only (d) Modified reflectance

See Color Figure 9

We developed a method for converting time-lapse photography captured with outdoor cameras into Factored Time-Lapse Video (FTLV): a video in which time appears to move faster (i.e., lapsing) and where data at each pixel has been factored into shadow, illumination, and reflectance components. The factorization allows a user to easily relight the scene, recover a portion of the scene geometry (normals), and to perform advanced image editing operations. Our method is easy to implement, robust, and provides a compact representation with good reconstruction characteristics.

Background and Objectives: Time-lapse photography, in which frames are captured at a lower rate than that at which they will ultimately be played back, can create an overwhelming amount of data. For example, a single camera that takes an image every 5 seconds will produce 17,280 images per day, or close to a million images per year. Image or video compression reduces the storage requirements, but the resulting data has compression artifacts and is not very useful for further analysis. In addition, it is currently difficult to edit the images in a time-lapse sequence, and advanced image-based rendering operations such as relighting are impossible. We developed a new representation for time-lapse video that efficiently reduces storage requirements while allowing useful scene analysis and advanced image editing.

Technical Discussion: Our method begins by locating the onset of shadows using the time-varying intensity profiles at each pixel. We identify points in shadow and points in direct sunlight to separate skylight and sunlight components, respectively. We then analyze these spatiotemporal volumes using matrix factorization. The results are basis curves describing the changes of intensity over time, together with per-pixel offsets and scales of these basis curves, which capture spatial variation of reflectance and geometry. The resulting representation is compact, reducing a time-lapse sequence to three images, two basis curves, and a compressed representation for shadows. Reconstructions from the data show better error characteristics than standard compression methods such as PCA.

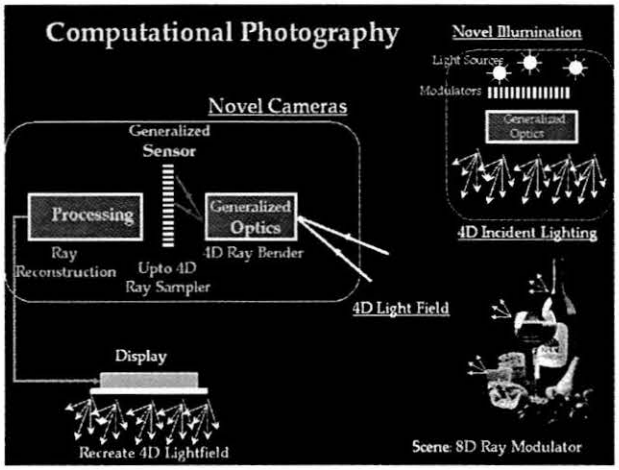
Collaboration: Szymon Rusinkiewicz, Princeton University.

Future Direction: We are currently investigating how the method can be generalized for cloudy and night time scenes.

Contact: Jay Thornton
http://www.merl.com/projects/ftlv/

Project Type: Research

Computational Photography



Computational photography combines plentiful computing, digital sensors, modern optics, actuators, and smart lights to escape the limitations of traditional film cameras and enables novel imaging applications. Unbounded dynamic range, variable focus, resolution, and depth of field, hints about shape, reflectance, and lighting, and new interactive forms of photos that are partly snapshots and partly videos are just some of the new applications found in Computational Photography.

Background and Objectives: In traditional film-like digital photography, camera images represent a view of the scene via a 2D array of pixels. Computational Photography attempts to understand and analyze a ray-based representation of the scene. The camera optics encode the scene by bending the rays, the sensor samples the rays over time, and the final 'picture' is decoded from these encoded samples. The lighting (scene illumination) follows a similar path from the source to the scene via optional spatio-temporal modulators and optics. In addition, the processing may adaptively control the parameters of the optics, sensor and illumination.

Technical Discussion: There are four elements of Computational Photography: (i) Generalized Optics; (ii) Generalized Sensor; (iii) Processing; and (iv) Generalized Illumination. The first three form the Computational Camera. Like other imaging fields, in addition to these geometry defining elements, Computational Photography deals with other dimensions such as time, wavelength and polarization.

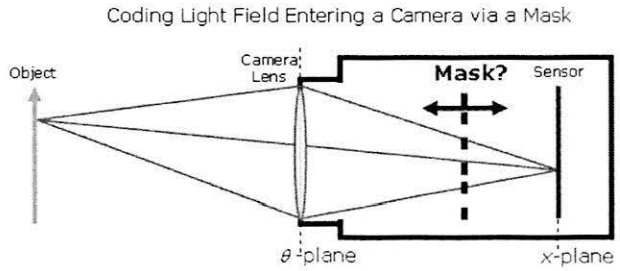
Ongoing MERL projects on computational photography: 1) Coded Exposure (Flutter Shutter Camera), to handle motion deblurring; 2) Coded Aperture for extended depth of field, digital refocusing; 3) Coded Illumination (Multi-flash camera) for detecting depth edges in real time, useful for numerous vision applications; and 4) Heterodyne Light Field Camera using a transmissive mask placed close to the sensor: For extended depth of field and capturing 4D light field.

Collaboration: Prof. Jack Tumblin, Dept of Computer Science, Northwestern University, and Prof. Shree Nayar, Columbia University, NY.

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

Project Type: Research

Coded Aperture for Light Field Acquisition and Digital Refocusing



We propose a new class of "Mask-enhanced" cameras for capturing scene properties beyond conventional 2D cameras. The new camera simply holds a transmissive mask inside a conventional camera. The modification is cheap and easy to install. Applications are show in image restoration due to out of focus blur, digital refocusing and computing the 4D light field of the scene.

Background and Objectives: Capture scene properties beyond conventional pixel intensities and recoverably encode scene information.

Technical Discussion: A new Fourier domain theory is developed that describe the effect of placing a single transmissive mask in the optical path of a conventional camera. Two designs are proposed. In the first design, a broadband mask is placed in the aperture to preserve high spatial frequencies. This allows full resolution digital refocusing and extended depth of field. In the second design, a narrowband mask is placed close to the sensor which allows capturing the 4D light field inside the camera. The narrowband mask acts as a carrier to modulate the 4D light field on to the 2D sensor, similar to heterodyning in radio communications.

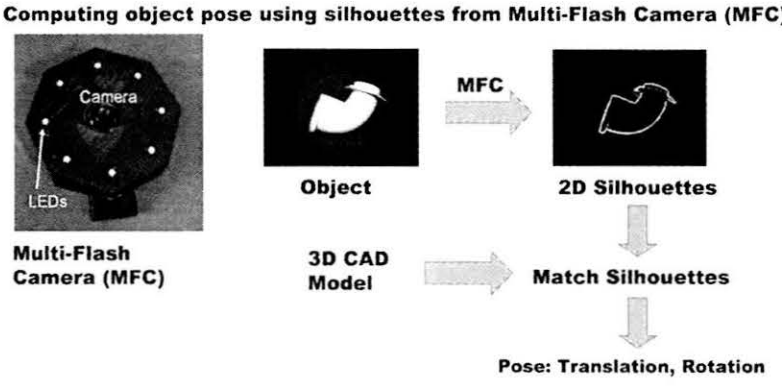
Collaboration: Prof. Jack Tumblin, Dept of Computer Science, Northwestern University.

Future Direction: Investigate mask installation in a hand-held digital camera for commercial purpose and investigate general ray modulators beyond 2D masks for encoding scene properties in other dimensions such as wavelength, time etc.

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

Project Type: Research

Multi-Flash Imaging for Object Pose Estimation



Pose estimation of objects is a common task in many computer vision applications. In this project, we use the silhouettes of the objects along with a CAD model for fast and robust pose estimation. The silhouettes/depth edges are reliably obtained using a multi-flash camera without the need of obtaining 3D information.

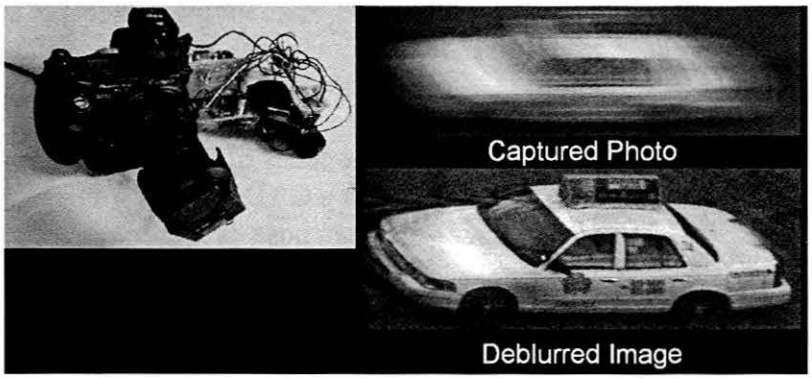
Technical Discussion: A multi-flash camera captures few (4 or 8) images using illumination from different directions. The shadow information is used to obtain the depth edges in real time. A 3D CAD model of the object is used to match the 2D silhouettes obtained by the multi-flash camera to estimate the pose of the object.

Future Direction: 1) Investigate segmenting silhouettes corresponding to the top-level object using shadows; and 2) Investigate pose estimation for specular objects.

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

Project Type: Research

Image Deblurring with Coded Exposure (Flutter Shutter Camera)



A new camera, flutter shutter camera is proposed to aid motion deblurring due to fast moving objects (cars etc) or cameras (aerial imaging). The modification amounts to simply opening and closing the shutter according to a pseudo-random code during the exposure time. The flutter shutter camera is also used for super-resolution of moving objects by utilizing the motion blur information.

Background and Objectives: Objective: Image deblurring is an ill-posed problem and traditional methods lead to noise and ringing artifacts in the deblurred image. The proposed camera can make deblurring a well-posed problem.

Technical Discussion: A traditional camera loses high spatial frequencies if there is a relative motion between the camera and the object during the exposure time. By opening and closing the shutter using carefully chosen binary codes, high spatial frequencies can be maintained. This makes motion deblurring a well-posed problem.

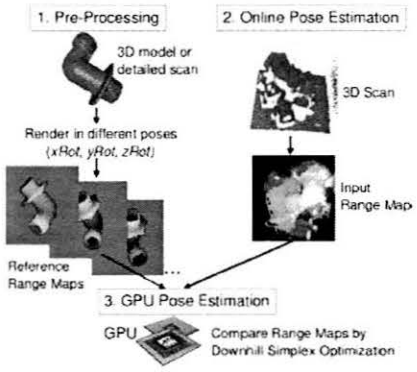
Collaboration: Prof. Jack Tumblin, Dept of Computer Science, Northwestern University.

Future Direction: 1) Investigate image priors for reducing noise in reconstruction; 2) Investigate continuous valued code and better codes by accounting for CCD photon noise; and 3) Utilize Flutter shutter videos for automatic motion estimation and deblurring.

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

Project Type: Research

GPU Pose Estimation



Object pose (location and orientation) estimation is a common task in many computer vision applications. Although many methods exist, most algorithms require a reasonably accurate initial pose and lack robustness to illumination variation, appearance change, and partial occlusions. We propose a GPU-based fast method for automatic pose estimation without manual setting of initial pose based on shape matching of a 3D model to a range image of the scene. Our algorithm is simple and accurately estimates the pose of partially occluded objects in cluttered scenes in about half a second.

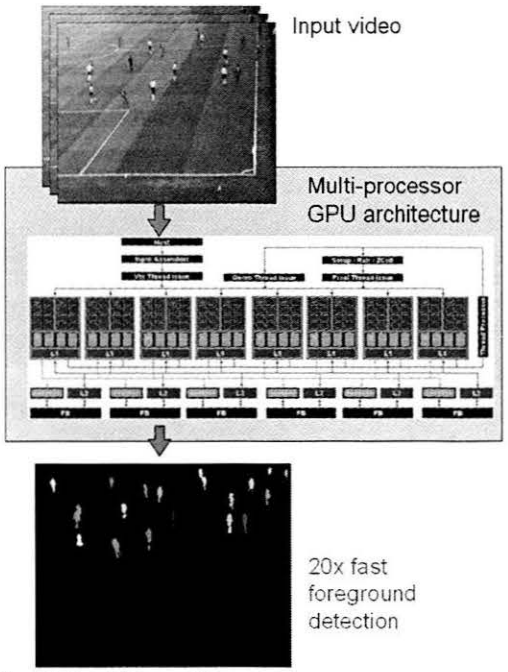
Background and Objectives: Pose estimation in scenes with clutter (due to unwanted objects and noise) and occlusions (due to multiple overlapping objects) is challenging. Pose estimation from range images – where each pixel contains an estimate of the distance to the closest object – is more robust to changes in illumination, shadows, and lack of features. Range images can be robustly acquired with active light systems. If a database of 3D models of objects is available, one can use model-based techniques, where the 3D model of the object is matched to the range image of the scene. We developed a novel model-based pose estimation algorithm for range images that runs entirely on modern Graphics Processing Units (GPUs). The massive data-parallel processing on GPUs (NVIDIA GeForce 8800 GTX) makes our method over 30 times faster than a comparable CPU implementation (on Pentium D945). Our method does not require manual setting of initial pose and accurately computes object poses for synthetic or laser scan data in about one second.

Technical Discussion: The figure shows an overview of our method. In a pre-processing step, we use a 3D model or detailed scan of the object and render it in different poses. Each pose is stored as a reference range map in texture memory. This task has to be performed only once per reference object. During online pose estimation, we acquire a 3D scan of the scene using an active light method (in our case a laser range scan). We smooth the 3D scan on the GPU using a median filter to compute the input range map. The task is now to find the best match between reference range maps and input range map through error minimization by pairwise comparisons. We devised a novel error function that uses the range values and Euclidean distance maps. The error function can be evaluated per pixel, which makes it suitable for efficient processing on GPUs. To efficiently minimize the error we developed a novel data-parallel version of the downhill simplex algorithm that runs entirely on the GPU.
All GPU code is implemented in NVIDIA'a Compute Unified Device Architecture (CUDA).

Collaboration: In Kyu Park, Inha University, Incheon, Korea.

Contact: Joseph Katz
http://www.merl.com/projects/gpu-pose/ **Project Type:** Research

GPU for Surveillance



See Color Figure 10

Our goal is to develop very fast image and video processing algorithms by taking advantage of Graphics processing units (GPU). We have already implemented MERL's state-of-science Bayesian background generation and foreground detection method. In comparison to the CPU version of the same algorithm, the speed of the GPU implementation is more than 20 times faster.

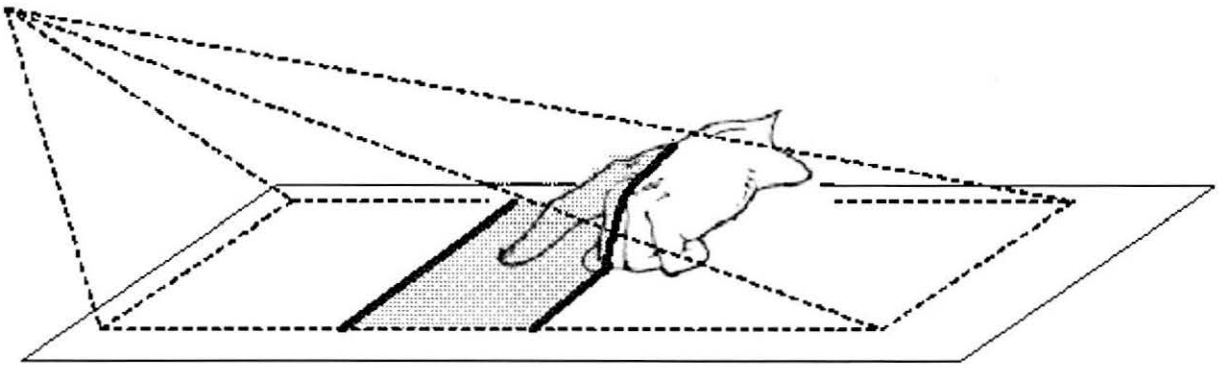
Background and Objectives: GPU's have made their way to home computers through video games and multimedia long time ago. With the increasing programmability, GPU's are capable of performing more than the specific graphics computations for which they were designed. GPU computing with CUDA™ (*) is a new approach to computing where hundreds of on-chip processor cores simultaneously communicate and cooperate to solve complex computing problems, especially for the time consuming video and 3D data analysis tasks.

Technical Discussion: One of the most important innovations offered by CUDA technology is the ability for threads on GPUs to cooperate when solving a problem. By enabling threads to communicate, CUDA technology allows applications to operate more efficiently. GPUs featuring CUDA technology have an on-chip Parallel Data Cache that developers can use to store frequently used information directly on the GPU. Storing information on the GPU allows computing threads to instantly share information rather than wait for data from much slower, off-chip DRAMs. This advance in technology enables users to find the answers to complex computational problems much more quickly than using traditional architectures.

* (CUDA™ is a trademark of NVIDIA.)

Contact: Fatih Porikli
http://www.merl.com/projects/gpusurveillance/ **Project Type:** Advanced Development

Vision for Interaction



We are developing a vision-based technique for interaction with a cell phone augmented with a small projector device.

Background and Objectives: Naturally, cellphones have very small amount of screen space to provide a full interactive experience. On the other hand, a promising direction is to increase the display size by using a small projector. This makes it difficult to interact with such a projected image and new sensing techniques need to be developed. We are conducting an initial proof of concept investigation into using computer vision algorithms to be able to provide an interactive paradigm while using a regular table top surface.

Technical Discussion: For the purposes of vision-based interaction with a projected image, one needs to establish a correspondence between observed image pixels and coordinates of these pixels in the projector space. We recover this correspondence by using an efficient 3D reconstruction technique using the projected image as a natural reference for 3D geometry of the scene.

Contact: Yuri Ivanov
http://www.merl.com/projects/vizinteration/

Project Type: Initial Investigation

Mechatronics

In the last part of this year, MERL embarked on the development of a new research program in mechatronics and control. The motivations for this new program are twofold. First, the design and control of electromechanical devices is central to many areas of MELCO's business, such as factory automation, aerospace, and transportation (automotive, elevator, and escalator). Second, with the rapidly increasing power of embedded computation in electromechanical systems, we felt there was the opportunity for synergy between research in mechatronics and control and MERL's existing research strengths in computer and information technology.

We are currently in the recruiting and project definition phase of this research program, so there are no specific projects to publicly report at the time of this writing. The overall goal of the program is to make fundamental innovations to extend the performance envelope of systems such as high performance factory automation machines, elevators and other state-of-the-art electromechanical control systems. Researchers at MERL will collaborate closely with MELCO's mechatronics and control R&D facilities in Japan.

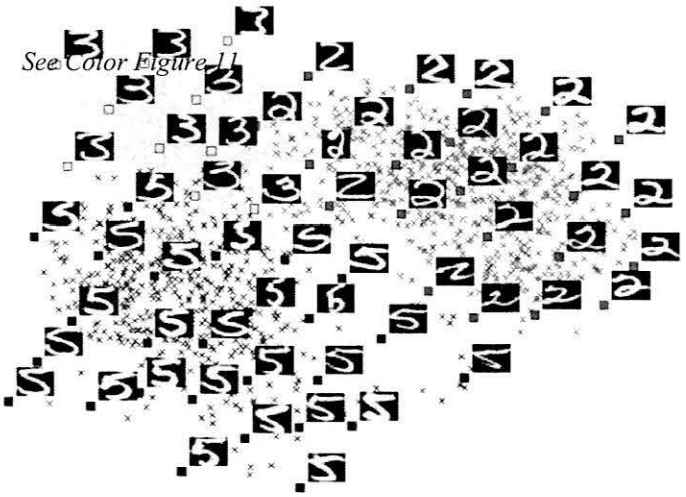
Algorithms

The Algorithms group at MERL develops solution methods for optimization problems involving very large numbers of variables. Typically these arise in inference problems involving images, video, or audio; network transport problems; coding and compression problems; or design problems. Usually these problems are characterized by very complicated probability distributions in extremely high dimensional spaces. Because classical approaches to these problems are intractable, our results can open new business opportunities where there are no competitive technologies.

Most of the group's work revolves around graph-based optimizations and inference, where the graph is a representation of the problem constraints and a probability distribution over possible solutions. Through formal analysis we identify tractable estimation or approximation schemes. This meshes with MERL's expertise in fields and technologies such as belief propagation, machine learning, computer vision, dynamic programming, convex optimization, coding and communications theory, and signal processing.

Project Descriptions

Dimensionality Reduction



See Color Figure 11

We developed methods for compressing high-dimensional signals that enable smooth interpolation and extrapolation between images, sounds, shapes, etc.

Background and Objectives: It may take millions of bytes to accurately record biometric data such as the shape of one's face, but it only takes a few hundred bytes to describe how one's face differs from similar faces. The distribution of all likely faces is presumed to form a smooth low-dimensional manifold. We have developed methods to model this manifold from data samples and assign it a coordinate system with which we can encode (compress) and decode (decompress) new samples. Navigating on this manifold makes it possible to interpolate and extrapolate. We are currently using this facility to generate ear shapes for an acoustic modeling study at MELCO.

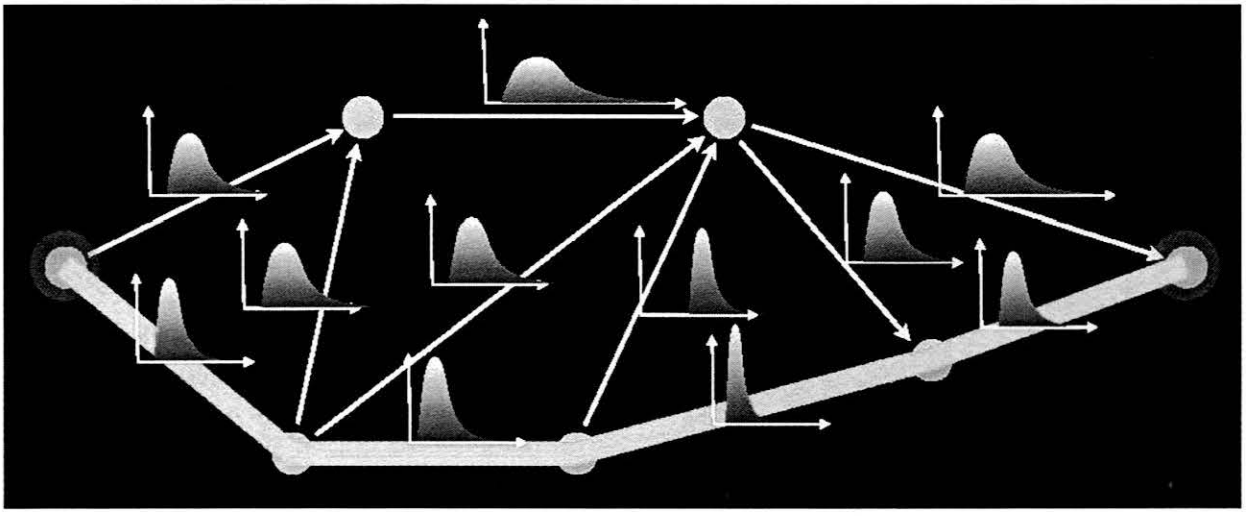
Technical Discussion: Given a few data samples (high dimensional vectors) and local distances between similar samples, we construct a convex optimization whose solution is an isometric mapping function taking the sample space into the low dimensional coordinate system.

Future Direction: We are looking at ways of extending the framework to handle vector samples that lack simple correspondence properties.

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

Project Type: Research

Stochastic Routing



We are developing optimal routing algorithms for networks where the costs vary probabilistically.

Background and Objectives: We study real-world networks such as traffic, communications, and logistics. In these settings, advancing a link is not a sure thing: The cost, time, and success will vary probabilistically. Classical routing algorithms can yield catastrophically bad outcomes. We are developing methods to optimize various measures of performance and bound the probability of failure.

Technical Discussion: We optimize expected utility where the user's utility is some nonlinear function of the realized total travel time or cost. We have proven that some such optimizations are NP-hard and/or inapproximable. For other useful cases we have found very efficient algorithms and approximation schemes. Two examples that are especially suitable for current car-navi hardware: Choosing a route that maximizes the probability of arriving before a deadline, and choosing a policy that minimizes time and avoids jams when the congestion state of the whole network changes stochastically.

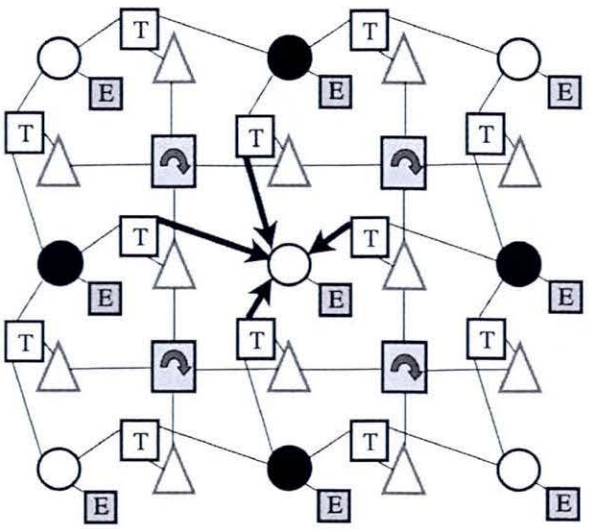
Collaboration: Jan Kara, Charles University (Prague).

Future Direction: We are now looking at wireless transmission problems where the link cost distributions are extremely heavy-tailed.

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

Project Type: Research

Phase Unwrapping for Interferometric Synthetic Aperture Radar



We developed a phase-unwrapping algorithm applicable to interferometric synthetic aperture radar (InSAR) systems. Our approach has the advantage of removing noise while simultaneously unwrapping phases. Our algorithm has been incorporated into MELCO's latest InSAR system, which has been successfully tested in the field.

Background and Objectives: Phase-unwrapping is a necessary step in obtaining a terrain height map from a satellite or airplane-based InSAR system. InSAR measurements provide a phase which is inherently ambiguous with respect to the addition or subtraction of units of 2*pi, and this ambiguity must be removed to recover a sensible height map. To further complicate the problem, InSAR phase measurements are also noisy. The phase-unwrapping problem can be attacked by finding the most probable set of "unwrapped" phases given the received phases, using a terrain model and a noise model. In our approach the terrain and noise models are explicitly defined by a "factor graph," and the most probable configuration is optimized iteratively using a combination of two algorithms, an iterated-conditional-modes (ICM) algorithm, and a minimum-cost-flow (MCF) algorithm. Our algorithm provably finds a locally optimal configuration.

Technical Discussion: Our phase-unwrapping algorithm is being used in the latest MELCO InSAR system, and is a key reason that the new system is able to generate much more accurate digital surface models. An experimental field observation was carried out in Tsukuba, Japan, and the new InSAR system incorporating MERL's phase-unwrapping algorithm achieved an accuracy of better than 50 cm at a spatial resolution of 30 cm. Detailed terrain features such as rice field terraces are accurately recovered.

Future Direction: We plan to further refine our algorithm and improve its efficiency.

Contact: Ali Azarbayejani
http://www.merl.com/projects/phaseunwrapping/ **Project Type:** Advanced Development

Color Figures

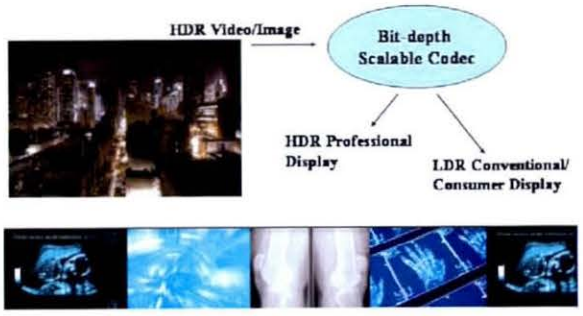


Figure 1 – see page 66

Figure 2 – see page 73

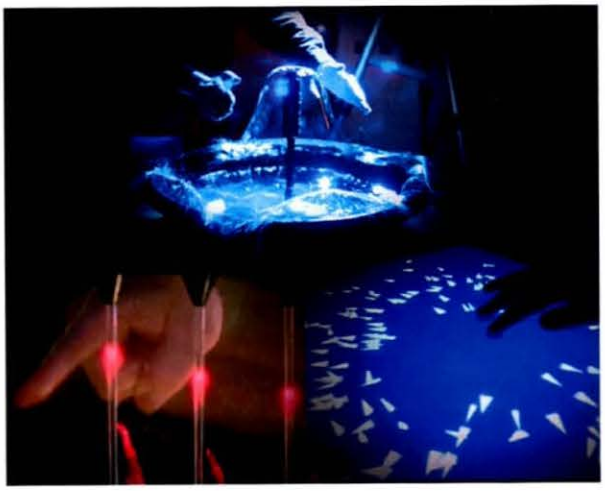
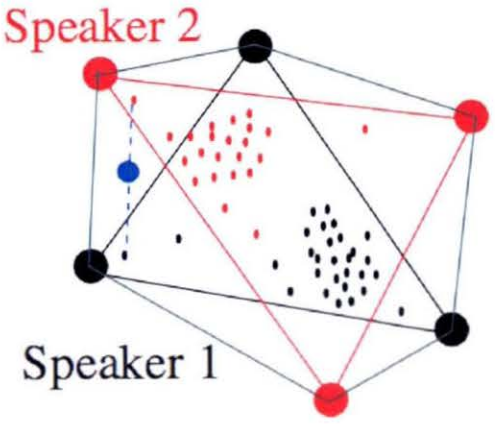


Figure 3 – see page 86



Schematized representations of speaker spectra in a latent variable space

Figure 4 – see page 88

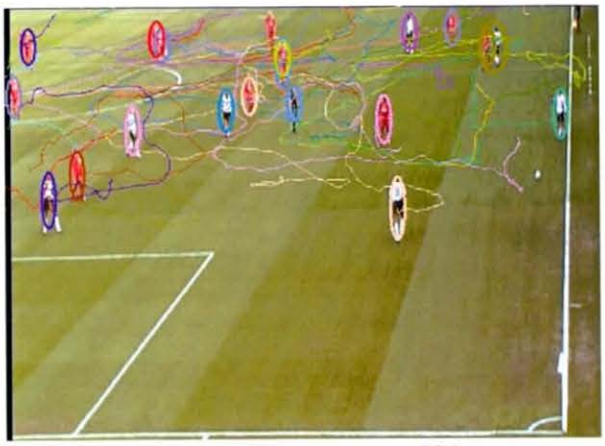


Figure 5 – see page 101

Figure 6 – see page 104

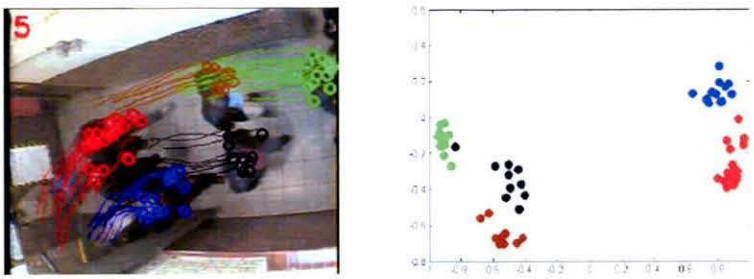


Figure 7 – see page 100

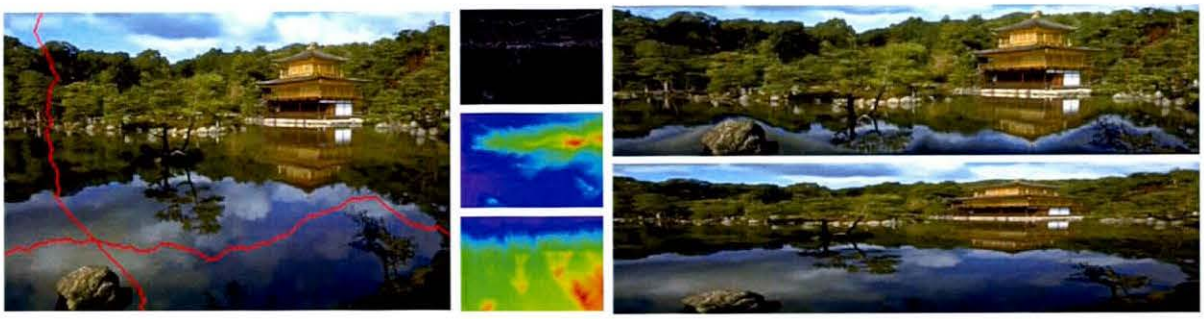


Figure 8 – see page 106

(a) Original (b) Reconstructed, no shadows (c) Sun illumination only (d) Modified reflectance

Figure 9 – see page 108

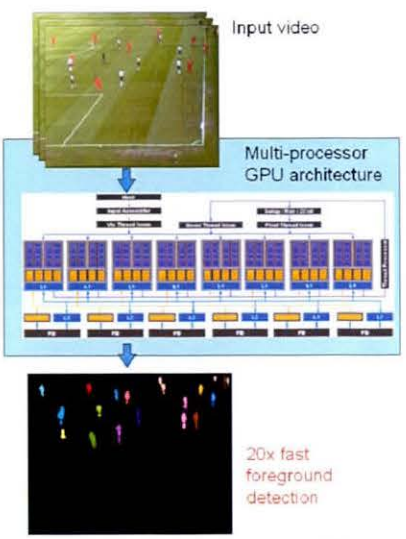


Figure 10 – see page 114

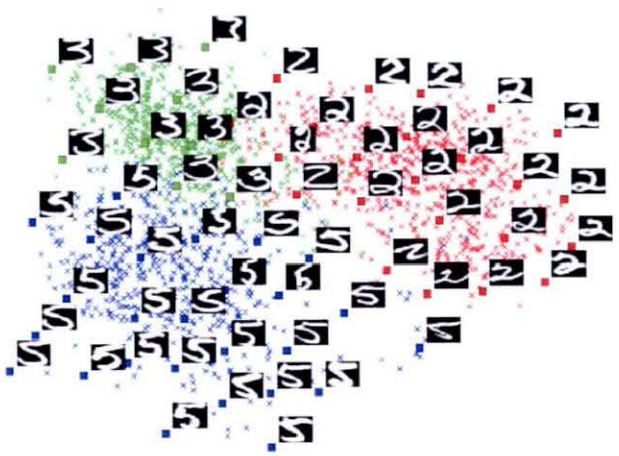


Figure 11 – see page 120