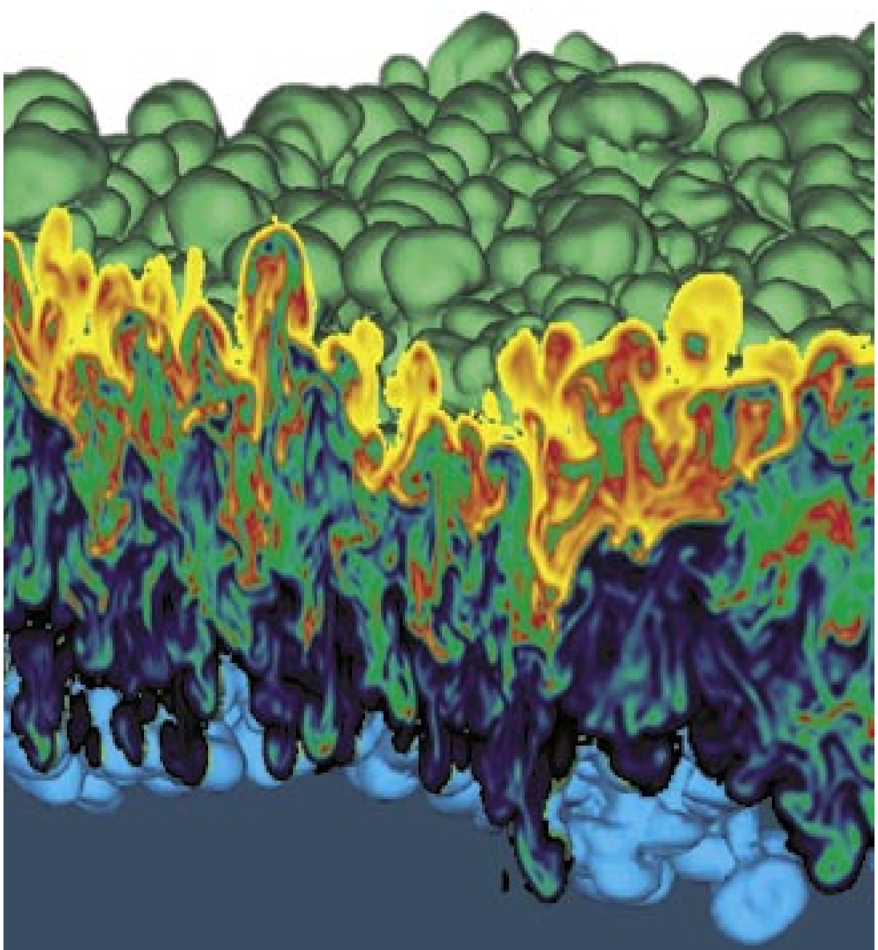
Laboratory Directed Research and Development (LDRD)

The ISCR supported three Exploratory Research in the Institutes (ERI) projects with Laboratory Directed Research and Development (LDRD) funds during FY 2004. These research grants are awarded to LLNL staff with the goal of developing ties to academia through co-funded research projects. Anticipating the emergence of data science as a cross-disciplinary theme, the ISCR has concentrated its efforts in developing technologies for large-scale and distributed data sets for the past several years. The current portfolio contains a project on providing a single interface to diverse data sources on the Web, a project on tracking objects in a succession of images, and a project in multiresolution representation of scientific datasets for scalability across a range of computer architectures. This portfolio, originally motivated by purely scientific applications, has already paid dividends in some of the Laboratory's new homeland security applications.



Enabling Large-Scale Data Access

Principal Investigators

Terence Critchlow and David Buttler, CASC

This project's goal was to develop an infrastructure capable of providing scientists with access to large numbers of data sources. To that end, we have developed an infrastructure that employs a user-specified description of a service of interest to crawl the Web. When an interface in the service class is identified, a wrapper that supports an XML-based query interface is automatically created. This year, we successfully demonstrated the application of our infrastructure on service classes from two vastly different domains — Basic Local Alignment Search Tool (BLAST) sequence similarity interfaces and interfaces to publication sites.

We have demonstrated the ability to automatically identify new data sources of interest while crawling the Web. We performed extensive testing of the infrastructure on hundreds of BLAST data sources and half a dozen publication sources. Next, we performed a minimal Web crawl, during which we identified and wrapped nine previously unknown BLAST sites and one new publication site. In addition, we identified approximately 100 Web pages that contained citations but were not queryable.

This work supports national security and other LLNL missions by benefiting ongoing programs at LLNL, such as Department of Homeland Security nonproliferation and detection efforts, that must utilize information from a wide variety of sources, including some that cannot be easily integrated using traditional techniques. Our infrastructure simplifies the process of creating an interface that combines local data with related information publicly available over the Internet, such as scientific publications.

We met and exceeded our FY 2004 milestones. In addition to demonstrating end-to-end automatic wrapper generation for BLAST interfaces, we extended the service class description to handle more complex data types, such as citations that cannot be easily represented by regular expressions. We performed several short Web crawls that successfully demonstrated the application of our infrastructure on both well-structured BLAST interfaces and poorly structured publication interfaces by identifying previously unknown sites in both domains. The expertise gained in this project will be applied to the Biodefense Knowledge Center Bio-Encyclopedia effort.

Publications

(1) D. Rocco, T. Critchlow. "Automatic Discovery and Classification of Bioinformatics Web Sources." *In Proceedings of the Georgia Tech Conference on Bioinformatics.* UCRL-JC-152980. 2003

(2) J. Caverlee, L. Liu, D. Buttler. "Probe, Cluster, and Discover:focused extraction of QA-Pagelets from the deep web." *In proceedings of the IEEE conference on Data Engineering.* 2004

(3) Wei Han. Ph.D. dissertation Georgia Institute of Technology 2003

(4) Dan Rocco. Ph.D. Dissertation, Georgia Institute of Technology 2004

Detection and Tracking in Video

Principal Investigator
Chandrika Kamath, CASC

Video cameras are used for monitoring and surveillance in several applications. We are developing robust, accurate, and near-real-time techniques for detecting and tracking moving objects in video from a stationary camera. This allows us to model the interactions among the objects, thereby enabling us to identify normal patterns and detect unusual events. Our algorithms and software include techniques to separate the moving foreground from the background, extract features representing the foreground objects, track these objects from frame to frame, and post-process the tracks for display. We focus on video taken under less-than-ideal conditions, with objects of different sizes moving at different speeds, occlusions, changing illumination, low resolution, and low frame rates.

The capability to detect and track in video supports the national security mission of LLNL by enabling new monitoring and surveillance applications for counterterrorism and counter-proliferation. This project will produce robust and accurate technology for video detection and tracking under less-than-ideal conditions with occlusions, fog or changing illumination, or at a low resolution or frame rate. This project will enhance existing algorithms to address these situations, allowing us to better understand their limitations, which in turn, will determine the conditions under which successful surveillance is possible. The algorithms and software are being applied to surveillance video, as well as spatiotemporal data from computer simulations.

During FY 2004, we

(1) Created a software infrastructure to handle streaming video data.
(2) Implemented several background subtraction algorithms and evaluated them on videos taken under different conditions.
(3) Proposed a new background subtraction method that outperforms other methods, especially on low-resolution, low-frame-rate video.
(4) Extracted features and used them in simple tracking algorithms.

We also filed a provisional patent on the new method and summarized our work in two papers. We are currently adapting the tracking algorithms to work under adverse conditions. We collaborated with the University of Colorado, Boulder, on tracking people, the University of California, San Diego, on tracking under occlusions, and a summer student on object representations for tracking.

Publications

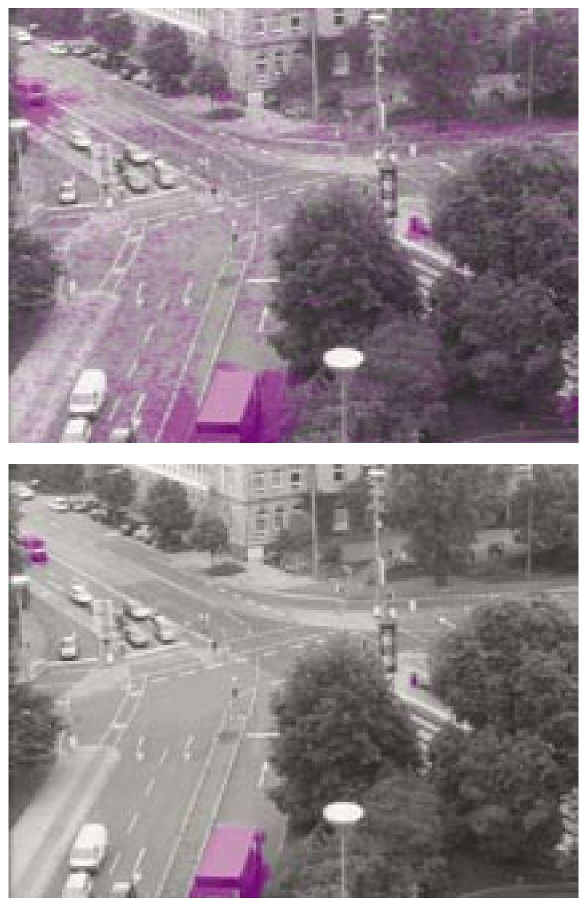
Cheung, S.C., C. Kamath, "Robust background subtraction with foreground validation for urban traffic video," *Eurasip Journal on Applied Signal Processing,* UCRL-JRNL-201916.

Moelich, M., "Autonomous motion segmentation of multiple objects in low resolution video using variational level sets," UCRL-TR-201054.

Gyaourova, A., C. Kamath, S.C. Cheung, "Block matching for object tracking," UCRL-TR-200271.

Cheung, S.C., C. Kamath, "Robust techniques for background subtraction in urban traffic video," *Video Communications and Image Processing Conference,* **Vol. 5308**, pp. 881-892, UCRL-CONF-200706.

Cheung, S.C., "Robust techniques for background subtraction," UCRL-ABS-200371.



Our new technique for background subtraction (bottom) is less sensitive to changes in illumination in comparison with current techniques (top). The pixels highlighted in purple indicate the moving objects in the frame.

ViSUS: Visualization Streams for Ultimate Scalability

Principal Investigator

Valerio Pascucci, CASC

We are developing a suite of progressive visualization algorithms and a streaming infrastructure to enable interactive exploration of large scientific data sets. The methodology optimizes the data flow in a pipeline of processing modules. Each module reads and writes a multi-resolution representation of a geometric model, providing the flexibility to trade speed for accuracy, as needed. The data flow is streamlined with progressive algorithms that map local geometric updates of the input into immediate updates of the output. A prototype streaming infrastructure will demonstrate the flexibility and scalability of this approach for visualizing large data sets on a single desktop computer, a cluster of personal computers, and heterogeneous computing resources.

In FY 2004, we brought ViSUS to a level of maturity and robustness allowing direct deployment for a number of targeted users. The main milestones achieved include developing new techniques that accelerate isosurface extraction with occlusion culling, graphics hardware, and view-dependent refinements. A test viewer has been developed for datasets from the HYDRA simulation. We provided a stable library with full implementation of our streaming technology that can be used by simulation codes for saving rectilinear grids in ViSUS IDX format. We released a new version of the Progressive Viewer with full slicing, isocontouring, and volume-rendering capabilities. We are working in collaboration with the MIRANDA team to start using the IDX format as output of choice for the Blue Gene/L runs.

The ViSUS project benefits the Laboratory at least at two levels. At the deployment level, the improved efficiency in the use of hardware resources reduces the cost of visualization-hardware infrastructures. At the scientific level, the developed technology reduces the overall time required for the design, simulation, and visualization cycle. The ability to remotely monitor large and expensive simulations saves computing resources through early termination and restart of erroneous test simulations. Runtime steering will be possible for simulation codes with mechanisms for dynamic modification of running conditions.

Use of our innovative, high-performance visualization techniques allows interactive display of very large data sets on simple desktop workstations and the monitoring (or steering) of large parallel simulations. This will have valuable applications to several LLNL missions, including stockpile stewardship, nonproliferation, energy security, and environmental management, that use large-scale modeling and simulations.

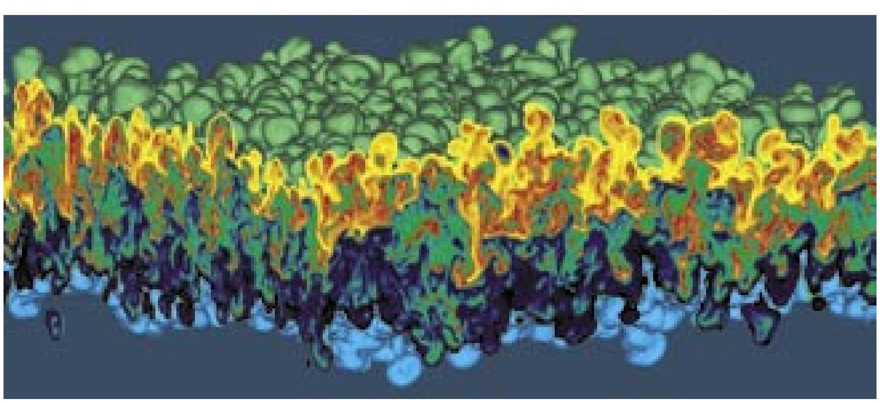
Publications

Laney, D. and V. Pascucci, "Progressive Compression of Volumetric Subdivision Meshes," *Proceedings of the International Symposium on 3D Data Processing, Visualization, and Transmission 2004*, pp. 293–300, UCRL-CONF-203679.

Pascucci, V. "Topology Diagrams in Scientific Visualization" Chapter in: *Surface Topological Data Structures: An Introduction for Geographical Information Science*, pp. 121–130, UCRL-200013-BOOK.

Pascucci, V., "Isosurface computation made simple: Hardware acceleration, adaptive refinement and tetrahedral stripping," *Proceedings of the Joint Eurographics - IEEE TVCG Symposium on Visualization*, pp. 293–300, UCRL-CONF-202459.

Van Kreveld, M.J., R.van Oostrum, C.L. Bajaj, V. Pascucci, and D.R. Schikore, "Efficient contour tree and minimum seed set construction" Chapter in: *Surface Topological Data Structures: An Introduction for Geographical Information Science*, pp. 71–86, UCRL-200018-BOOK.



The ViSUS Progressive Viewer generates a snapshot of the MIRANDA hydrodynamics code computing a Rayleigh–Taylor turbulent mixing of fluids.