Scalable Visualization Toolkits for Brains to Bays

PROJECT LEADERS

ARTHUR OLSON
The Scripps Research Institute (TSRI)

BERNARD PAILTHORPE
SDSC

ARTHUR TOGA
UCLA

CARL WUNSCH
MIT

PARTICIPANTS

JON GENETTI, DAVID NADEAU
SDSC

MICHEL SANNER
TSRI

CHANDRAJIT BAJAJ
University of Texas, Austin

NICOLE BORDES, ALEX DE CASTRO, RICHARD CHARLES, JOHN MORELAND
SDSC

ANNA OMELCHENKO
TSRI

ARIEL SHAMIR, SANGHUN PARK, XIAOYU ZHANG
University of Texas, Austin

BERND HAMANN, IKUKO TAKANASHI
UC Davis

PAUL THOMPSON
UCLA

SCOTT BADEN
UC San Diego

DETLEF STAMMER
Scripps Institution of Oceanography

JOEL SALTZ
University of Maryland/ Johns Hopkins University

JOERG MEYER
Mississippi State University

Accurate rendering of detailed biomedical, geophysical, or astronomical data sets up to and beyond a terabyte in size can strain even a supercomputer's capability. Yet scientists also need tools that will handle smaller tasks, and using a specialized code package that runs *only* on a supercomputer to visualize megabyte data sets is like cracking a walnut with a sledgehammer. NPACI's Scalable Visualization Toolkits alpha project is creating versatile software to render and interact with data sets on systems ranging from desktop computers to supercomputers. "Researchers need a scalable interface to take advantage of NPACI's evolving computing capabilities," said Arthur Olson, leader of NPACI's Interaction Environments thrust area. "A multiplatform toolkit for visualizing terabyte-scale volumetric, multi-modal, time-varying data sets will benefit the entire scientific community. The project has reached a stage at which we can demonstrate some of these benefits."

The Scalable Visualization Toolkits alpha project is integrating renderers, user interfaces, and data-orchestration utilities from various Interaction Environments projects into a coherent whole. The system architecture is layered: upper layers provide an object-oriented interface to arbitrarily large multi-dimensional data sets and perform visualization operations upon the data, while lower layers provide data I/O and storage. The toolkits include components for analyzing, filtering, compositing, and rendering data.

"We are developing these tools in coordination with NPACI scientific application groups," said Bernard Pailthorpe, associate director for Scientific Visualization at SDSC and alpha project manager. "In particular, Neuroscience and Earth Systems Science thrust researchers have found conventional visualization tools to be inadequate."

DEFINITION AND RENDERING

Several of the project's tools concern defining and rendering volumes simultaneously with surfaces. One toolkit, the Volume Imaging Scalable Toolkit Architecture (VISTA), produces composite visualizations from multiple, multi-modal data sets, including volumes, geometry, and scene descriptions. Leveraging technology from MPIRE, an earlier package for 3-D volume rendering, VISTA is an application program interface that can be linked into existing code packages. It incorporates true perspective rendering, luminosity and opacity as voxel parameters, and extended scene definition capabilities.

Chandrajit Bajaj leads the project's Vis Tools parallel data analysis and rendering development efforts at the University of Texas' Center for Computational Visualization (CCV). The VisualEyes high-performance volumetric visualization tools integrate feature detection analysis with real-time surface and volume rendering. The VisualEyes architecture allows parallel, out-of-core, data analysis, and feature detection servers to feed multiple client-side rendering processes. "Scientists and engineers are increasingly dependent on visualization to bring forth the hidden information in their equations and calculations," Bajaj said. "They are also realizing that visualization, interactive exploration, and visual querying of their results are indispensable in their quest for new discoveries or validation of their theories."

www.npaci.edu/Alpha/ScalableVis

CCV contour spectrum and fast isocontour tools analyze physical parameters and automatically compute and render time-varying isocontours; surfaces interpolating time-series data points on a 3-D mesh with the same parameter values can be interactively visualized with varying colors and transparencies. Other topology analysis methods identify points, curves, and surface features that mark transitions in 3-D data sets. VisualEyes recently has been enhanced to composite volume renderings with isocontours of multi-modal, time-series volumetric data.

A volume scene graph toolkit gives high-level support for multiple, overlapping data sets. "Volume scenes are well-suited to applications that need to correlate several 3-D data sets," said SDSC's David Nadeau, who implemented the basic Volume Tools infrastructure to organize data flow and computation. "Scenes might compare and contrast overlapping volumes from multiple simulation runs, or an application might combine overlapping volumes from several scans of a specimen using different technologies—MRI, PET, and cryosectioning, for example."

ACCESS AND CONTROL

Arthur Olson and Michel Sanner at The Scripps Research Institute (TSRI) direct client-side integration of the visualization toolkits. Using Python, an interpreted language comparable to Perl but with fundamental object-oriented language features, they have created two high-level tools that run on various computers and operating systems. Déjà Vu is a workstation environment for scripting, visualization, interactive exploration, and manipulation of volumetric data; its graphical interface enables users to interactively change visualization parameters and to identify and explore regions of interest. Working in conjunction with Déjà Vu, the Python Volume Viewer is a toolkit that takes advantage of hardware-oriented direct volume-rendering technologies.

Development of a paged-array toolkit that gives access to arrays of data too large to fit in computer memory is now complete. Layered onto this toolkit is a mesh toolkit that supports manipulation of arbitrary, N-dimensional, paged data sets.

Data I/O for these toolkits will be compatible with the Active Data Repository (ADR) and DataCutter from the Programming Tools and Environments thrust and the SDSC Storage Resource Broker (SRB) from the Data-Intensive Computing thrust. Access to large data sets on computational grids via the SRB has been developed at SDSC under the direction of Data-Intensive Computing thrust leader Reagan Moore. ADR optimizes storage, retrieval, and processing of very large multi-dimensional data sets and supports common database operations, data retrieval, memory management, process scheduling, and user interaction. Under the direction of Programming Tools and Environments thrust leader Joel Saltz, the ADR server has been ported from an IBM SP at the University of Maryland to Blue Horizon and has been extended to use tape archives such as the High Performance Storage System (HPSS).

BRAINS TO BAYS AND BEYOND

Arthur Toga directs UCLA's Laboratory of Neuro Imaging (LONI), where his research team develops tools and data sets to measure the structure and function of healthy and diseased brains. Most of the neuroimaging and brain mapping is conducted in 3-D; representing feature variability over time or across populations increases the number of dimensions to four or even five. "We are developing a modeling environment to allow researchers to more easily study and understand the dynamic evolution of brains as they change through childhood growth, normal aging, and deterioration from pathological conditions such as Alzheimer's disease," said UCLA assistant professor Paul Thompson, head of the LONI modeling team. LONI researchers are providing hundreds of gigabytes of multi-modal human brain data as test cases for the Scalable Visualization Toolkits effort (Figure 1).

Carl Wunsch at MIT and Detlef Stammer at UCSD's Scripps Institution of Oceanography are modeling the general circulation of Earth's oceans, including the oceanic heat and fresh water fluxes and other factors that are key to understanding climate changes. The MIT/SIO Ocean State Estimation project analyzes World Ocean Circulation Experiment hydrographic and current data and TOPEX/POSEIDON satellite observations in worldwide and regional models, dividing the surface of Earth into thousands of simulation cells. The models map the flow field and track temperature, salinity, oxygen content, nutrients, carbon and nitrogen markers, and sea surface level. The computational fluid dynamics simulations can produce terabyte-scale output files, which also serve as test cases for the toolkits.

AT SC2000 IN DALLAS

At SC2000, the alpha project team will demonstrate volume exploration with the Python Volume Viewer and a hardware-accelerated direct volume rendering system. The demo will feature interactive navigation of a downsampled LONI brain data set, with a "volume zoom" function—automatic download to the client of subvolumes of high-resolution data—and user manipulation of color and opacity to determine optimal parameters for high-resolution rendering on the server. Researchers also will demonstrate interactive rapid isocontouring on the client computer, used in downloading geometry from the large server database.

An animation created with CCV's integrated surface and volume renderers will depict the results of time-series simulations from the Earth Systems Science thrust—vertical and oblique views of scientifically important regions in the North Atlantic and South Atlantic, based on preliminary 30-layer high-resolution oceanographic simulations by Stammer and Arne Biastoch of SIO.

Other demonstrations will show "volume-cleaning"—cropping and elimination of extraneous voxels—by software developed by Bernd Hamann and UC Davis colleagues and Joerg Meyer of Mississippi State University, and interactive use of a scene graph viewer with data subsetting developed at SDSC. UCSD's Scott Baden will demonstrate the KeLP parallel programming libraries enhanced and adapted for use with the toolkits.

"We're making visible progress," Pailthorpe said. "This collaborative effort between computer scientists and application scientists in multiple domains is addressing challenges at the frontiers of scientific visualization." *—MG* ▼

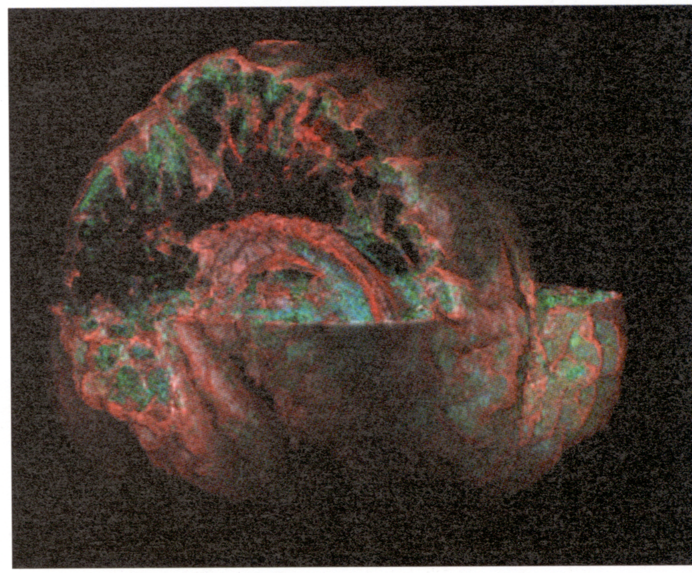

FIGURE 1. SLICING INTO A 3-D BRAIN

These 3-D renderings of a human brain were produced by the participants in the Scalable Visualization Toolkits alpha project. The data originated as 753 photographs of cross-sections of an actual human brain from the Laboratory of Neuro Imaging at UCLA. The data were "cleaned" of extraneous elements by interactive volume rendering software developed by Bernd Hamann of UC Davis and Joerg Meyer of Mississippi State University. The resulting data set, more than 6 GB in size, was then reconstructed into a 3-D scene definition by volume scene tools developed by the SDSC Scientific Visualization group. The top visualization was rendered by the VISTA software toolkit developed by John Genetti of SDSC. The other image is a screen capture of client-side real-time interactive exploration of the 3-D data using the Python Volume Viewer, developed by Arthur Olson and Michel Sanner of TSRI, which can identify structures and relationships of interest hidden within the volume for later high-resolution imaging.