BNL-68461

High-Performance Computing for the Study of Earth and Environmental Science Materials Using Synchrotron X-ray Computed Microtomography

HUAN FENG
Department of Earth and Environmental Studies, Montclair State University, Upper Montclair, NJ 07043, USA
fengh@mail.montclair.edu

KEITH W. JONES
Environmental Sciences Department, Brookhaven National Laboratory, Upton, NY 11973, USA
kwj@bnl.gov

MICHAEL MCGUIGAN, GORDON J. SMITH, JOHN SPILETIC
Information Technology Division, Brookhaven National Laboratory, Upton, NY 11973, USA
mcguigan@bnl.gov, smith3@bnl.gov, spiletic@bnl.gov

Synchrotron x-ray computed microtomography (CMT) is a non-destructive method for examination of rock, soil, and other types of samples studied in the earth and environmental sciences. The high x-my intensities of the synchrotron source make possible the acquisition of tomographic volumes at a high rate that requires the application of high-performance computing techniques for data reconstruction ta produce the three-dimensional volumes, far their visualization, and for data analysis. These problems are exacerbated by the need to share information between collaborators at widely separated locations over both local and tide-area networks. A summary of the CMT technique and examples of applications are given here together with a discussion of the applications of high-performance computing methods to improve the experimental techniques and analysis of the data.

1. Introduction

Materials studied in the earth and environmental sciences are generally very inhomogeneous and complex materials. Investigation of the three-dimensional properties of these materials is essential. However, there are relatively few ways that these properties can be measured using non-destructive methods. These methods include laser confocal microscopy, magnetic resonance imaging, and x-ray computed tomography. The use of x-ray computed tomography is particularly powerful since it gives information on x-ray atteuuation coefficients and thus can distinguish between different minerals, pore space, and liquid-filled pore space in specimens that can range from a few millimeters to many centimeters in size. Our purpose here is to describe the application of computed tomography techniques to these problems based on the use of synchrotron radiation and to discuss the application of high performance computing to improve the technique.

2. Instrumentation and Method

A schematic diagram of the CMT apparatus at the Brookhaven National Synchrotron Light Source (NSLS) is shown in Figure 1 [7]. The x rays are detected with a thin scintillator of CsI(Na) or YAG:Ce. A mirror/lens system is used to focus light from the scintillator onto a charge-coupled device (CCD) camera. Blurring effects caused by scattering of the x rays in the scintillator are minimized by the small depth-of-field of the magnifying lens. The spatial resolution of the system is energy dependent and of the order of 0.005 mm. The CCD cameras used employed CCD chips with 13 17 x PO35 and 3072 x 2048 pixels. In practice, the pixels are often binned to reduce the size of the tomographic volumes and thereby ensure practicable times for data acquisition and reconstruction. Monoenergetic beams are used at energies to about 20 KeV, and filtered white beam is used for measurements at higher energies to obtain higher x-ray intensities. Typical data collection times are of the order of Z-2 hours.

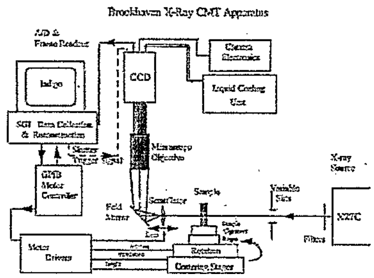

Figure 1. A schematic diagram of the apparatus.

Data acquisition produces a set of camera frames taken as the sample is rotated in a series of steps determined by the number of pixels desired in the volume. The procedure covers an angular range of 180 degrees with respect to the incident x-ray beam. The number and size of files depend on the sample size and desired spatial resolution.

The data reconstruction proceeds in 3 phases. Phase 1 applies a whitefield normalization, any filters needed, and writes files containing data from all views for a single slice (horizontal row of pixels), one file per slice. Phase 2 processes each slice independently. It applies the view-by-view air value normalization, optionally applies a. filter to reduce the ring artifacts, computes the location in the images of the center of rotation, and converts the data to a sinogram. Phase 3 is the actual reconstruction. It generates a square array with dimensions of the horizontal row size. After this, the reconstruction is completed and the visualization process begins. This is a much more varied process and depends strongly on the particular sample being analyzed.

3. Experimental Results and Discussion

We have used the CMT apparatus to investigate many different materials relevant to the earth sciences. In particular, saudstones are of wide interest, and typical data is presented here to show a specific example of the usefulness of synchrotron CMT.

The data can be presented in several ways. A volume representation showing a 3-D view of sandstone from the Vosges region of France is shown in Figure 2. The grain structure of the material is clearly visible. A view of a single section through the sample is shown in Figure 3. This type of view helps to highlight the pore-grain relationships.

The data can then be processed according to the measured attenuation meffficients to display the data in binary form representing either solid or pore space. Analysis of this data them gives the two-dimensional correlation function, porosity? permeability, and tortuosity. The measured microgeometry also can be used as a realistic basis for fluid flow calculations [2].

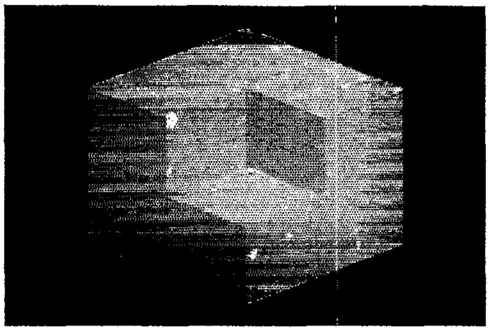

Figure 2. Isosurface rendering of a three-dimensional tomographic volmne of a sample of sandstone from the Vosges, 'The color scale indicates the values of the measured absorption coefficients. The pore space is shown as blue.

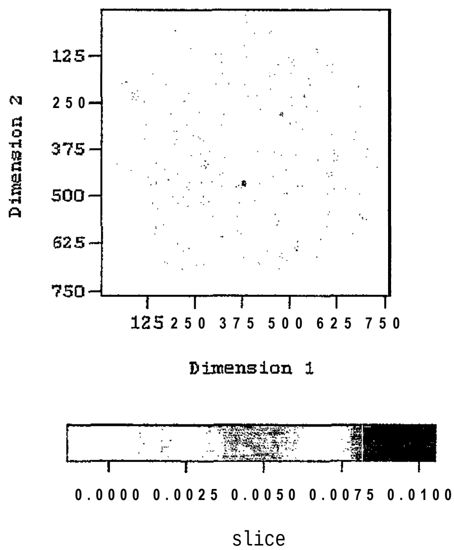

Figure 3. Single section through the sandstone volume shown in Figure 2. Absorption coefficients are indicated by the color scale.

4. High Performance Computing

The present system used at BNL and others, similar in design at the NSLS, APS, ESRF, and other synchrotrons, have demonstrated their worth in varied experiments on environmental and earth science topics. However, the system performance and usefulness can be vastly improved by use of high-performance computing technology.

First, present day CCD cameras can produce 15-20 frames per second with a size of about 1000 x 1000 pixels, Assuming that about 1500 frames will be necessary to acquire data for a tomogram, approximately 6 gigabytes of data are produced per second (2 bytes per pixel), The goal is then to carry out all steps of the reconstruction process in the time needed to acquire the data, or about 1500 frames/1 5 frames/s or 100 s.

Second, in order to be able to adjust experimental parameters in near real time, it is necessary to produce data displays on about the same time scale. There needs to be control of the display by the experimenter so that different views are feasible and thresholding can be set to display changes in pore structures or fluid motion.

There are other aspects of the problem that include high-speed networking, rapid data storage, and remote viewing that present challenges to computer science. Here we concentrate our attention on the first two points, data reconstruction and viewing. We have developed methods for doing these tasks on the time scales required by recourse to parallel computation techniques.

X-ray computed microtomography is a highly computer intensive and memory intensive application. The large volumes and small grid spacing required for micrometer resolution push the limits of even the most powerful workstations. For this reason we are applying high performance parallel computation and data compression for remote access of tomographic data sets. The two main areas that need to be addressed are reconstruction and visualization. Reconstruction takes many projections obtained from the high resolution CCD camera and uses a FFT transform method developed by Robert Marr at Brookhaven to form a three-dimensional gridded data set. Visualization of the data set either as a volume or an isosurface is used to closely inspect the sampb and extract new features. This is typically done within a complete visualization package such as OpenDX or VTK.

Reconstruction of large data sets can take several hours. Visualization on large grids can also take tens of minutes just to form a particular view of an isosurface. The use of parallelization to address these problems is one way to achieve necessary speedups for fast reconstruction and visualization. Basically one divides up the data set across multiple processors and reassembles the final result by synchronization and communication among processors. There are two main protocols for parallelization: Parallel Virtual Machine (PVM) and the Message Passing Interface (MPI). MPI is the more recent of the two protocols and is becoming a standard particularly on clusters of Linux computers, Parallel reconstruction using MPI is currently being used at the APS, and this technique can also be applied to data from the NSLS beam line. The results of applying parallel visualization using MPI to the NSLS beam line data are described here.

4.1 Parallel Visualization

Parallel visualization refers to the use of multiple computers for the graphical depiction of large data sets. It has its origins in parallel rendering where one breaks up the image to be rendered into smaller pieces and has each processor render its own part with all the pieces brought together for the final composite image [6, 8]. For example, Baily [5] has ported Pixar's RenderMan rendering software to the parallel Intel paragon machine. The free ray tracing program POV has also been ported to Linux clusters using MPI and PVM and achieved near linear speedups comparable with more expensive specialty supercomputers. Recently, parallel visualization has gone a step beyond this by breaking up the data set itself This data parallel model allows large data sets of high resolution, such as the x-ray tomography data sets discussed here, to be manipulated by using the cumulative memory and processing power of a Linux cluster, The data parallel model has been implemented in OpenDX 4.1.2 and VTK using MPI [1, 3].

4.1.1 OpenDX

OpenDX is an open source visualization package that can be applied to a wide variety of data sets [1]. Currently, we use the software to give a quick view of x-ray tomographic data sets, extract isosurfaces and slices, and convert from NetCDF data format to a data format that can be read into VTK. Recently, a port of the software using MPI has been achieved, and we are planning to apply this version to x-ray tomography. Figure 5 shows a screen save of the OpenDX so&ware we have developed. OpenDX builds applications by drag& modules of specific functionality into the visual programming editor canvas and linking them together to form a network. The network shown uses an isosurface module, an export module to do the NetCDF conversion, and an image module for the final rendering. The data set shown is a high resolution x-ray tomographic data of the thigh bone of a rat used in osteoporosis studies.

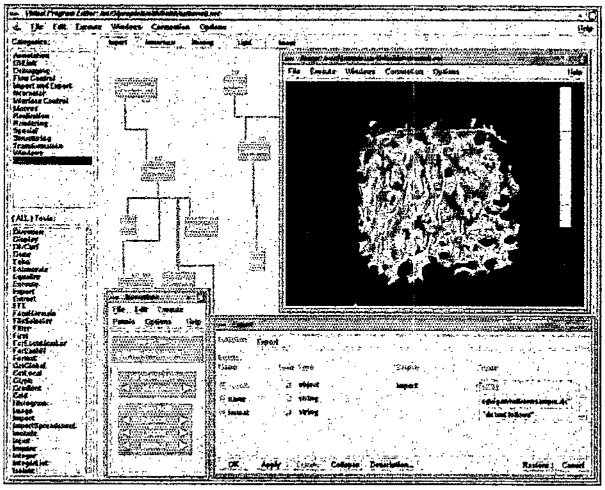

Figure 5. OpenDX application for isosurface extraction and conversion from NetCDF data format to VTK format.

4.1.2 VTK

The 'Visualization Toolkit (VTK) is a multiplatform visualization package with C++, TK/TCL and Python bindings. It is known for its high quality implementation of the latest algorithms in computer graphics. Its most recent version includes a parallel MPI implementation of a subset of its modules [3, 4]. Here we discuss three programs built on this subset which we have applied to x-ray tomographic data s&s.

The first program, ParaIso, breaks up the data set according to the command line arguments. It then computes isosurfaces for each piece on separate processors and then renders the final image with colors to indicate the work of the separate processors. An example for a tomographic data set is shown in Figure 6. The computation was done on a 4 Xeon processor 1400 Linux server from SGI.

The second MPI program DataParallelism breaks up the data more directly by using an image reader module. It includes it's own sample data and explicit timing functions to measure the performance of the parallel isosurface computation. The results for the sample data are shown in Figure 7 and indicate nearly a factor of three speedup using all four processors.

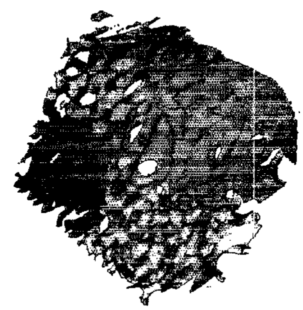

Figure 4. Output from the VTK program ParaIso. The four colors indicate the portion of the isosurface computed by each of the four processors on the Linux cluster.

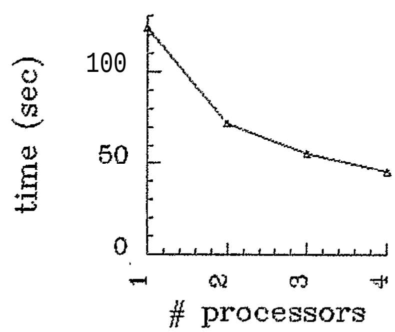

Figure 7. Graph showing the speedup using the parallel MPI VTK software DataParallelism to compute isosurfaces.

The third program, ParaView, is potentially the most powerful of the three. It includes a Tk widget control panel for interacting with the parallel MPI VTK application. This includes functionality to input various data sources, control over isosurface values, and thresholding for volumetric data. We are currently modifying the code to directly input x-ray tomographic data. A screen shot of the user interface applied to sample data is shown in Figure 8.

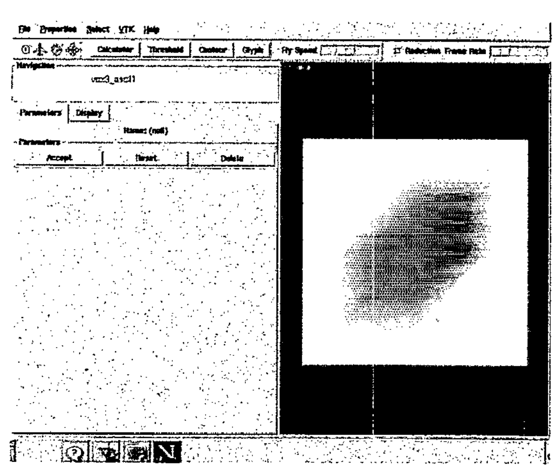

Figure 8. User interface for the parallel MPI VTK program, ParaView, shown with sample data set.

4.2 Remote Visualization

The use of large computer facilities to perform x-ray tomographic reconstruction and visualization is of no use unless the final rendering can be delivered to the scientist's desktop. This can mean delivery over a high-speed intranet or more remote delivery connecting dispersed researchers over the Internet or wide area network. We are currently testing two software packages designed to perform remote visualization. Both packages use compression algorithms to speed up transmission to the desktop.

The first remote visualization software, VizServer from SGI, transmits OpenGL visualization remotely from a multipipe SGI Onyx2 computer to SGI, Sun, or Linux clients. The software was able to achieve frame rates useful for interactive viewing (approx. 10 frames/sec) for typical data sets. The Onyx2 we used had 2 GB of Ram., 6 processors and 2 pipes. As only one pipe is available for vizserving, this means at most one user at a time can access the data remotely from this system

A second remote visualization software has recently been released by TGS as part of OpenInventor 3.0. OpenInventor is a high level C++ and Java graphics toolkit built on top of OpenGL. This software works on a variety of platforms and doesn't require separate pipes for each user. It delivers OpenInventor applications directly to the desktop, and these can include x-ray tomographic data in Inventor and VRML format.

Another approach to remote visualization is the construction of a VRML server to deliver 3-D models directly over the Internet. VTK itself can be used to construct such a server in conjunction with a VRML browser. Currently we are using Cosmoplayer and parallelgraphics VRML browsers. The later can deliver on a variety of platforms including wireless PDA devices.

4.3 Stereoscopic Viewing

To understand the three-dimensional structure from high-resolution x-ray computed tomography it is useful to have a stereoscopic presentation of the data, Currently we are using a passive system with two Barco projectors and polarized filters [9]. The projectors are connected to an Onyx2 computer and the visualization is rear projected on a special screen that preserves polarization. The final visualization is suitable for group viewing in a visualization theatre with inexpensive polarized glasses. At the desktop, stereoscopic viewing can be done using a page-flipping method and active glasses. Both methods can be used with OpenDX or VTK visualization toolkits. The parallel graphics VRML browser can also be operated for remote stereoscopic visualization on the Internet.

5. Conclusions

In this paper we briefly described the use of synchroton CMT for investigation of earth and environmental sciences samples and then described improvements to the CMT system by application of high performance computing methods. We used OpenDX and MPI versions of VTK to perform the visualizations and achieved a factor of three increase in speed using a four processor Linux cluster. We also studied the application of visualization server software for desktop delivery and achieved reasonable frame rates for typical data sets. In the future, we will apply MPI versions of OpenDX and the use of VRML servers in order to make these high performance techniques widely available.

Acknowledgments

We wish to thank C. Law, B. Geveci, and S. Murtagh for assistance with parallel VTK and C. Holmes for assistance with SGI vizserver as well as Betsy Dowd for belp with CMT data sets. Work supported under DOE Contract DE-AC02-98CH10886.

References

1. Abram, G. and Treinish, L. An extended data-flow architecture for data analysis and visualization. Proceedings of Visualization, 263, IEEE Computer Society Press (1995).
2. Adler, P. M., Thovert, J.-F., Jones, K. W., Peskin, A, M., Siddons, D. P., Andrews, B., and Dowd, B. Determination of topology and transport in porous media using synchrotron computed microtomography (abstract). Presented at 1996 Spring Meeting of The American Geophysical Union, Baltimore, Maryland (1996).
3, Ahrens, J., Law, C., Shroeder, W., Martin, K., and Papka, M. A parallel approach for efficiently visualizing extremely large, time-varying datasets. LANL Technical Report LAUR-00-1620.
4. Ahrens, J., Martin, K., Geveci, B., Law, C., and Papka, M. Large-scale visualization using parallel data streaming. (preprint)
5. Bailey, M. Parallel RenderMan. http://www.sdsc.edu , (1996).
6. Crockett, T. W. Parallel rendering. In SIGGRAPH '98 "Parallel Graphics and Visualization Technology" Course #42. (1998)
7. Dowd, B. A., Andrews, A. B., Marr, R. B., Siddons, D. P., Jones, K. W., and Peskin, A. M. Advances in x-ray computed microtomography at the NSLS. Presented at 47th Annual Denver X-Ray Conference, Colorado Springs, Colorado, August 3-7, 1998. In Advances in X-Ray Analysis, Vol. 42, Plenum Publishing Corp., New York, New York. (1999)
8. Ellsworth D., Polygon Rendering for Interactive Visualization on Multicomputers, PhD thesis, University of North Carolina at Chapel Hill. (1997)
9. Smith, G., Andrews, B. ITD's BNL Visualization Lab, http://www.itd.bnl.gov/visualization/vis3.html, ITD, BNL, Upton, NY (1997).