Multiresolution Data Access Within The *VisIt* Visualization Environment

Andrew Foulks and R. Daniel Bergeron
Department of Computer Science University of New Hampshire
Durham, NH 038040
email: {rafoulks,rdb}@cs.unh.edu

***Abstract*— Interactive visualization of very large data sets remains a challenging problem to the visualization community. One promising solution involves representing the data at multiple resolution levels in both space and time. Low resolution data is used to give the scientist a large scale overview of the data. Finer resolution data is used to show smaller regions of interest chosen by the scientist exploring the data.**

We have developed multiresolution data support software whose goal is to facilitate interactive visualizations of very large scientific datasets. We have also implemented a database plugin for the *VisIt* visualization environment that allows all *VisIt* visualization modules to access our multiresolution data support system. Using our software, the scientist can view the data at a coarse resolution, can then select a region of interest, followed by a zoom into that region, viewing the data at a finer resolution.

I. Introduction

One significant challenge remaining in the scientific visualization community is the size of the datasets that must be analyzed. Typically, today's simulations generate far more data than the scientist can effectively understand. Since understanding such large and complex datasets is a critical component of understanding the science behind the data, this problem represents a nettlesome bottleneck in scientific research. In this paper, we address this situation by providing the scientist with tools that facilitate interactive visual exploration of very large time series datasets. [1]

A. Application

Interactive visualization of data is a very effective technique for gaining insight into the phenomena behind simulation data. It is the huge size of these science datasets that limit the *interactivity* in the visualization environment. Often space science disciplines must deal with data that is much too large to be stored in main memory or to be represented on a display screen of only a few million pixels. One promising solution involves data representation at multiple resolution levels. Coarser resolutions are used for large-scale overview visualizations, and increasingly finer resolutions can be accessed as the scientist "zooms in" to regions that encompass a more limited spatial and/or temporal range.

Some of the very largest datasets come from space science research that requires the analysis of *time varying* scientific datasets. Fortunately, the temporal domain also provides a natural mechanism for limiting data access that can ease the data size issue to a certain extent. Our research focuses on interactive visualization of *time series* of spatial datasets such as the data produced by magnetohydrodynamics (MHD) simulations. The output produced is quite often many hundreds of gigabytes of three-dimensional time series data. The combination of both temporal and spatial data provides opportunities for the development of visualization applications, to the extent that the application can easily access a data representation based on a resolution in both the spatial and temporal domains that is appropriate for the current task.

[1]This work was supported in part by NASA under grant AISR05-0071, and the National Science Foundation under grants IIS-0082577 and IIS-9871859.

B. Project Overview

We are currently developing *multiresolution data management tools* that help support interactive visualization of very large multiresolution and multisource scientific datasets. Our goal is to integrate our software into existing visualization applications tailored to the needs of specific research areas and tasks. In particular, we focus on supporting visualization software using multi and adaptive resolution data to aid in interactivity. This paper describes the integration of our multiresolution data management support into the *VisIt* visualization environment [1] developed by Lawrence Livermore Laboratories.

II. Multresolution Scientific Data Representation

We have developed a multiresolution scientific data model [2] that incorporates spatial and temporal semantics with localized error and we have implemented a prototype database system (the Granite Scientific Database System) based on that model [3], [4], [5], [6], [7], [8]. The model's semantics are common to many scientific applications and therefore are valuable to a variety of disciplines. The model supports integrated views of multi-file datasets that may be distributed over a network. Both point-based and cell-based data organizations are supported.

We have developed a tool that can generate multiple resolutions for time series data in both the *spatial* and *temporal* domains. The spatial and temporal resolutions are integrated into a single comprehensive data representation. Our prototype implementation is in our *STARgen* and *STARview* tools [9]. *STARview* supports multiple spatial wavelet decompositions

of each step of a time series dataset, which can itself be a member of a multiresolution time series decomposition.

III. MHD Simulation Data Visualization

Space science researchers have built and simulated numerical models of the solar wind and its interaction with the earth's magnetosphere for only the past 20 years or so [10], [11]. The data we are currently working with comes from simulations run at the Space Science Research center at the University of New Hampshire. Their simulation of solar wind activity records several physical attributes, including bulk plasma velocity, current density, magnetic field, and pressure. The data is a 3D time series in which points are sampled on a stretched cartesian grid [12]. The particular dataset used for the examples in this paper contains 87 recorded time steps spanning a numerical simulation of 5220 time steps. The total size of this dataset is 15GB.

IV. *VisIt* Integration

The primary goal of our visualization research is to aid the scientist in gaining knowledge from very large datasets.

VisIt is a general purpose visualization environment aimed at giving researchers scientific visualization tools for scientific datasets [1]. *VisIt* is built largely upon the Visualization Toolkit (VTK) [13] libraries, extending the interface and providing a comprehensive environment to the scientist. *VisIt* supports many different types of data and has a modular architecture that allows users to build data plugins to access other types of data. We have implemented a database plugin for *VisIt* that can read our multiresolution hierarchy and provide *VisIt* renderers with multiple resolutions of data.

We have integrated our multiresolution data software into *VisIt* so that any rendering plugin can use the multiresolution data. Existing plot plugins in *VisIt* should not have to be recompiled to be made "multiresoluion aware". To achieve this task, we have separated our multiresolution data access module from the visualization subsystem.

Our software consists of two *VisIt* plugins. The first is a *Database* plugin, which has the responsibility of reading and importing data in our multiresolution format for use with *VisIt* plots. The second plugin is an *Operator* plugin, which presents the user with a widget, implemented as an *Operator Attribute*, to control the level of refinement from which the current plot plugin gets its data. By checking the box "Auto Update" in the *VisIt* user interface, changes to the current resolution will automatically re-invoke the *VisIt* rendering pipeline, so that the current plot is regenerated with the chosen level of refinement.

Technically speaking, our multiresolution control widget does not "operate" on the data in the sense defined by the *VisIt* framework (i.e., as a run-time filter). Instead, this widget acts only as a front end controller to the database plugin, directing it to fetch data at user specified resolutions. A more proper implementation would not use an Operator for such a task. Instead, a generic controller widget associated only with the Database plugin should be implemented. We are told that future versions of *VisIt* will support such a feature, and it is our intent to use that functionality when it becomes available. Figure 1 shows *VisIt* visualizations for two different resolutions of one time step from Geospace General Circulation Model (GGCM) density simulation data [14], [15].

V. Conclusions and Future Work

We have developed a data model and tools that help the scientist to achieve interactive visualization of very large scientific data sets. To demonstrate our model, we have integrated our preliminary implementation with the existing visualization toolkit, *VisIt*. Renderers in *VisIt* are not aware of the multiresolution or distributed nature of the data. We believe our environment can significantly improve the scientists' research toolkit. Current multiresolution data access is controlled by an operator plugin. In the future we hope to extend our data model integration by giving the scientist access to error data directly in the visualizations. Error, or uncertainty information is important because it can aid the scientist by showing regions in the data in which the data changes rapidly. These regions of interest can be viewed at a finer resolution for closer examination. We also would like to extend the data model to include both spatial and temporal *adaptive resolution* datasets.

References

[1] "Visit visualization environment." [Online]. Available: http://www.llnl.gov/visit/home.html

[2] P. J. Rhodes, R. D. Bergeron, and T. M. Sparr, "A data model for distributed multiresolution multisource scientific data," in *Hierarchical and Geometric Methods in Scientific Visualization*. Heidelberg, Germany: Springer-Verlag, 2001.

[3] ——, "Database support for multisource multiresolution scientific data," in *SOFSEM 2002: Theory and Practice of Informatics, 29th Conference*, ser. Lecture Notes in Computer Science, W. I. Grosky and F. Plasil, Eds., vol. 2540. Milovy, Czech Republic: Springer, 2002, pp. 99–114.

[4] P. J. Rhodes, X. Tang, R. D. Bergeron, and T. M. Sparr, "Iteration aware prefetching of large multidimensional scientific datasets," in *Proceedings of the 17th International Scientific and Statistical Database Management Conference*, June 2005.

[5] ——, "Out of core visualization using iterator-aware multidimensional prefetching," in *Proceedings of SPIE: Visualization and Data Analysis*, vol. 5669, San Jose, California, January 2005, pp. 295–306.

[6] X. Tang, "Integrating block oriented data iteration with multidimensional caching for large scientific datasets," Master's thesis, Department of Computer Science, University of New Hampshire, Durham, New Hampshire, September 2004.

[7] L. Ye, "Representation of adaptive resolution rectilinear cell data," Master's thesis, Department of Computer Science, University of New Hampshire, Durham, New Hampshire, September 2004.

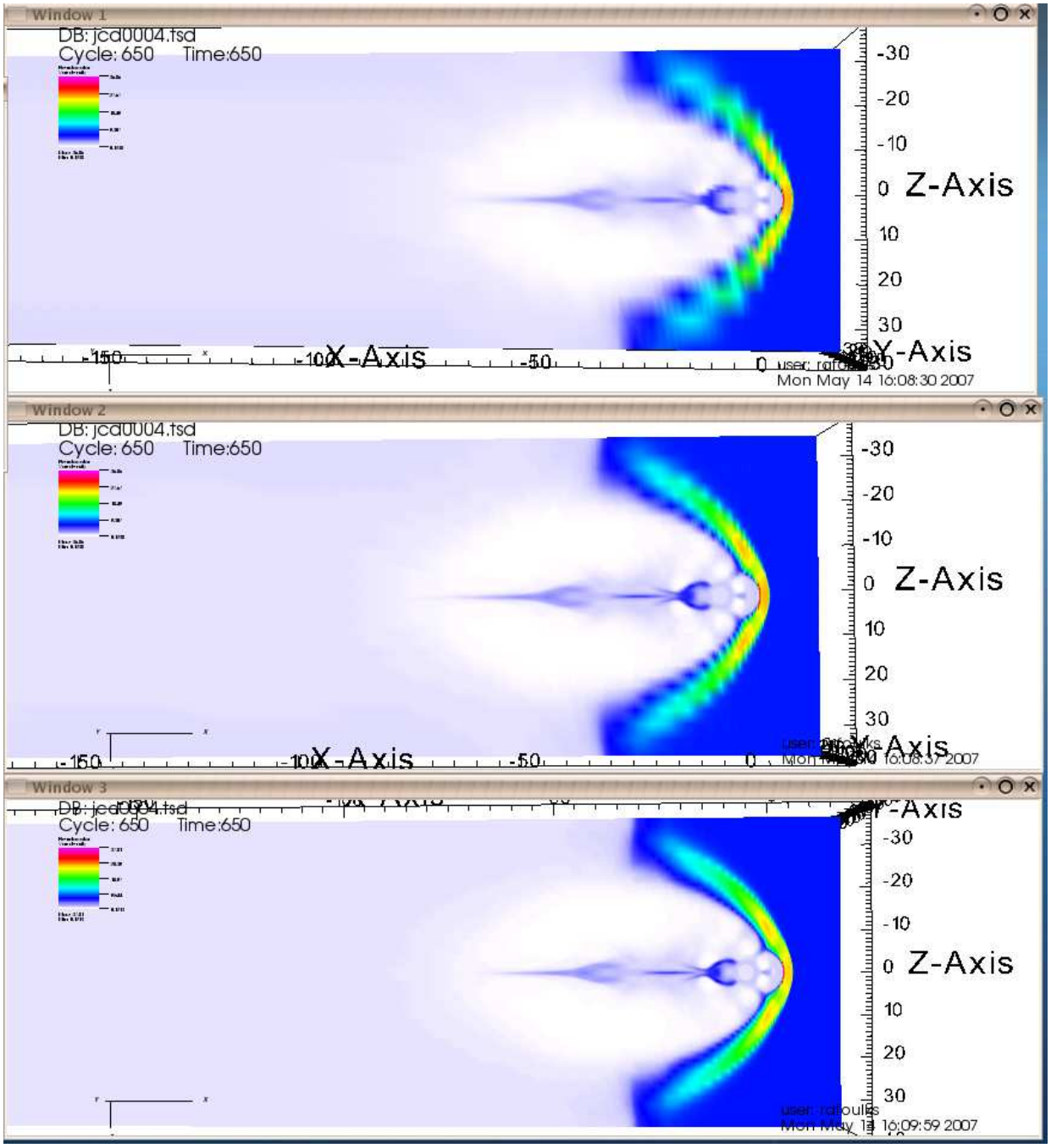
[8] F. Jiang, "Implementation and performance evaluation of block-oriented adaptive resolution for multisource scientific datasets," Master's thesis, Department of Computer Science, University of New Hampshire, Durham, New Hampshire, December 2002.

[9] A. Foulks, R. D. Bergeron, and J. P. McHugh, "Starview: A multiresolution time series data visualizer," in *Proceedings of SPIE: Visualization and Data Analysis*, vol. 5669, San Jose, California, January 2005, pp. 187–198.

[10] J. N. Leboeuf, T. Tajima, C. F. Kennel, and J. M. Dawson, "Global simulation of the time dependent magnetosphere," *Geophys. Res. Lett.*, vol. 5, no. 609, 1978.

[11] J. G. Lyon, S. H. Brecht, J. A. Fedder, and P. J. Palmadesso, "Computer simulation of a geomagnetic substorm," *Phys. Rev. Lett.*, vol. 46, no. 1038, 1981.

[12] J. Raeder, "Global mhd simulations of the dynamics of the magnetosphere: Weak and strong solar wind forcing," in *Proceedings of the Second International Conference on Substorms*, Geophysical Institute, University of Alaska Fairbanks, 1995, p. 561.



Above: Low resolution data (98x28x28)
Middle: Medium resolution data (196x56x56)
Below: High resolution data (392x112x112)

Fig. 1. MHD Density (GGCM data)

[13] W. J. Schroeder, K. M. Martin, and W. E. Lorensen, "The design and implementation of an object-oriented toolkit for 3d graphics and visualization," in *VIS '96: Proceedings of the 7th conference on Visualization '96*. Los Alamitos, CA, USA: IEEE Computer Society Press, 1996, pp. 93–ff.

[14] J. Raeder, J. Berchem, and M. Ashour-Abdalla, "The geospace environment grand challenge: Results from a global geospace circulation model," *Journal of Geophysical Research*, vol. 103, no. 14787, 1998.

[15] J. Raeder, "Global geospace modeling: Tutorial and review," in *Space Plasma Simulation, Lecture Notes in Physics*, vol. 615. Springer Verlag, Heidelberg, 2003.