Visualization of a Tornado-producing Thunderstorm: A Study of Visual Representation

David Bock
National Center for Supercomputing Applications
Urbana, IL, USA
dbock@illinois.edu

Leigh Orf
University of Wisconsin
Madison, WI, USA
leigh.orf@ssec.wisc.edu

Robert Sisneros
National Center for Supercomputing Applications
Urbana, IL, USA
sisneros@illinois.edu

ABSTRACT

We present a study of visual representations for large-scale scientific data. While non-photorealistic rendering techniques encapsulate significant potential for both research and deployment, photorealistic visualizations remain king with regard to accessibility and thereby acceptability. Therefore in this work we endeavor to understand this phenomenon via the creation of the analogous visualizations in a setting that is readily comparable. Specifically, we employ identical non-color parameters for two visualizations of scientific data, CM1 simulation output data. In the first data elements are colored according to a natural diagrammatic scale and in the other physically-based techniques are deployed. The visualizations are then presented to a domain scientist to illicit feedback toward classifying the effectiveness of the different design types.

Keywords

CM1; Scientific Visualization

1. INTRODUCTION

Visualization is utilized in practically every step of the large-scale scientific workflow, from the early stages of simulation validation on through to the high quality visualizations deployed to disseminate information to the general public. From beginning to end, this pipeline seems to increasingly involve photorealistic renderings over those that are non-photorealistic. While each stage of the process may hold equal value to any other, breakthroughs and successes are certainly associated with the later stages. Available research into photorealistic visualizations [4, 5, 12, 3, 2] belies the importance placed thereon in practice. Alternately, research in non-photorealistic techniques are quite well represented in literature; we refer the interested reader to the more than 600 references contained Sousa's indexed taxonomies [13]. In this work we create representative visualizations of each and in collaboration with a domain science representative evaluate their efficacies.

The domain best represented with regard to both research in and deployment of realistic visualizations is climate science. We focus on a subset of this domain; we visualize the data resulting from a simulation of those thunderstorms capable of producing tornadoes. Supercells are long-lived thunderstorms exhibiting persistent rotation, and while most supercells do not produce tornadoes, a small fraction of all supercells produce long-lived extremely devastating tornadoes. In a recent simulation, a long-lived EF5 tornado, on the ground for nearly 2 hours, was simulated [9, 8, 10]. This breakthrough simulation represents the first of its kind and produced nearly 100 TB of data, and is the target of ongoing study. The simulation provides an ideal testbed for comparing different visual representations of the storm, with the dual goals of providing photorealistic representations of the storm that can be compared to photography and video in the field, and elucidating features of the storm beyond what photorealism can provide.

We first describe our target application in Section 2. We follow with a discussion of visualization specifics including used hardware, deployed software, and design considerations in Section 3. We conclude with a brief discussion of insights gained from our study of visual representations in Section 4.

2. APPLICATION

The thunderstorm simulation was conducted on the Blue Waters supercomputer, utilizing CM1 [1], a cloud model designed for idealized studies of thunderstorms and related phenomena. CM1 was written to exploit massively parallel hardware architectures, utilizing non-blocking communication and a hybrid MPI-OpenMP parallelization approach. The 3D model grid for the simulation contained $2200 \times 2200 \times 380$ (1.8 billion) grid points with data saved every 1 second throughout the period explored in this work. In order to explore the inner workings of the thunderstorm at this extremely high temporal resolution, new I/O code was added to the model utilizing the HDF5 libraries, with the core driver being used to buffered writes to memory [7]. The simulation was run for 3 model hours, utilizing 20,000 integer cores, taking roughly 3 wallclock days to complete.

On each shared-memory compute node a single HDF5 file containing 50 time levels was buffered to memory before being flushed to disk. This approach drastically reduced the overhead associated with the conventional approach of having 16 cores per node write a single file during each history write cycle. Middleware was developed to provide a window into the model's native output format, which included multiple HDF5 files spread across multiple directories, and to allow easy conversion to monolithic file formats such as netCDF. We utilized this middleware to convert this data in parallel to a format more amenable to deploying our renderer in a time parallel fashion.

3. VISUALIZATION

Due to the ever-increasing size of datasets, it is becoming critically important to be able to process, render, and visualize data on the same system it is computed. For this reason, we have ported our batch renderer to the Blue Waters system. Using our graphical interface, visualization scenes can be designed and developed separately and on a local machine and sent to Blue Waters to be processed and rendered using the standard queue and batch system. In the remainder of this section, we will discuss the software behind the two visualizations created, represented by the different transfer functions of Figure 1.

3.1 Software

In order to graphically represent data from the CM1 cloud model we use an in-house, custom rendering system. This system employs a physically-based, distributed ray-tracer to visualize high-dimensional, volumetric scientific data along with other integrated scene elements represented as polygonal surfaces, quadrics, lines, and particles. Our renderer incorporates a variety of techniques [11] in order to develop representations that maximize insight and understanding of scientific data. Such techniques include the use of physically-based properties for both surfaces and volumes including bi-directional reflectance and transmission functions, volume scattering with real-world absorption, scattering, and emission properties, and volumetric phase functions. To provide additional perceptual and visual cues, our system provides the following techniques:

- ray-traced, depth-map, or deep-map shadows
- motion blur
- depth-of-field
- ambient occlusion
- environment lighting

To design and develop visualization scenarios, a graphical, front-end interface is provided for scene construction, animation, and rendering preview with additional features including:

- GPU (GLSL) hardware shading
- key-frame animation editor
- and an extensive color-map and transfer function editor.

A command-line version of the renderer allows for batch rendering of scenes on a variety of different computing environments and platforms.

Our rendering system has been used to visualize data from a variety of computational simulations including density currents, turbulent boundary layers, hurricane development, tornado-producing thunderstorms, binary neutron stars, 3D Rayleigh-Taylor instabilities, wind-farm simulations, and AMR (adaptive mesh refinement) galaxy clusters. This work has been featured in the book Visual Computing, publications including Computer Graphics World and Astronomy Magazine, a variety of papers, articles, reports, poster sessions, presentations at Supercomputing, XSEDE, and SIGGRAPH.

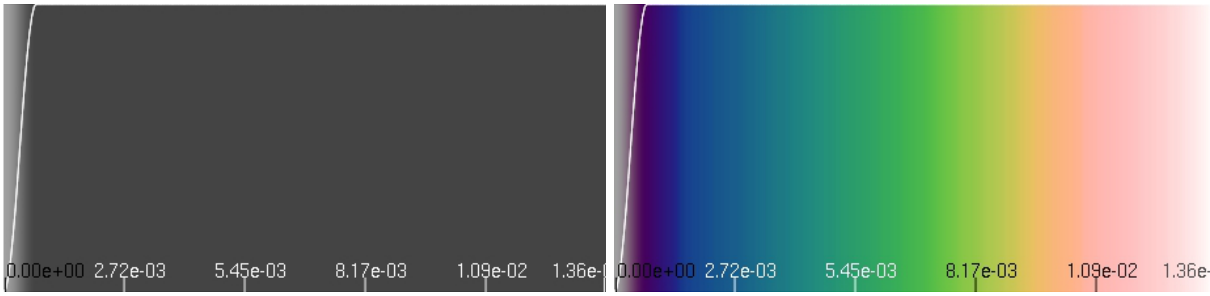
3.2 Representation Designs

"Scientific visualization is the process that transforms data to representations that leverage our innate channels of communication associated with the assimilation of information. These representations provide a more effective method in which to gain insight and understanding of the science, the simulation, and the results." [6] As a visual medium, scientific visualizations must consider the perceptual aspects of how the mind processes images and movement. As a result, critical importance must be considered when designing visualization presentations. Artistic choices of specific design elements (e.g. tone, texture, color, perspective, depth, and light and shadow) and design principles (e.g. balance, emphasis, scale, and composition) should be cognizant of the perceptual response. In addition to design choices, it is also important to consider the type of data and the simulation being represented. Computational models often simulate some real-world, physically observable phenomenon. Since models computationally simulate such physical phenomenon, accompanying visualizations might be better designed that visually simulate the same physical phenomenon while remaining true to the underlying data. Effective visualizations map symbolic data values into the more cognitive, recognizable domain.

One of the goals of our system is to provide rendering techniques with which to better visually simulate the physical phenomenon being represented. We demonstrate and compare two different visualizations from our rendering system on data from a computational model simulating a tornado-producing thunderstorm. In the first visualization, we render the data with the goal of attempting to visually simulate the physical properties associated with such a storm. We volume render the data using physically-based techniques including volumetric phase functions, volume scattering properties of absorption and scattering, lighting, self-shadowing, and ground shadows. We also carefully select the aforementioned design elements to attempt to visually simulate the actual physical phenomenon. In the second visualization, we volume render the data using a standard color-based mapping without lighting, shadows, or other physically-based properties.

4. DISCUSSION

Figure 2(a) conveys the characteristics readily observable in an actual storm. The mammatus clouds (that hang down like pouches from the anvil) especially look realistic. However, since we are not rendering snow, hail, and rain, the image itself does lack realism. Realistically, the storm almost certainly wouldn't be visible in the background since it would be obscured by heavy rain/hail from the foreground storm. Nonetheless the visualization represented in this fig-



(a) Photorealistic (b) Non Photorealistic

Figure 1: The two colormaps used for our two visual representations. The alpha transfer function is represented as the white curves in both. Alpha is used only to remove the very low values of the storm. Note that there is no color variation for the photorealistic colormap. The differences in color are purely from lighting, deep-map shadows, Mie-Murky phase function, and scattering properties.

ure is invaluable; this visualization is key to verifying the underlying physical phenomena simulated in the storm. It is only then may one infer the scientific relevance of the resulting data from a simulation run.

In Figure 2(b) one can see the sharp gradient in cloud mixing ratio that is revealed by the green colorations in the lower-center of the image. This is along the periphery of the storm's updraft. We are likely seeing evidence of very rapid condensation occurring within the updraft, which is "laid bare" to some extent at mid-levels when viewed from the south. At lower levels, the sharper gradients are obscured by the horizontally-oriented shelf-like cloud that occurs along the leading edge of the gust front, and other dynamical processes. The fact that the mammatus clouds and anvil kind of wash out in dark blue is evidence of the lower concentration of cloud ice that is spread over a larger area.

However, regarding the visualization in Figure 2(b), scientist feedback made it clear that what is shown is not scientifically exciting. The foremost suggested improvement was to apply this type of visualization to fields such as vorticity and pressure, i.e. those without clear physically observable representations. We can therefore conclude that as long as data corresponds to a process that is physical and visually observable that photorealistic techniques are the most perceptually valuable, regardless of the amount of data discernible in a visualization.

5. ACKNOWLEDGMENTS

The simulation was completed, data stored, and visualization rendered on the supercomputer employed by the Blue Waters sustained-petascale computing project, supported by NSF Grant OCI 07-25070.

6. REFERENCES

[1] G. H. Bryan and J. M. Fritsch. A benchmark simulation for moist nonhydrostatic numerical models. *Mon. Weather Rev.*, 130(12):2917–2928, 2002.

[2] Y. Dobashi, T. Yamamoto, and T. Nishita. Interactive rendering of atmospheric scattering effects using graphics hardware. In *Proceedings of the ACM SIGGRAPH/EUROGRAPHICS conference on Graphics hardware*, pages 99–107. Eurographics Association, 2002.

[3] K. Garg and S. K. Nayar. Photorealistic rendering of rain streaks. *ACM Transactions on Graphics (TOG)*, 25(3):996–1002, 2006.

[4] R. Hufnagel, M. Held, and F. Schröder. Large-scale, realistic cloud visualization based on weather forecast data. In *Proc. IASTED Int. Conf. Comput. Graph. and Imaging (CGIMâĂŹ07)*, pages 54–59, 2007.

[5] J. Kniss, C. Hansen, M. Grenier, and T. Robinson. Volume rendering multivariate data to visualize meteorological simulations: a case study. In *Proceedings of the symposium on Data Visualisation 2002 (VISSYMâĂŹ02)*, pages 189–194, 2002.

[6] B. H. McCormick, T. A. DeFanti, and M. D. Brown. Visualization in scientific computing, 1987.

[7] L. Orf. An I/O strategy for CM1 on Blue Waters. Technical report, Nov. 2013.

[8] L. Orf, B. Wilhelmson, R Lee, and C. Finley. Genesis and maintenance of a long-track EF5 tornado embedded within a simulated supercell, Nov. 2014.

[9] L. Orf, R. Wilhelmson, and L. Wicker. A numerical simulation of a Long-Track EF5 tornado embedded within a supercell, Feb. 2014.

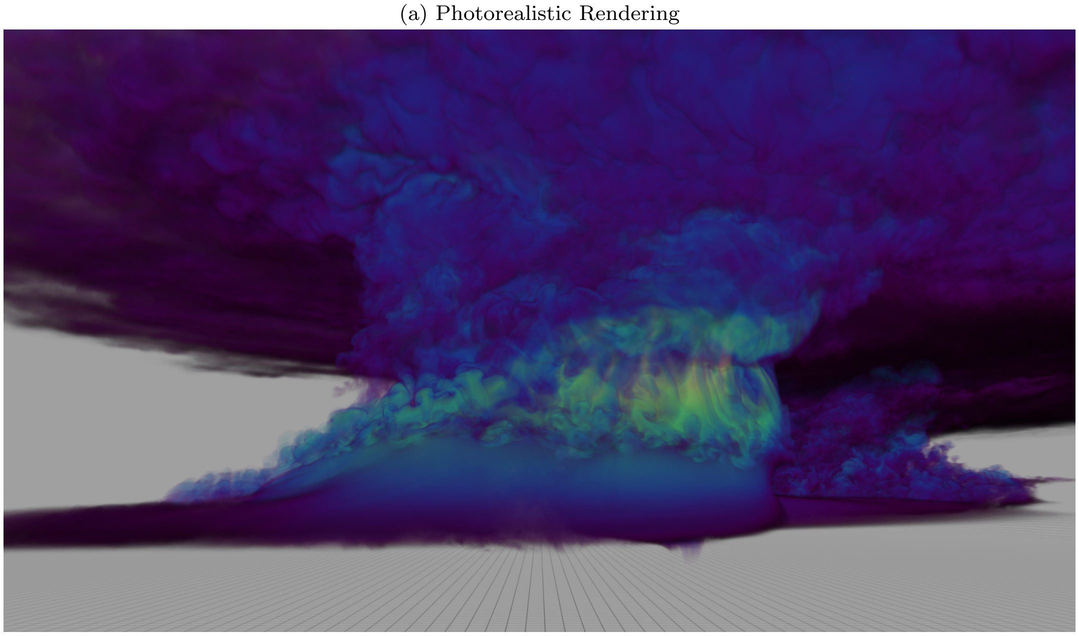
[10] L. Orf, R. Wilhelmson, and L. Wicker. Visualization of a simulated Long-Track EF5 tornado embedded within a supercell thunderstorm. In *Supercomputing*, Nov. 2014.

[11] M. Pharr and G. Humphreys. *Physically based rendering: From theory to implementation*. Morgan Kaufmann, 2004.

[12] S. Roettger and T. Ertl. Fast volumetric display of natural gaseous phenomena. In *Computer Graphics International, 2003. Proceedings*, pages 74–81. IEEE, 2003.

[13] M. C. Sousa. Indexed taxonomies of non-photorealistic rendering. *Theory and Practice of Non-Photorealistic Graphics*, 2003.

(a) Photorealistic Rendering



(b) Non Photorealistic Rendering

Figure 2: Sample frames from the video illustrating general differences between our two visual representations.