Scientific Visualization in Digital Fulldome Show "Solar Superstorms"

Advanced Visualization Lab (AVL), NCSA, University of Illinois
Donna Cox, Bob Patterson, Stuart Levy, AJ Christensen, Kalina Borkiewicz, Jeff Carpenter

ABSTRACT

The Advanced Visualization Lab worked with astrophysics data scientists to create data-driven cinematic scientific visualizations as part of a narrative journey through the dynamic processes of the sun in a digital fulldome production titled "Solar Superstorms", narrated by Benedict Cumberbatch, as part of the CADENS NSF grant. The visualizations described in this paper represent different parts of the story of the lifecycle of stars, the dynamic processes of our sun, and how they affect life on Earth. The paper discusses the challenges the AVL team faced and the solutions they devised in registering datasets from different data scientists, exploring massive supercomputer datasets in real-time virtual environments, and rendering extremely high-resolution images of spatially dense time-evolving data efficiently.

1   INTRODUCTION

The AVL is a research team that specializes in creating high-quality cinematic scientific visualizations of supercomputer simulations for public outreach. As part of the CADENS "Centrality of Advanced Digitally ENabled Science" NSF grant, the team learned that science educators were seeking material that addressed the science of the sun. In collaboration with co-PI and writer-producer Tom Lucas, the team talked with many members of the science community and curated collaborations with many data scientists across the country. Their recently completed film "Solar Superstorms" which is narrated by Benedict Cumberbatch visualizes these scientists' datasets in a narrative sequence that begins with the birth of stars in the early universe and travels through the convective and magnetically active regions of the sun to finally arrive home and observe Earth's space weather.

The following sections discuss each of the supercomputer datasets that the team visualized, what challenges they presented, and how the team addressed them.

2   EXPOSITION

2.1   The Formation of First Stars and Galaxies
Simulation data from John H. Wise of Georgia Institute of Technology's Center for Relativistic Astrophysics

John Wise's simulation of the early universe (from ~20 million - 500 million years after the Big Bang) shows formation of the first generations of stars, as their light ionizes much of the space around them. Three features were targeted in this visualization: the dense filaments along which stars form (measured by "baryon overdensity", blue-whitish where density of ordinary matter much exceeds the average); the large regions of gas which become ionized and heated as new stars begin emitting lots of ultraviolet light, and later cool off after those stars'

lives end (measured by temperature, in orange/yellow/white); and the heavier elements mixed into the gas after those dying stars explode as supernovae, forming small bubbles of high "metallicity density". A second ("supernova") simulation examines the life and death of a single first-generation ("Population III") star, looking in detail at the processes that mix the heavy elements into their environment. Those "metals" in turn make it easier for later generations of stars to form. John Wise computed both simulations using the Enzo code, on Blue Waters at NCSA.

Combining these fields in the previsualization environment, and in the final renders was a challenge. Another persistent challenge was the non-uniform cadence of the simulation output which made the visualization stutter when mapping data files directly to successive movie frames.

To create interactive previews of scientific datasets, the AVL uses an original open-source software tool called Partiview, which was primarily authored by AVL programmer Stuart Levy. Partiview is able to render particles, lines, triangle meshes, and star-like data and interactively update their visual attributes based on data attributes. To flexibly navigate through these Partiview environments, explore 3D data topology, and develop complex camera choreography in an immersive stereoscopic 4K environment, the team links Partiview with another original software tool called Virtual Director, authored by past AVL programmer Matt Hall and Marcus Thiebaux. Virtual Director is a remotely collaborative camera control and design tool capable of performing or keyframing path splines and editing them during interactive data playback.

A Partiview scene was created for the shot by visualizing the particles in John's data, which gave the team a general idea where the large scale structure would be. The team relied heavily on low-resolution render tests to perfect the camera view.

Because of the size of John Wise's data, the AVL team chose to engineer a new pipeline using the Blue Waters supercomputer to facilitate higher parallelism than their dedicated visualization cluster was capable of, to take advantage of the high speed network between the data storage location and the compute nodes, and to develop a method of cinematic production rendering of restrictively large data directly on a supercomputer.

This new supercomputer-based pipeline relied upon a supercomputation data analysis and visualization software package called "yt", and the AVL interfaced closely with founding developers Matt Turk and Sam Skillman to ensure that the image output would be as high quality and data-accurate as possible. AVL employed this experimental volume renderer and adapted this into the development pipeline. These tasks included integrating color transfer functions and Virtual Director camera path data into yt, integrating a fish-eye lens shader for digital fulldome rendering, and coordinating OpenMP functionality to take advantage of the parallelism on the Blue Waters supercomputer.

As mentioned above, integrating three data fields into a single image render was a challenge, as multi-field rendering is not a feature of yt. The AVL achieved this multi-field look through two different methods. In the supernova shot, they built a combined color ramp that represented the color transfer functions of two fields, baryon overdensity and temperature, and adjusted the combined transfer function to write the first half of the color ramp to the temperature field, and the second half to the density field. At render time, the color of the higher-normalized magnitude field was chosen. Because the multi-field colormap method did not allow

for the accumulation of color in voxels shared among fields, a 2D image compositing method was preferred for the galaxy evolution shot where that limitation was apparent. This same compositing method was used to add a metallicity layer to the supernova shot. Using a consistent camera path, each of the three fields was rendered on its own and then layered with the others. This compositing method also provided the flexibility to adjust color, contrast, and optimize visibility of the features after the time-consuming data rendering process was complete.

To compensate for the irregular cadence of the data output, the AVL used a script to arrange timesteps based on their relative position in the predetermined frame length. Data files with a low redshift were discarded, and files with high redshift were aligned to their correct spot along the timeline. To fill in the gaps, an image interpolation technique called "motion flow" was employed in the compositing software Nuke on a per layer basis. This technique was then used to create more in-between frames throughout the sequence to achieve a slow-moving embodied experience for digital fulldome audiences.

Using these techniques, the AVL was able to iteratively produce high quality cinematic renders of thousands of John Wise's multi-gigabyte datasteps in approximately 71,500 compute-node hours on the Blue Waters supercomputer.

2.2 Convection and Dynamo Action in a Sun-like Star

Simulation data from Mark Miesch of the National Center for Atmospheric Research, and Nick Nelson of Los Alamos National Labs

In a star like the sun, heat is transported outward by convection over the outermost part of the star, from about 70% of its radius to its surface. The convective circulation carries along any magnetic field which is present in the gas. This simulation looks at large-scale magnetized convection in the body of a rotating sun-like star, from about 70% to 96% of its radius. The star's rotation organizes the convective cells, which form long rolling "bananas", extending from mid-northern to mid-southern latitudes. We identify them from their high vorticity (up to one rotation per day or more) which is seen in the yellow/orange regions of the sequence. Tracer particles show the fluid circulation, colored yellow/orange by vorticity. Buoyant loops in the magnetic field, seen in the final shot in blue, are carried along by these "bananas".

Mark Miesch and Nick Nelson ran global 3D simulations of convection and dynamo action in a Sun-like star to reveal that persistent wreaths of strong magnetism can be built within the bulk of the convection zone. These examine the characteristics of buoyant magnetic structures that are self-consistently created by dynamo action and turbulent convective motions in a simulation with solar stratification. Miesch and Nelson simulated spherical grid data using the ASH simulation code at the Laboratory for Computational Dynamics at the University of Colorado, Boulder, and on Pleiades at NASA's Advanced Supercomputing Division.

The AVL team used a commercial visual effects tool called Houdini as its primary data processing environment to create derivative data assets which could demonstrate the particularly interesting features of the simulation. To allow the commercial tool to read and process scientific data, AVL programmer Kalina Borkiewicz wrote a C++ plugin using the Houdini Development Kit, that was able to read numerous file types and grid topologies, and produce Houdini-native assets that could be updated in the interactive graphic user interface.

Attributes that could be updated include different varieties of volume domain edge falloff, asset volume resolution, and asset volume type. Using this, the AVL could adjust the asset volume to low resolution while experimenting with techniques. The volume type allowed the team to read the data into a rectangular grid, or into a sparse volume grid format called OpenVDB for render and processing optimization.

Using this Houdini plugin, the AVL was able to read the velocity vector field in as an optimized sparse VDB volume and derive a scalar VDB volume for the vorticity values (|curl velocity|). Likewise, the Houdini plugin was used to create a scalar current field VDB volume (|curl B|).

Tracer particles were advected through the velocity volume to show the motion of the "bananas" and polar vortices. The tracer particles' trajectories were reshaped by subtracting out the bulk rotation of the sun at a typical depth, so as to better show the local curvature of the circulation. Working with the scientists, AVL also advected particles in the magnetic volume in areas of high magnetic magnitude and regions identified by the scientists as likely candidates to illustrate a feature called buoyant loops, where waves of magnetic field rise toward the solar surface and create sunspots. These flow lines and magnetic field lines were then exported as geometric assets to be used in the pipeline.

Three separate shots were created in the interactive Virtual Director environment using different assets. The first two focused on the vorticity data, and interactively displayed the time-evolving flowlines, as well as isosurfaces created in Houdini at varying thresholds. The third shot focused on the buoyant loops and interactively displayed the magnetic field lines and isosurface geometries.

Houdini was also used to create a data-driven solar core, using temperature and fusion energy functions from a standard solar model. Radial and spatial algorithmic noise was added to give a sense of parallax and reduce occlusion of the far side of the convection data. Flow line and magnetic field line data was splatted into sparse VDB volumes. To fit the scene in the memory of the render nodes, AVL implemented a camera view frustum culling technique to remove lines outside the camera's viewpoint, and a camera depth culling technique which allowed the team to split the lines into different depth layers of different resolutions. An external script also created a filtering process by which the team could remove lines that didn't contribute significantly to tracing the important vortices across their lifetime.

So that the shots could go above the surface of the sun, a sun surface asset was created using Matthias Rempel's solar surface convection supercomputer simulation. Matthias's data was interpolated in image compositing software, tiled to cover a larger percentage of the sun's surface, and displaced based on intensity. Matthias's data also included magnetic vectors, which were brought into Houdini as RGB values and then translated into small volumetric pointers on the surface of the sun. A halo was also put around the sun, mapped with imagery observed through NASA's Solar Dynamics Observatory telescope.

These many different datasets were rendered in several image layers from a unified camera path, and composited together using a commercial software tool called Nuke. Transitions past the solar surface utilized camera depth data and custom-drawn mattes to create an organic feeling between datasets.

2.3 Near-Solar Surface Magneto-Convection

Simulation data from Bob Stein, et al of Michigan State University

Bob Stein used results of global solar dynamo simulations to determine spatially and temporally evolving bottom boundary conditions for a magneto-convection simulation of the top of the solar convection zone (a slab of the sun's surface 48000 km wide and 20000 km deep). Magnetic flux emergence in forming solar active regions is the driver of flares and coronal mass ejections that produce the dangerous storms in Earth's space weather. The simulation ran on Pleiades at NASA's Advanced Supercomputing Division. Pat Moran, a visualization expert at NASA Ames, processed the output of these simulations through an advanced algorithm to produce spatially consistent visually meaningful field lines by selecting for lines that ultimately end up connecting with a particular active region on the sun's surface.

From the simulation's magnetic field, AVL constructed a current field (|curl B|), and visualized it volumetrically; the highest-current regions are concentrated into thin sheets, shown in yellow/orange. Embedded within this volume are Pat Moran's bundles of magnetic field lines, in blue/cyan, colored by magnetic field strength.

To create an interactive time-evolving representation of this high resolution volume for camera choreography, the team seeded a uniform grid of points in the volume and colored them by the values of the nearest voxels. Additionally, Pat Moran's field lines were rendered as geometric objects in a different color range.

In Houdini, Pat Moran's field lines were splatted into sparse VDB volumes to create an integrated volumetric appearance. Volumetric rendering of the current field was also customized in Houdini. During the simulation's course, the grid resolution doubles, and the distribution of current changes significantly with time - it didn't work well to use any fixed mapping of current magnitude to visual color and opacity. AVL tied the current's transfer function to its histogram at each time, such that quantiles defined by e.g. the top 1%, 0.5%, 0.1%, etc. of the current distribution were colored consistently.

The team's solar surface treatment utilizing Matthias Rempel's solar convection simulation and NASA's Solar Dynamics Observatory imagery was repurposed for this sequence to create a unified context across shots. The team was careful in this shot to align the sunspot in Rempel's data with the emerging active regions beneath the solar surface in Stein's simulation. Nuke was used to composite the rendered image layers.

2.4 A Homologous Sequence of Coronal Mass Ejections
Simulation data from Yuhong Fan of the National Center for Atmospheric Research

Yuhong Fan's magnetohydrodynamic simulation of the development of a homologous sequence of two coronal mass ejections (CMEs) demonstrates their so-called cannibalistic behavior. These CMEs originate from the repeated formations and partial eruptions of kink unstable flux ropes as a result of continued emergence of a twisted flux rope across the lower boundary into a pre-existing coronal potential arcade field. The simulation shows that a CME erupting into the open magnetic field created by a preceding CME has a higher speed. The second CME is cannibalistic, catching up and merging with the first into a single fast CME before exiting the domain. Yuhong Fan simulated her non-uniform spherical grid data on the Yellowstone supercomputer at the National Center for Atmospheric Research.

This simulation shows the explosive power of ionized plasma that can be projected from the sun toward Earth. AVL shows Fan's magnetic field lines in blue/cyan, traced between points on the Sun's surface, and a current field (|curl B|), computed from the simulation's magnetic field, in yellow/orange. As the CMEs expand, note that the current is concentrated into a thin, rippling sheet - reminiscent of auroras on Earth, where current sheets are also involved.

To create an interactive time-evolving representation of this high resolution volume for camera choreography, the AVL team again seeded a uniform grid of points in the volume and colored them by the values of the nearest voxels. Additionally, Yuhong Fan's field lines were rendered as geometric objects.

Because the data cadence was not as high as the frame rate of the video, the team developed data interpolation techniques to create smooth visual playback. The Houdini plugin was augmented to do temporal b-spline interpolation. This method created noticeable stepping artifacts at low resolution in a uniform grid, but by pushing the resolution much higher by implementing a sparse VDB volume representation and convolving the voxels, the artifacts were removed. The magnetic field lines were challenging to interpolate because their shapes changed dramatically between datasteps. To interpolate these, the team created a method of "rubber band interpolation" that would non-uniformly stretch the field lines so they could associate both similar shapes and parametric coordinates. After this interpolation, some lines were still so problematic that they were culled using a volatility score that was calculated in an external script. Deriving velocity from the interpolation technique still produced distracting results when motion blurred, so the team decided to calculate line "echoes" to show motion of a line over time. The lines and their echoes were accumulated in areas of overlap and then splatted into sparse VDB volumes to create an integrated volumetric appearance, colored by the magnetic magnitude on the original data lines, and faded based on an echo's age.

Matthias Rempel's solar convection simulation and NASA's Solar Dynamics Observatory imagery was used in this shot as well to provide a consistent context. The team was careful to place the center of the Fan's coronal mass ejection over Rempel's sunspot with advice from the scientists. Nuke was used to composite the rendered image layers.

2.5 Kinetic Simulation of the Magnetosphere
Simulation data from Homa Karimabadi, et al, University of California at San Diego

Homa Karimabadi's science team conducted the most detailed 2D and 3D global hybrid simulations of the interaction between solar wind and the magnetosphere. These simulations focus on how ion kinetic processes at the bow shock, a standing shock wave in front of the magnetosphere, drive turbulence in a region behind the shock called the magnetosheath. Massive electromagnetic storms on the Sun can reach the Earth, wreaking havoc on our technological systems. Homa's team simulated their data using a particle-in-cell method to produce rectangular multigrid data on the Blue Waters supercomputer.

The team faced a unique challenge in representing the scientists' data as thoroughly as possible while working within the constraints of limited data output. The scientists were able to output hundreds of timesteps of 2D slices from the volume data, but only a few dozen full 3D descriptions of the volume, all sampled from an early part of the simulation evolution. Because of the nature of the compute-expensive simulation, the scientists could not afford to re-run the

simulation to output intermediate data. After testing several techniques, the team chose a hybrid 3D and 2D approach. The sequence begins with a slow step-dissolving treatment of the 3D data, rendered from the perspective of a continuously moving camera. Then it fades into a treatment of the 2D data, played at the correct relative cadence to match it with the 3D data, and smoothly interpolated across every frame using an image-based technique called "motion flow" that is a feature of commercial compositing software package Nuke. Finally, the team used a supercomputation data analysis and visualization software package called "yt" to trace magnetic field lines through the 3D data. This line data was cycled through the shot's duration to visually associate the 3D and 2D data, and to give a sense of environment once the 3D volume data had disappeared.

Previsualization of this scene focused on the 2D data which had incredible detail, so the Partiview scene was optimized for playback on an image plane representing the 2D data. Geometric magnetic field lines, and the Earth and sun were added as well. This allowed the team to choreograph a camera path that started by facing the lit side of the Earth, and then turning to simultaneously reveal the sun and reveal the turbulence in the magnetosphere that the sun was causing.

The team faced a challenge in rendering the 3D data with embedded geometric features. After attempting a pipeline utilizing Z-buffered occlusion maps in yt that had aliasing limitations, it was discovered the 3D data did not have highly resolved features from its early stage in the simulation, and so it did not look visually different when its volume resolution was reduced. Therefore, the team used a commercial visual effects tool called Houdini to integrate all of the data assets that had been created together. The 3D data was downsampled from the raw data on the Blue Waters supercomputer into a compressed Houdini-native data format using a custom Python script. The magnetic field lines were filtered for artifacts and then splatted into sparse VDB volumes to create an integrated volumetric appearance. The 2D data was rendered on a plane and augmented with a color transfer function and falloff mattes to make it appear to be an organic flying 2D bow shock, removing artifacts that suggested it was a 2D slice from a rectangular grid. The sun was constructed from a glare image oriented toward the camera. The Earth was textured with night and day photography on opposite sides, oriented to show the camera as much human civilization as possible, and an atmosphere was added.

The 3D data and line data were step-dissolved in Nuke using two full render passes with the datasteps updating every second, one with the datasteps offset by one step. These two layers with matching camera positions allowed the team to cross-fade each adjacent pair of datasteps slowly over the course of a full second. Other image layers including the 2D data, sun, Earth, and background starfield were then included in the final composite.

ACKNOWLEDGEMENTS