Visual Insights into High-Resolution Earthquake Simulations

Amit Chourasia, Steve Cutchin, Yifeng Cui, and Reagan W. Moore
San Diego Supercomputer Center

Kim Olsen and Steven M. Day
San Diego State University

J. Bernard Minster
Scripps Institution of Oceanography

Philip Maechling and Thomas H. Jordan
Southern California Earthquake Center

This study focuses on the visualization of a series of large earthquake simulations collectively called TeraShake. The simulation series aims to assess the impact of San Andreas Fault earthquake scenarios in Southern California. We discuss the role of visualization in gaining scientific insight and aiding unexpected discoveries.

Each year, estimated losses from earthquakes are measured in billions of dollars and thousands of fatalities.[1] While researchers have studied earthquakes for a long period of time, recent advances in computation now enable large-scale simulations of seismic wave propagation.

Because there's no doubt about the utility of developing a system to help us better understand earthquakes, we've done extensive work on what we call the TeraShake project (see http://epicenter.usc.edu/cmeportal/TeraShake.html). In this project, we used a multitude of visualizations to interpret simulation output, and we recently completed one of the first large-scale simulations done on the Southern San Andreas Fault.

Earthquake threat

The San Andreas Fault has a history of producing earthquakes with moment magnitude exceeding 7.5.[2] The most recent was the 1906 magnitude 7.9 event[3] that caused tremendous damage to San Francisco. The 1857 (Fort Tejon) earthquake ruptured a 360-kilometer stretch from Parkfield to Wrightwood. However, the San Bernardino Mountains segment and the Coachella Valley segment of the San Andreas Fault haven't seen a major event since 1812 and 1690, respectively.[4]

Average recurrence intervals for large earthquakes with surface rupture on these segments are around 146 +91 – 60 years and 220 ±13 years, respectively.[5] A major component of the seismic hazard in Southern California and northern Mexico arises from the possibility of a large earthquake on this part of the San Andreas Fault.[6] Because no strike-slip earthquake of similar or larger magnitude has occurred since the first deployment of strong-motion instruments in Southern California, much uncertainty exists about the ground motions expected from such an event. Compounding this problem is the fact that this is a densely populated area. We therefore conducted our TeraShake earthquake simulation in an effort to try and reduce this uncertainty.[7]

TeraShake simulation

TeraShake is a large-scale finite-difference (fourth-order) simulation of an earthquake[8] event based on Olsen's Anelastic Wave Propagation Model code, conducted in the context of the Southern California Earthquake Center (SCEC) Community Modeling Environment. The 600 × 300 × 80-km simulation domain (see Figure 1) extends from the Ventura Basin and Tehachapi region to the north and to Mexicali and Tijuana to the south. It includes all major population centers in Southern California, and is modeled at 200-meter resolution using a rectangular, 1.8-giganode, 3,000 × 1,500 × 400 mesh.

The simulated duration is 250 seconds, with a temporal resolution of 0.011 seconds and a maximum frequency of 0.5 Hz, for a total of 22,727 time steps. We conducted this study in two phases (see Figure 2). We modeled TeraShake-1 using a kinematic source description (see the bottom image in Figure 2b) for the San Andreas Fault rupture. Four production runs with different parameter and input settings were made for phase one. TeraShake-2 (Figure 3b on page 30) added a physics-based dynamic rupture component to the simulation, which we ran at a high 100-meter resolution, to create the fault's earthquake source. This is physically more realistic than the kinematic source description (see Figure 2b) used in TeraShake-1. We ran three production runs for TeraShake-2 with various parameter settings.

Data products and challenges

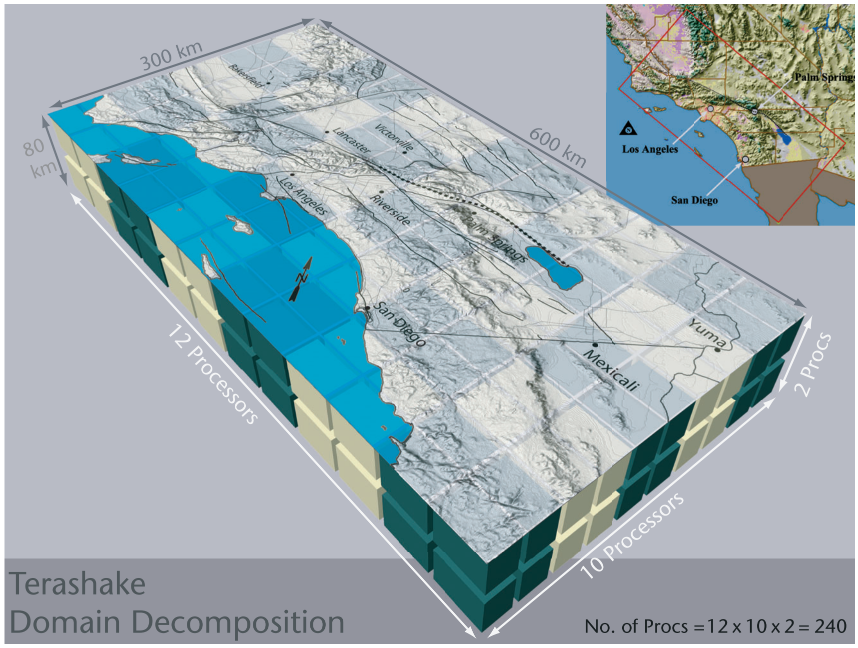
To date, all of the TeraShake simulations have consumed 800,000 CPU hours on the San Diego Supercomputer Center's (SDSC's) 15.6-teraflop Datastar supercomputer and generated 100 Tbytes of output. The TeraShake application runs on 240 to 1,600 processors

 Published by the IEEE Computer Society

and generates up to about 100,000 files and 10 to 50 Tbytes of data per simulation.

The output data consists of a range of data products including surface velocity components (2D in space, 1D in time) and volumetric velocity components (3D in space, 1D in time). The system records surface data for every time step and records volumetric data for every 10th or 100th simulation time step. It derives other products (including velocity magnitude, peak cumulative velocity, displacement, and spectral acceleration) from the output data.

This simulation data poses significant challenges for analysis. The foremost need is to computationally verify the simulation progress at runtime and thereafter seismologically assess the computed data. After completing the simulation run, the results from the multivariate data needs to be intuitively and quickly accessible to scientists at different geographical locations. Features of interest or key regions should be identified for further analysis. This requires that we integrate the geographical context with the data when it's presented.



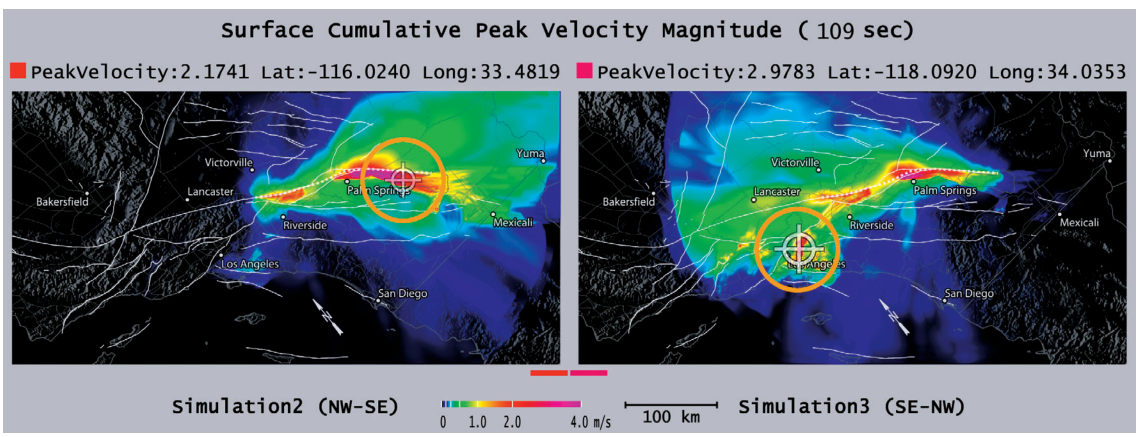
1 The top right inset shows the simulation region 600 km long and 300 km wide, indicated by the red rectangle. In the center the topography, fault lines, and city locations are visible. This also shows the domain decomposition of this region into 240 processors.

Visualization process

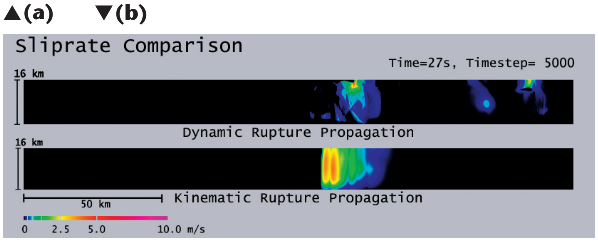
From the beginning of the study, we've collaborated on the visualization process. At the first stage, we color mapped the surface data, and it went through several iterations of refinement, based on feedback from scientists, to capture and highlight the desired data range and features. Because most of the data is bidirectional (velocity components in positive and negative directions), we finally chose the color ramp to have hot and cold colors on both ends. We decided that the uninteresting range in the middle would be black (see Figure 3a).

The next step was to provide contextual information; this process also underwent several revisions as better overlay maps become available with time (see Figure 3b). Concurrently, we designed different transfer functions to capture features of interest through direct volumetric rendering (see Figure 3c). Finally, we developed alternate methods like topography deformation (see Figure 3d) and self-contouring (see Figure 4a) to aid in analysis.

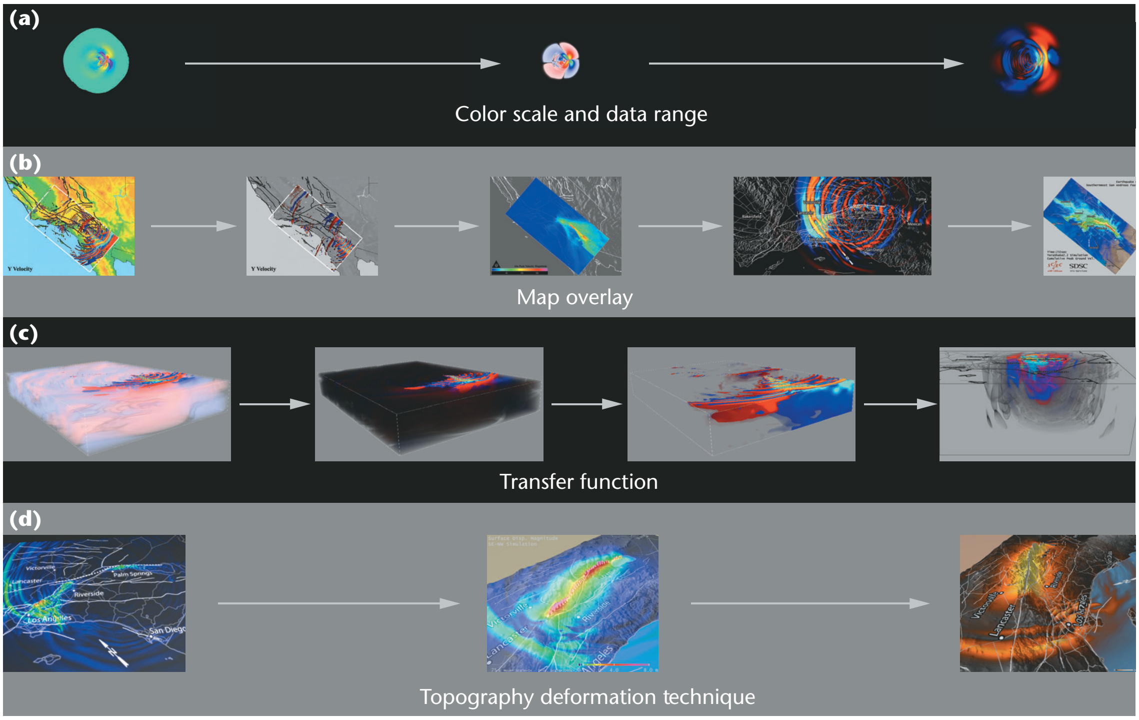
We experienced several challenges along the way. The foremost difficulty we experienced was handling the



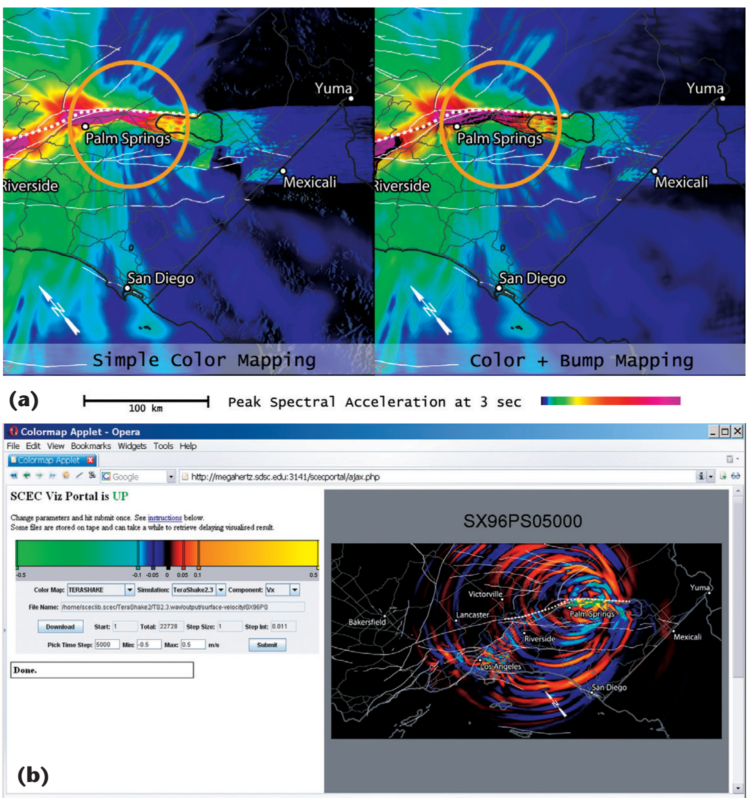
▲(a) ▼(b)



2 (a) Image comparing maps of maximum ground velocity in two different simulated earthquake scenarios on the San Andreas Fault. Simulation2 is for rupture toward the southeast, while Simulation3 is for rupture toward the northwest. (b) Snap shots of fault-plane velocity for TeraShake-2 (top) and Terashake-1 (bottom), respectively, showing the differences in source characteristics for the two cases.



3 **Iterative refinements of the visualization incorporating feedback from scientists.**



4 **(a) Map showing the advantage of the self-contouring technique in contrast to the simple color mapping. In the region highlighted by the orange circle, scientists identified a star-burst pattern. This indicates an unusual radiation of energy worthy of further investigation, which went unnoticed with simple color mapping. (b) Screen shot of the visualization Web portal. Users can select different data sets and change visualization parameters to create data maps.**

sheer size of data ranging from 10 to 50 Tbytes for each simulation. Keeping this data available for a prolonged period on a shared file system or moving it around was impractical. Interactive visualization tools for this capacity of temporal data are either virtually nonexistent or require specialized dedicated hardware. Furthermore, the disparate geographic location of scientists also made these tasks difficult. Thus, we settled on providing packaged animations through the Web, which we refined over time through feedback.

Visualization techniques

We used existing visualization techniques and combined them with off-the-shelf software to create meaningful imagery from the data set. We classify our visualization process into four categories: surface, topographic, volumetric, and static maps.

Surface visualization

In this technique, we processed the 2D surface data via direct 24-bit color maps overlaid with contextual geographic information. Annotations and captions provided additional explanations. We encoded the temporal sequence of these images into an animation for general dissemination. We used Adobe's After Effects for compositing and encoding the image sequences. Layering the geographic information with high-resolution simulation results provided precise, insightful, intuitive, and rapid access to complex information. Seismologists need this to clearly identify ground motion wave-field patterns and the regions most likely to be affected in San Andreas Fault earthquakes. We created surface visualizations for all 2D data products using this method.

We also needed to compare multiple earthquake scenarios to understand rupture behavior. In Figure 2a, the left image shows the areas affected by a rupture traveling northwest to southeast; the right image shows the corresponding results for a rupture moving southeast to northwest. Both images have geographical and contextual information (fault lines and freeways) overlain. We also show seismic information of time, peak velocity, and instantaneous location visually via a graphical cursor and in text. The goal is to gain an understanding of the region most heavily impacted by such an earthquake and the degree of damage.

The side-by-side approach for this visualization provides scientists with an important visual intuition regarding similarities and differences between fault rupture scenarios. In particular, such comparisons reveal significant differences in the ground motion pattern for different rupture directions, and in one case, it shows wave-guide effects leading to strong, localized amplification. By using animations, the scientists were able in some instances to detect instabilities in the absorbing boundary conditions, and to identify alternate conditions to remedy the problem (see Figure 2b). In other instances, specific physical behaviors were observed—for example, that the rupture velocity was exceeding the S wave speed at some locations. Such supershear rupturing is of great interest to seismologists, because it generates different types of ground motions than subshear ruptures.[9]

Topographic visualization

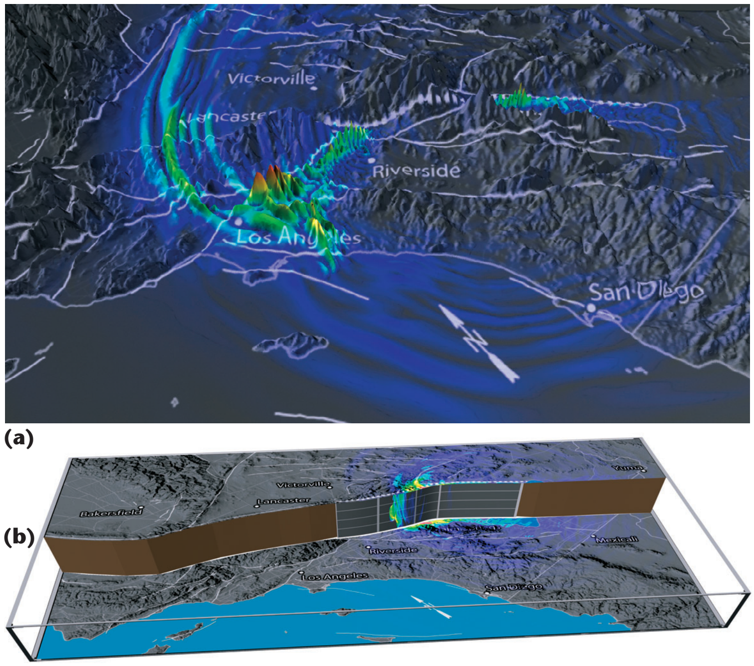
This process uses dual encoding of the surface velocity data as both color mapping and as displacement mapping. We used the surface velocity data to create a color-mapped image; and the displacement magnitude calculated from the surface velocity data to generate a gray-scale image. The system uses the gray-scale image as a displacement map to create terrain deformation along the vertical axis (see Figure 5a) using Autodesk's Maya (see http://www.autodesk.com/maya).

The animation corresponding to Figure 5a lets the scientist gain a better understanding of the kind of waves propagating the model. Another example is use of cross-sectional views (see Figure 5b). In this view, we lower the southwest side of the fault's surface to the bottom so that we can see the rupture on the fault. This kind of visualization helps seismologists connect surprising features in the ground motion with features in the rupture propagation. Currently, we're working to develop a method to display true 3D deformations based on three components of surface velocity data to provide more realistic insight.

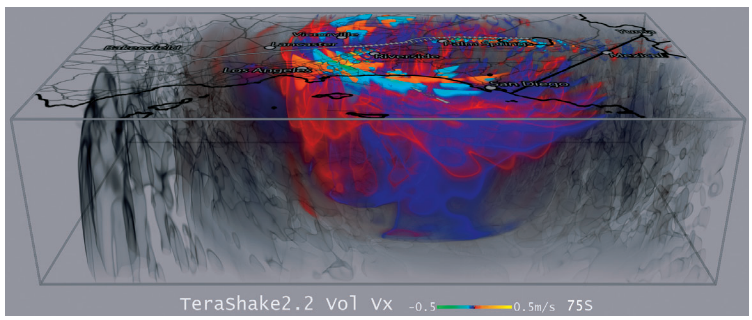
Volumetric visualization

The bulk of the data our simulation yielded is volumetric. This is by far the most significant for analysis, as it holds the most information content. We performed direct volume rendering[10] of the volumetric data set and composited it with contextual information to provide a holistic view of the earthquake rupture and radiated waves (see Figure 6).

Our initial work helped the seismologists see the general depth extent of the waves. For example, depending on the wavefield's component, waves propagating pre-



5 (a) Snap shot showing deformation along the vertical axis of the terrain-based velocity magnitudes. (b) Cross-sectional view showing the slip rate on the vertical fault plane; the velocity magnitude and surface topography are cut in two halves on the horizontal plane.



6 Snap shot showing volume-rendered velocity in *y* (shorter horizontal) direction. The *z* axis (vertical) has been scaled twice.

dominantly in the shallow layers can be identified as surface waves. Such waves typically contain large amplitude and long duration and can be particularly dangerous to certain structures. However, more research in addition to existing work[11-13] needs to be done to represent multivariate data in a unified visual format.

Additional challenges in volumetric visualization include how to visually present to the user a global understanding of the seismic waves' behavior while at the same time allowing them to examine and focus on localized seismic activity. This challenge is important, as often only reviewing the global behavior of the seismic wave hides important localized behaviors while at the same time simply focusing on localized activity often hides how this localized motion impacts the waves' overall behavior.

Static maps

Some data products like spectral acceleration depict the ground's vibration characteristics at different frequencies (see Figure 4a). Peak ground velocities and peak ground displacements are nontemporal and require visual representation for better understanding.

Self-contoured maps. We developed a technique to highlight features in 2D by using bump mapping. (Wijk[14] provides a detailed analysis of this approach.) Encoding the spectral acceleration levels using both color maps and bump maps reveals subtle transitions between ground motion levels within localized regions with similar spectral acceleration properties. The color and bump encoding technique brought out variations in the data that weren't previously visible (see Figure 4a).

Watching the earthquake simulation evolve over time helps viewers gain insight into both wave propagation and fault rupture processes.

Web portal. We wanted to ensure that TeraShake could offer scientists hands-on analysis capabilities. This is important, because scientists need to be able to take this kind of approach to gain a better understanding of the output data. However, the size of TeraShake's data poses a significant problem for accessibility and analysis. We therefore developed a Web front end where scientists can download the data and create custom visualizations over the Web directly from the surface data. The portal uses Linux, Apache, PHP, and Java technology for Web middleware. On the back end it relies on specialized programs to fetch data from the archive, visualize, composite, annotate, and make it available to the client browser.

Visualization tools and results

We used SDSC's volume rendering tool Vista, based on the Scalable Visualization Toolkit, for visualization rendering. Vista employs ray casting[10] with early ray termination for performing volumetric renderings. We visualized surface and volume data with different variables (velocities and displacements) and data ranges in multiple modes. The resulting animations have proven valuable not only to domain scientists but also to a broader audience by providing an intuitive way to understand the results. The visualizations required significant computational resources. So far, the visualizations alone have consumed more than 10,000 CPU hours and more than 30,000 CPU hours on SDSC's Datastar and TeraGrid IA-64, respectively.

The use of multiple data sets with different visual representations (see Figures 2a, 2b, 4a, 4b, 5, and 6) help the seismologists understand key earthquake concepts like seismic wave propagation, rupture directivity, peak ground motions, and the duration of shaking. The strong visual impact leads the viewer from the global context of earthquake hazards to the hazards in a specific region and then into the details about a specific earthquake simulation. Watching the earthquake simulation evolve over time helps viewers gain insight into both wave propagation and fault rupture processes. The simulation also illustrates the earthquake phenomena in an effective way to nonscientists. We've performed more than 100 visualization runs, each utilizing between 8 and 256 processors in a distributed manner. The results have produced more than 130,000 images and more than 60 unique animations (available at http://visservices.sdsc.edu/projects/scec/terashake/).

Insights gained through visualization

Large-scale simulations are playing an increasing role in the understanding of regional earthquake hazard and risk. These developments are comparable to those in climate studies, where the largest, most complex general circulation models are being used to predict the effects of anthropogenic global change. By visualizing the TeraShake simulations, scientists have been able to gain the following insights:[15]

- In a northwestward propagating rupture scenario, the wave propagation is strongly guided toward the Los Angeles basin after leaving the San Andreas Fault (unexpected). (See the image on the right in Figure 2b and Figure 5a.)
- The sediment-filled basin acts as an amplification source for trapped waves. A strong amplification is observed in the LA basin long after the initial rupture (unexpected). (See the image on the right in Figure 2b and Figure 5a.)
- Contiguous basins act as energy channels, enhancing ground motion in parts of the San Gabriel and Los Angeles basins (see the image on the right in Figure 2b and Figure 5a).
- The visualizations identify regions with particularly strong shaking (see Figure 2).
- The visualizations validate an input rupture model and instability identification.
- The visualizations observe star-burst patterns in the spectral amplification maps (unexpected). (See Figure 4a.)

Discussion

The role of contextual information has been pivotal; the surface visualizations have been quite helpful to scientists. However, encoding the rendered image sequence into animations has been a bottleneck, because of the serial, time-consuming, and lossy compression process. Further, it requires carefully selecting codecs for broader accessibility by scientists on different platforms.

In the past, researchers have successfully used some visualization techniques—such as plotting the surface peak ground motions on a map—to analyze wave propagation simulations. Thus, such basic techniques have proven to provide useful scientific information and are useful for initial simulation assessment. However, without the exploratory visualizations applied to the TeraShake simulations, such as the color and bump encoding technique, the origin of some features in the results would have been unclear. These new methods

An Earth-Shaking Moment

To see the rippling effect of a major earthquake in Southern California visit http://csdl.computer.org/comp/mags/cg/2007/extras/g5028x1/mov.

should be considered in future large-scale simulations of earthquake phenomena. While these methods are under development, future efforts should concentrate on documentation of the procedures to promote widespread use.

In addition to the scientific insight gained, these methods can provide important instructional learning materials to the public. For instance, Project 3D-View, a NASA-funded program, uses a TeraShake animation for teacher training; TeraShake could potentially become a part of the curriculum in thousands of classrooms. Also, the National Geographic Channel featured one of the animations in their documentary, *LA's Future Quake*.

Because large temporal data sets pose significant challenges for analysis, it's better if we can automate visualization techniques and methods applied in a planned way. Interactive visualization of large temporal data sets seems useful but is nontrivial and often impractical. Domain-specific feature-capturing algorithms coupled with visualization could play an important role for analysis. Animations, though noninteractive, can often serve the purpose for gaining insight when created in a thoughtful manner. Off-the-shelf software such as Maya can augment scientific visualization tools. And, multiple representations of the same data products with different techniques can be valuable assets. Using high-dynamic-range (HDR) imagery for amplifying fidelity and precision in visualizations seems promising but the lack of ubiquitous HDR display hardware and the plethora of tone-mapping methods make this task difficult.

Future directions

Planned future work includes online analysis of surface data by remote Web clients plotting synthetic seismograms. We also plan to investigate data mining operations, spectral analysis, and data subsetting.

The TeraShake simulation project has provided some insights on the IT infrastructure needed to advance computational geosciences, which we'll examine further. We would like to use 16-bit imagery instead of our current 8-bit imagery to increase visualization fidelity. Additionally, we want to integrate a GIS database to overlay a variety of contextual information that would enable further analysis. We're currently testing our integration of surface imagery with virtual globes like Google Earth and NASA's World Wind. ■

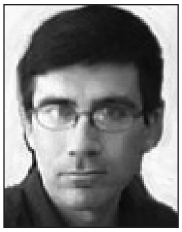
Acknowledgments

We thank Marcio Faerman, Yuanfang Hu, Jing Zhu, and Patrick Yau for their contributions. We also thank the scientists and engineers from SDSC and SCEC who funded this work and made it possible. SCEC is funded by the US NSF Cooperative Agreement EAR-0106924 and US Geological Survey Cooperative Agreement 02HQAG0008. The SCEC contribution number for this article is 1085.

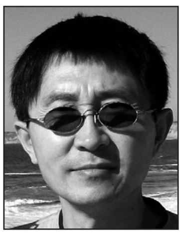
References

1. Federal Emergency Management Agency (FEMA), "HAZUS 99 Estimated Annualized Earthquake Losses for the United States," *FEMA Report 366*, Feb. 2001, p. 33.
2. K. Sieh, "Prehistoric Large Earthquakes Produced by Slip on the San Andreas Fault at Pallett Creek, California," *J. Geophysical Research,* vol. 83, no. B8, 1978, pp. 3907-3939.
3. B.A. Bolt, "The Focus of the 1906 California Earthquake," *Bulletin Seismological Soc. Am.*, vol. 68, no. 1, 1968, pp. 457-471.
4. R. Weldon et al., "Wrightwood and the Earthquake Cycle: What a Long Recurrence Record Tells Us about How Faults Work," *Geological Seismology Am. Today*, vol. 14, no. 9, 2004, pp. 4-10.
5. Working Group on California Earthquake Probabilities, "Seismic Hazards in Southern California: Probable Earthquakes, 1994 to 2024," *Bulletin Seismological Soc. Am.*, vol. 85, no. 2, 1995, pp. 379-439.
6. A.D. Frankel et al., "Documentation for the 2002 Update of the National Seismic Hazard Maps," *US Geological Survey Open File Report*, 2002, pp. 020-420.
7. K.B. Olsen et al., "Strong Shaking in Los Angeles Expected from Southern San Andreas Earthquake," *Geophysical Research Letters,* vol. 33, citation no. L07305, 2006; doi:10.1029/2005GL025472.
8. Y. Cui et al., "Enabling Very-Large Scale Earthquake Simulations on Parallel Machines," *Proc. Int'l Conf. Computational Science, Part I,* LNCS 4487, Springer, 2007, pp. 46-53.
9. E.M. Dunham and R.J. Archuleta, "Near-Source Ground Motion from Steady State Dynamic Rupture Pulses," *Geophysical Research Letters*, vol. 32, citation no. L03302, 2005; doi:10.1029/2004GL021793.
10. J.T. Kajiya and B.P.V. Herzen, "Ray Tracing Volume Densities," *Proc. Siggraph,* ACM Press, 1984, pp. 165-174.
11. P. Chopra, J. Meyer, and A. Fernandez, "Immersive Volume Visualization of Seismic Simulations: A Case Study of Techniques Invented and Lessons Learned," *Proc. IEEE Visualization*, IEEE CS Press, 2002, pp. 171-178.
12. H. Yu, K.L. Ma, and J. Welling, "A Parallel Visualization Pipeline for Terascale Earthquake Simulations," *Proc. ACM/IEEE SC Conf.*, IEEE CS Press, 2004, p. 49.
13. A. Uemura, C. Watanabe, and K. Joe, "Visualization of Seismic Wave Data by Volume Rendering and Its Application to an Interactive Query Tool," *Proc. Parallel and Distributed Processing Techniques and Applications*, CSREA Press, 2004, pp. 366-372.
14. J.J. vanWijk and A.C. Telea, "Enridged Contour Maps," *Proc. IEEE Visualization,* IEEE CS Press, 2001, pp. 69-74.
15. C. North, "Toward Measuring Visualization Insight," *IEEE Computer Graphics and Applications,* vol. 26, no. 3, 2006, pp. 6-9.

Amit Chourasia *is a visualization scientist at the San Diego Supercomputer Center (SDSC) at the University of California, San Diego (UCSD). His research interests are in scientific visualization, visual perception, and computer graphics. Chourasia received an MS in computer graphics technology from Purdue University. He's a member of the Geological Society of America. Contact him at amit@sdsc.edu.*



Steve Cutchin *manages the visualization services group at SDSC. His research interests include scientific visualizations of natural phenomena, volume rendering, and techniques for improved perceptual representation of large-scale data sets. Cutchin earned his doctorate studying collaborative systems at Purdue University. Contact him at cutchin@ sdsc.edu.*

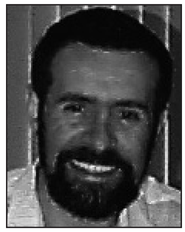


Yifeng Cui *is a computational scientist at SDSC. His research interests include high-performance computing on both massively parallel and vector machines, large-scale distributed modeling, data-intensive computing, performance optimization and evaluation, tracer hydrology, climatology, and hydroclimatology. Cui received a PhD in hydrology from the University of Freiburg, Germany. Contact him at yfcui@sdsc.edu.*

Reagan W. Moore *is the director of data and knowledge systems at SDSC. He coordinates research efforts in data grids, digital libraries, and persistent archives. His current research is on the development of rule-based data management systems. Moore is also the principal investigator for the development of the Storage Resource Broker data grid, which supports internationally shared collections. Moore also has a PhD in plasma physics from UCSD. Contact him at moore@sdsc.edu.*

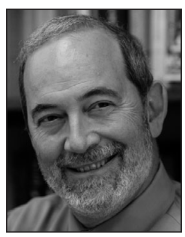
Kim Olsen *is an associate professor at the Department of Geological Sciences at San Diego State University (SDSU). His research interests include 3D simulation of wave propagation, strong ground motion, earthquake dynamics, parallel and high-performance computing, and visualization. Olsen has a PhD in geophysics from the University of Utah. He's a member of the Seismological Society of America and the American Geophysical Union. Contact him at kbolsen@sciences.sdsu.edu.*

Steven M. Day *is the Eckis Professor of Seismology at SDSU and a visiting research geophysicist at UCSD. His principal research interest is computational earthquake physics. Day has a PhD in earth science from UCSD, and he serves on the board of the Southern California Earthquake Center (SCEC). Contact him at day@moho.sdsu.edu.*



J. Bernard Minster *is a professor of geophysics at the Institute of Geophysics and Planetary Physics (IGPP) of the Scripps Institution of Oceanography, UCSD, and senior fellow at the San Diego Supercomputer Center. His research interests include determining the structure of the Earth's interior and imaging the Earth's mantle and crust using seismic waves. Minster has a PhD in geophysics from the California Institute of Technology. Contact him at jbminster@ ucsd.edu.*

Philip Maechling *is the information technology architect at SCEC. His research interests include the application of high-performance computing to seismic hazard analysis, real-time earthquake monitoring, and scientific workflow technologies. Maechling received his BS in applied physics from Xavier University in Cincinnati, Ohio. He has authored and coauthored journal articles and book chapters on computer science research topics including distributed object technology, grid computing, and scientific workflow technologies. Contact him at maechlin@usc.edu.*



Thomas H. Jordan *is the director of the SCEC and the W.M. Keck Foundation Professor of Earth Sciences at the University of Southern California. He studies earthquakes and earth structure. Jordan received a PhD from the California Institute of Technology. Contact him at tjordan@usc.edu.*