Editors: Jim X. Chen, jchen@cs.gmu.edu
R. Bowen Loftin, loftin@tamug.edu

VISUALIZATION RESEARCH CHALLENGES

A Report Summary

By Robert Moorhead, Chris Johnson, Tamara Munzner, Hanspeter Pfister, Penny Rheingans, and Terry S. Yoo

NEARLY 20 YEARS AGO, THE US NATIONAL SCIENCE FOUNDATION (NSF) CONVENED A PANEL TO REPORT ON VISUALIZATION'S POTENTIAL AS A NEW TECHNOLOGY.[1] IN 2004, THE NSF AND THE US NATIONAL INSTITUTES OF HEALTH (NIH) convened the Visualization Research Challenges (VRC) Executive Committee—made up of this article's authors—to write a new report. Here, we summarize that report, available in full at http://tab.computer.org/vgtc/vrc/ and in print.[2]

The VRC report aims to

- evaluate the progress of the maturing visualization field,
- help focus and direct future research projects, and
- provide guidance on how to apportion national resources as research challenges change rapidly in the fast-paced information technology world.

We explore the state of the field, examine visualization's potential impact on areas of national and international importance, and present our findings and recommendations for the future of this growing discipline. Our audience is twofold: visualization research's supporters, sponsors, and application users on the one hand, and visualization researchers and practitioners on the other. Our findings and recommendations reflect information gathered from visualization and applications scientists during two VRC workshops, as well as input from the larger visualization community.

Visualization's Value

Advances in computing science and technology have engendered unprecedented improvements in scientific, biomedical, and engineering research; defense and national security; and industrial innovation. Continuing and accelerating these advancements will require comprehending vast amounts of data and information produced from myriad sources.[3] Visualization—namely, helping people explore or explain data through software systems that provide a static or interactive visual representation—is critical to achieving this goal. Visualization designers can exploit the high-bandwidth channel of human visual perception and help us comprehend information orders of magnitude more quickly than we would reading raw numbers or text.

Visualization is fundamental to understanding models of complex phenomena, such as multilevel models of human physiology from DNA to whole organs, multicentury climate shifts, international financial markets, or multidimensional simulations of airflow past a jet wing. The "Characterizing Flow Visualization Methods" sidebar summarizes a study of several flow visualization techniques, whereas the "Fusion Plasma Physics" sidebar explains how visualization enables experiments of complex phenomena. Visualization reduces and refines data streams, letting us winnow huge volumes of data in applications—for example, public health surveillance at a regional or national level that tracks the spread of infectious diseases. The "Virus Structures" sidebar shows how visualization helps structural biologists understand how virus structure correlates with strength. Visualizations of such application problems as hurricane dynamics and biomedical imaging are generating new knowledge that crosses traditional disciplinary boundaries. Visualization can also provide industry with a competitive edge by transforming business and engineering practices into more understandable procedures.

Although well-designed visualizations can help people enormously, naive visualization attempts are all too often ineffective or even actively misleading. Designing effective visualizations is a complex process that requires a sophisticated understanding of human information-processing capabilities, both visual and cognitive, and a solid grounding in the considerable body of work that already exists in the visualization field. Further research in visualization—and the transfer of effective visualization methodologies into the working practice of medicine, science, engineering, and business—will be critical in handling the ongoing information explosion.

Although visualization is itself a discipline, advances in visualization lead

 Copublished by the IEEE CS and the AIP 1521-9615/06/$20.00 © 2006 IEEE

Characterizing Flow Visualization Methods

For decades, researchers have been developing and publishing visualization techniques that advance the state of the art. However, disproportionately few quantitative studies compare visualization techniques. Figure A shows the differences between flow visualization methods, showing six methods for visualizing the same 2D vector field.[1] Subjects who participated in the user study performed several tasks, including identifying the type and location of critical points in visualizations. Assuming roughly equal importance for all tasks, the streamline visualization performed best overall: on average, subjects were fastest and most accurate when using it.

This study produced both quantitative results and a basis for comparing other visualization methods, creating more effective methods, and defining additional tasks to further understand trade-offs among methods. A future challenge is to develop evaluation methods for more complex 3D time-varying flows.

Reference

1. D.H. Laidlaw et al., "Comparing 2D Vector Field Visualization Methods: A User Study," *IEEE Trans. Visualization and Computer Graphics*, vol. 11, no. 1, 2005, pp. 59–70.

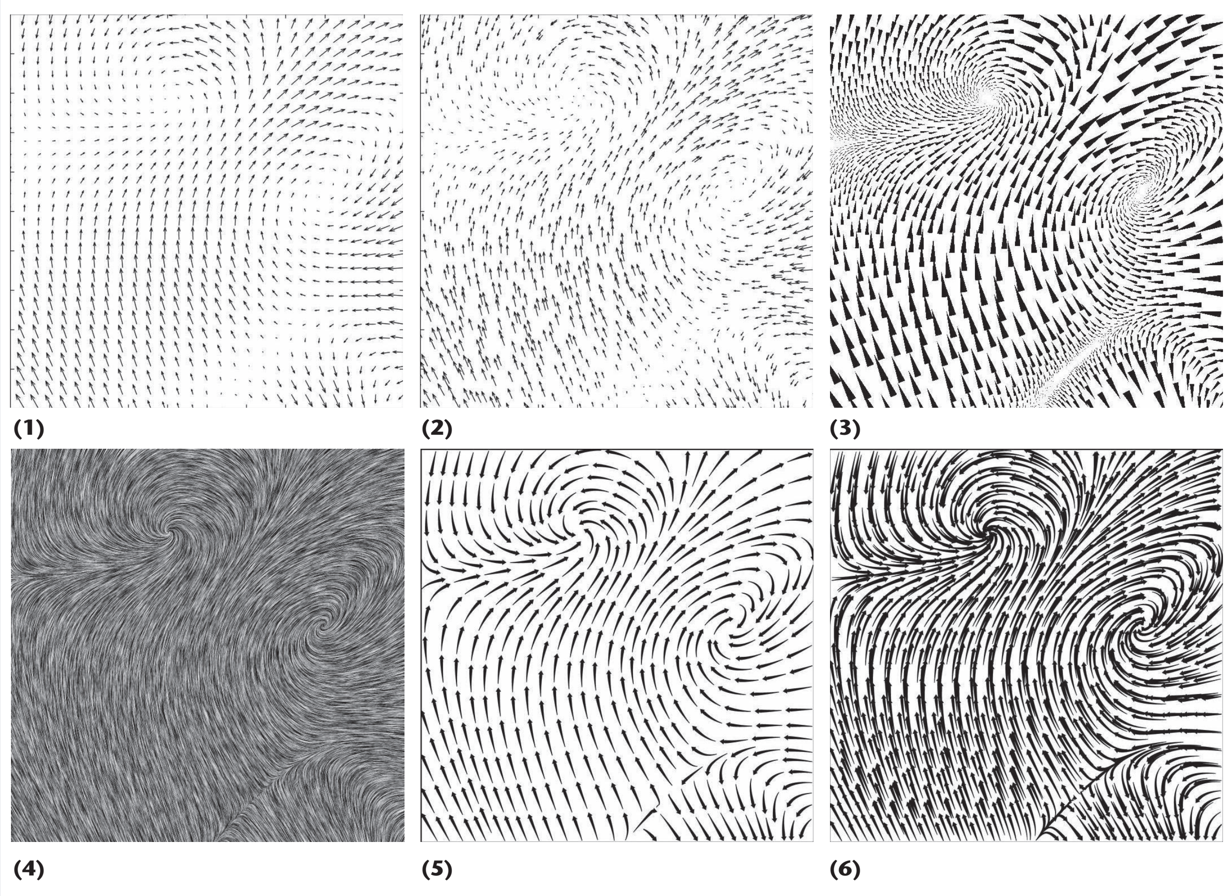

Figure A. Six methods for visualizing the same 2D vector field. They include (1) icons on a regular grid; (2) icons on a jittered grid; (3) layering method inspired by oil painting; (4) line-integral convolution; (5) image-guided streamlines; and (6) streamlines seeded on a regular grid. (Figure courtesy of David Laidlaw, Brown University.)

inevitably to advances in other disciplines. Just as knowledge of mathematics and statistics has become indispensable in subjects as diverse as the traditional sciences, economics, security, medicine, sociology, and public policy, so too is visualization becoming indispensable in enabling researchers in other fields. Like statistics, visualization is concerned with analyzing and interpreting information, both quantitatively and qualitatively, and with presenting data in a way that most clearly conveys their salient features. Both fields develop, understand, and abstract

Fusion Plasma Physics

A tokamak is a doughnut-shaped chamber used in fusion research. During tokamak experimental operation, when a plasma is heated and confined magnetically, events occasionally occur that rapidly terminate the plasma discharge. For future experiments, such as the International Thermonuclear Experimental Reactor (ITER), the stored energy will be approximately 100 times greater than in present-day devices. Thus, these disruptions[1] have the potential to severely damage the material wall, especially if the heat flux is highly localized. Figure B shows the results of a simulation of a particular disruption in the DIII-D tokamak. The figure presents visualizations of the temperature isosurfaces, the magnetic field lines, and contours of the heat flux on the wall, which show toroidal and poloidal localization as a result of the magnetic field lines' topology. The localization results from the plasma instability compressing the flux surface in the plasma's core and then transporting the resulting "hot spots" to the wall. Visualizing the data in 3D shows where the plasma wants to preferentially bulge, enabling the development of improved disruption mitigation techniques.

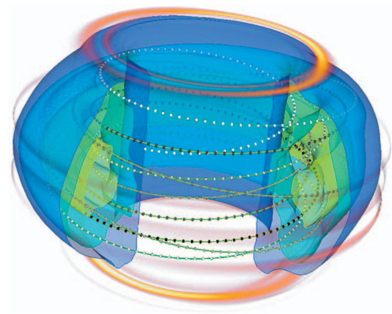

Figure B. Tokamak disruption simulation. The image presents visualizations of the temperature isosurfaces, the magnetic field lines, and contours of the heat flux on the wall.

Reference

1. S.E. Kruger, D.D. Schnack, and C.R. Sovinec, "Dynamics of the Major Disruption of a DIII-D Plasma," *Physics of Plasmas*, vol. 12, no. 056113, 2005.

data-analytic ideas and package them as techniques, algorithms, and software for various application areas.

However, despite visualization's importance to discovery, security, and competitiveness, support for research and development in this critical, multidisciplinary field has been inadequate. Unless we recommit ourselves to supporting visualization research, development, and technology transfer, we'll see a decline in the progress of discovery in the other important disciplines that depend on visualization. As these disciplines lose their ability to harness and interpret information, the discovery rate itself will decline. In the inevitable chain reaction, the US will lose its competitive edge in business and industry.

State of the Field: Infrastructure

One of the 1987 NSF report's[1] key emphases was the need for a national infrastructure to enable visualization research and application. Many of the report's needs have been satisfied, while others remain a challenge.

Visualization Hardware

Many of the original report's hardware concerns have been allayed over time and by Moore's law. Processors with what used to be considered supercomputer-class power are now available in commodity desktop PCs that cost a few thousand dollars. Graphics performance that used to require special-purpose workstations costing tens or hundreds of thousands of dollars is now available as a commodity graphics card for a few hundred. Fast and cheap hardware aimed at the business and entertainment mass markets allows unprecedented access to computational and graphics power for visualization—a boon for both visualization researchers and end users. The latest generation of programmable graphics pipelines on these cards are flexible enough to spark an explosion of sophisticated processing techniques that are now feasible in real time, benefiting visualization users who need to explore large data sets.

Developers have also made great advances in visualization-specific hardware. Volume rendering is extremely computationally intensive. A great deal of academic and industrial research in software volume-rendering algorithms helped the field mature and finally made hardware creation feasible. Grant-funded university research that began at the State University of New York (SUNY) led to the development of an actual product through Mitsubishi Electric Research Lab (MERL), culminating in the successful spin-off company TeraRecon (www.terarecon.com).

In contrast, display technology improvements have historically lagged far behind the Moore's law curve. In the past 20 years, cathode-ray-tube (CRT) displays have little more than doubled in physical display size and resolution and have retained the same weight and form factor. In the past, our ability to design user interfaces has been constrained because monitors are relatively heavy and expensive. However, recent breakthroughs in flat-panel and projector technology have broken the CRT's stranglehold. The combination of low cost, high resolution, and freedom from CRTs'

VIRUS STRUCTURE

Determining a virus's 3D structure is the first step toward understanding its virulent function. A new experimental imaging approach uses cryo-electron microscopy (cryo-EM) for elucidating single-particle macromolecular structures (such as viruses) at the highest quasi-atomic resolution, typically around 10 Å. Figures C1 through C3 show visualizations using combined surface and volume rendering of the same half-shell model of the Rice Dwarf Virus (RDV).[1] The different colors show the nucleo-capsid shell from the outside (Figures C1 and C3) and the inside (Figure C2), elucidating the local and global complexity of each structural protein unit's quasi-symmetric packing. You can see each unit, exhibiting a trimeric fold, further visualized from two different views (Figures C4 and C5), as well as with different colors (Figure C6), to show the three different conformations of the monomeric protein chain forming the trimeric structure unit. All of the structure units are automatically segmented from a reconstructed 3D electron-density map. Previously, structural biologists have largely attempted these 3D ultrastructure elucidation steps manually. The quasi-atomic models of these viruses and other macromolecular machines, constructed via this structure-elucidation pipeline, provide microbiologists and drug-discovery scientists with crucial insight into molecular interactions within macro-assemblies. Such complexes' large physical size and complexity, combined with cryo-EM's very low signal-to-noise ratio, still present significant computational and experimental challenges in this research.

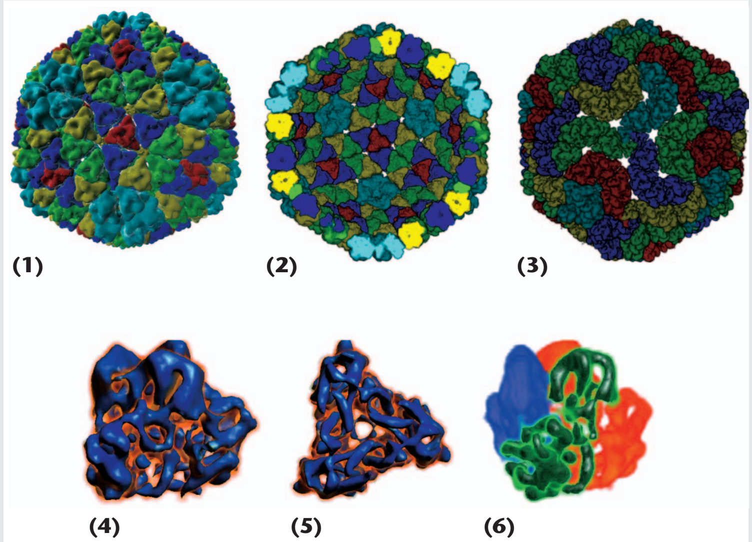

Figure C. The Rice Dwarf Virus. These visualizations combine surface and volume rendering for the same half-shell model of the virus. The different colors show the nucleo-capsid shell from (1 and 2) the outside and (3) the inside. Each structural protein unit is further visualized from (4 and 5) two different views as well as with (6) different colors.

Reference

1. Z. Yu and C. Bajaj, "Automatic Ultra-Structure Segmentation of Reconstructed Cryo-EM Maps of Icosahedral Viruses," *IEEE Trans. Image Processing: Special Issue on Molecular and Cellular Bioimaging,* vol. 14, no. 9, 2005, pp. 1324–1337.

weight and bulk constraints have led to an explosion of computer-driven displays in new contexts that go beyond simply replacing a user's CRT with a sleek flat-panel display that has a smaller footprint.

The primary limitation in interactive visualization interfaces is the number of available pixels; we're pixel-bound, not CPU-bound or even render-bound, and currently face a shortage. High-resolution displays will let us investigate new and exciting areas of the interface design space as displays approach paper-like resolution. We can create wall-sized displays with a resolution of dozens or even hundreds of megapixels by tiling projectors' output. Although active surfaces will still be relatively expensive in the near term, a longer-term vision includes gigapixel displays that will eventually be as cheap, lightweight, and ubiquitous as wallpaper. Physically large displays that encompass an observer's entire viewing field enable applications that use peripheral vision as well as the foveal vision we currently use with medium-sized desktop displays. Small-gadget displays will have the one-megapixel resolution that we typically associate with desktop displays. Small handhelds are ubiquitous and portable, and, when networked, can be used as control panels for a shared large display.

Networking

The visualization field has benefited greatly from the Internet's expansion and commoditization. However, as data sets continue to grow, it's still difficult—and sometimes prohibitive—to move large-scale data sets across even the fastest networks to visualize data locally. As such, we need advances both in networking and remote and collaborative visualization algorithms, such as view-dependent algorithms, image-based rendering, multiresolution techniques, importance-based

The Visualization Toolkit

In 1993, three visualization researchers from General Electric's corporate R&D center began to develop an open-source visualization system, known as the Visualization Toolkit (VTK). VTK was initially envisioned as a teaching and research collaboration tool, hence its release under an open-source license. The software gained rapid acceptance, in part due to the sophistication of its object-oriented design and software process, but also because of the user community that formed around it. VTK is now used worldwide and has helped spawn several small companies and derivative products. Kitware (www.kitware.com), for example, was formed in 1998 to support VTK, subsequently creating products based on the toolkit, including the open-source ParaView parallel visualization system and the proprietary VolView volume-rendering application.[1] VTK continues to evolve with contributions from researchers in academia, US national laboratories, and businesses, and is used in dozens of commercial software applications.

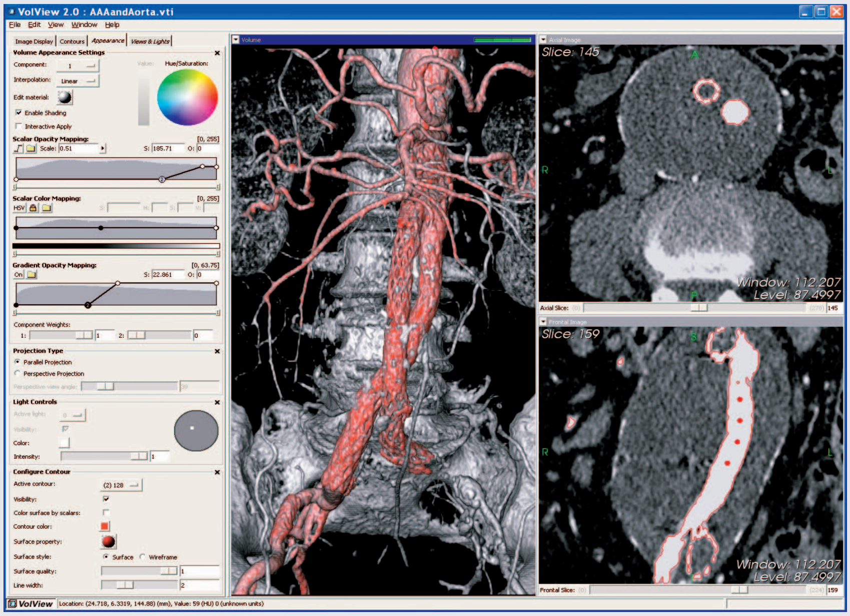

Figure D. The VolView visualization software toolset. This example shows VolView being used to visualize the spinal column, blood vessels, and tissue slides. (Figure courtesy of Kitware.)

methods, and adaptive resource-aware algorithms.

Visualization Software

Dataflow toolkits, introduced more than 15 years ago, heralded the first generation of general-purpose visualization software. Various application-specific and open-source packages have appeared since, and many are in widespread use today. Several are delineated in the new VRC report.[2]

A trade-off exists between quickly creating a one-off prototype that suffices for a research paper but is too brittle for general use and devoting the time to create releasable code at the expense of progressing on the next research project. One benefit for researchers of releasing code to a user community is that real-world use typically spawns new research challenges strongly tied to real problems. Such ties are extremely important for a maturing visualization field. With open-source models, the user community itself sometimes takes over some or all of the support burden. Releasing software doesn't have to be a gargantuan task; often, people who find that a particular piece of research software closely matches their needs are happy to use software that is less polished than a commercial product.

The Visualization Toolkit (VTK; www.vtk.org) system[4] began as an open-source initiative within General Electric's Global Research division, and has rapidly moved into mainstream use in universities, national laboratories, and industrial research labs worldwide (see the "Visualization Toolkit" sidebar for examples). It continues to accelerate development by providing reusable software, relieving programmers from reinventing necessary infrastructure. The spin-off company Kitware (www.kitware.com) is built around an open-source business model, in which customers can pay for support and customization while the community continues to develop the free codebase. Similarly, the Insight Segmentation and Registration Toolkit (ITK; www.itk.org) was an NIH open-source software initiative intended to support a worldwide community in image processing and data analysis. It can interface openly with visualization platforms such as VTK and SCIRun, the University of Utah's visualization system, which has also moved to open-source infrastructure software to ease its integration into public and private research.

Funding

We note with particular concern the recent article by ACM President David Patterson[5] that documents the decline in both industrial and military funding for basic research. Moreover, an increasingly low percentage of NSF pro-

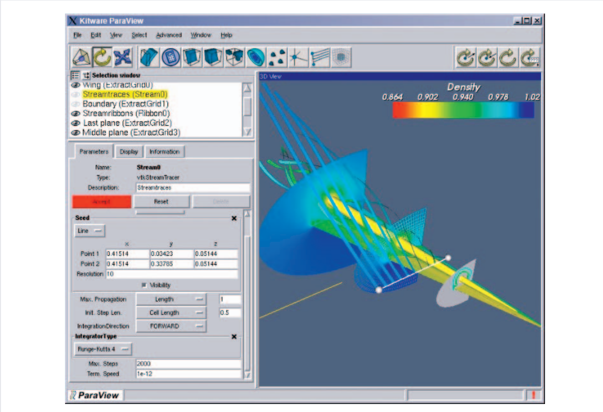

FIgure E. The ParaView toolset. Here, ParaView is being used to visually analyze a computational fluid dynamics data set. (Figure courtesy of Kitware.)

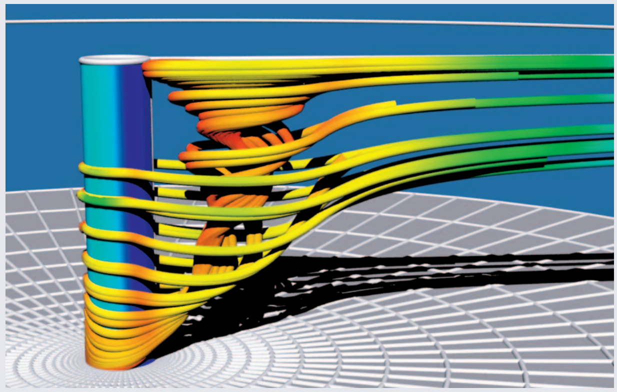

Figure F. The Visualization Toolkit (VTK). This image shows the flow of liquid oxygen across a flat plate with a cylindrical post perpendicular to the flow. (Figure courtesy of Kitware.)

Figure D demonstrates volume rendering using VolView, whereas Figure E shows a computational fluid dynamics visualization using ParaView. Figure F uses the LOx Post data set, simulating the flow of liquid oxygen across a flat plate with a cylindrical post perpendicular to the flow. This model studies the flow in a rocket engine, where the post promotes the mixing of the liquid oxygen.[2]

References

1. W.J. Schroeder, K. Martin, and B. Lorensen, *The Visualization Toolkit: An Object Oriented Approach to Computer Graphics*, 3rd ed., Kitware, 2004.
2. S.E. Rogers, D. Kwak, and U.K. Kaul, "A Numerical Study of Three-Dimensional Incompressible Flow around Multiple Posts," *Proc. AIAA Aerospace Sciences Conf.*, AIAA paper 86-0353, 1986, pp. 1–12.

posals are getting funding, and there appears to be a bias in proposal reviewing toward low-risk incremental work rather than attacking grand challenges.

One important exception to this downturn is the new US National Visualization and Analytics Center (NVAC) initiative,[6] which will spend several hundred million dollars in the next five years on visual analytics. Much of this funding will be focused on the national security domain. Although considerable fundamental research will arise from this short-term program, NIH and NSF must ensure that the health and science domains are sufficiently funded, and that long-term research continues to put the field on a solid scientific foundation.

Recommendations

We explored these infrastructure and funding findings, as well as other findings, with help from area experts in two separate workshops held in Bethesda, Maryland, in September 2004 and Salt Lake City, Utah, in May 2005. Based on our panelists' deliberations, we make the following recommendations.

Principal Agency Leadership Recommendation

NSF and NIH must make coordinated investments in visualization to address the 21st century's most important problems, which are predominantly collaborative and cross disciplines, agencies, and sectors. Both NSF and NIH can and should provide leadership to other federal funding partners and to the research communities they support. To achieve this, they should modify their programmatic practices to better engage visualization capabilities across disciplines important to scientific and social progress, and encourage and reward interdisciplinary research, open practices, and reproducibility in technical and scientific developments. Such agency leadership is critical if we're to meet US needs in critical areas and maintain competitiveness worldwide.

Short-Term Policy Recommendation

NIH and NSF can immediately implement funding policy changes to encourage evaluation of visualization and collaboration between visualization and other fields, without requiring new funding initiatives. Likewise, journals and conferences should update their publication review criteria to reflect how important evaluation is in determining techniques' success and characterizing their suitability. Methodological rigor should be expected when user studies are proposed or reviewed, and as necessary, visualization researchers should collaborate with those trained in fields such as human–

computer interaction, psychology, and statistics. Peer review of proposals and publications should also reward visualization driven by real-world data and tasks, in close collaboration with target users. In other fields, review protocols should encourage domain scientists who create partnerships with visualization researchers, just as engagement with statisticians is considered normal practice in areas such as biomedical research.

Midterm Direction Recommendation

NIH and NSF should establish pilot programs to combine efforts and create collaborative development between visualization and other research domains. Funding for such programs should contribute proportionately to both the visualization research and the domain specialty. This should help improve the penetration of emerging technologies into new domains, increasing their facility to move data and share results through visualization. Awards in this area should be dedicated to open access of source code, availability of research data to the worldwide community, and reproducibility of the technical and scientific developments.

For the long term, we recommend a coordinated and sustained national investment in a spectrum of centralized and distributed research programs to promote foundational, transitional, and applied visualization research in support of science, medicine, business, and other socially important concerns. This investment is critical if the US wants to remain competitive in a global research and development community that has increasing resources. In addition to funding transitional research, such programs should emphasize foundational research and integration of methodologies from other fields, and collaboration with domain specialists who provide driving problems in areas of national concern. A long-term funding commitment is required to create and maintain curated data collections and open-source software to promote open science. Characterizing how and why visualizations work, systematically exploring visual representations' design space, developing new interaction approaches, and exploiting the possibilities of novel display hardware will be particularly important areas of emphasis.

We recommend investment in a spectrum of centralized and distributed research programs.

Acknowledgments

We're indebted to the expert and visionary VRC workshop panelists for sharing their considerable talents and insights, and to NIH and NSF for their sponsorship. In particular, we thank panelists Maneesh Agrawala, Liz Bullitt, Steve Cutchin, David Ebert, Thomas Ertl, Steve Feiner, Felice Frankel, Bob Galloway, Alyssa Goodman, Mike Halle, Pat Hanrahan, Chuck Hansen, Helwig Hauser, Karl Heinz Höhne, Ken Joy, Arie Kaufman, Daniel Keim, David Laidlaw, Ming Lin, Leslie Loew, Bill Lorensen, Wayne Loschen, Patrick Lynch, Kwan Liu Ma, Alan MacEachren, Chris North, Art Olson, Catherine Plaisant, Jerry Prince, Will Schroeder, Jack Snoeyink, John Stasko, Barbara Tversky, Matthew O. Ward, Colin Ware, Ross Whitaker, Turner Whitted, and Jarke van Wijk.

References

1. B.H. McCormick, T.A. DeFanti, and M.D. Brown, *Visualization in Scientific Computing*, Nat'l Science Foundation, 1987.
2. C. Johnson et al., *NIH-NSF Visualization Research Challenges Report*, IEEE Press, 2006.
3. P. Lyman and H.R. Varian, *How Much Information*, 2003; www.sims.berkeley.edu/how-much-info-2003.
4. W. Schroeder, K. Martin, and B. Lorensen, *The Visualization Toolkit: An Object-Oriented Approach to Computer Graphics*, 3rd ed., Kitware, 2003.
5. D.A. Patterson. "The State of Funding for New Computer Science Initiatives in Computer Science and Engineering," *Comm. ACM*, vol. 48, no. 4, 2005, p. 21.
6. J.J. Thomas and K.A. Cook, eds., *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, Nat'l Visualization and Analytics Center, 2005.

Robert Moorhead is the Billie J. Ball Professor of Electrical and Computer Engineering at Mississippi State University (MSU) and the director of the Visualization, Analysis, and Imaging Lab. He is also the associate director for research for the MSU GeoResources Institute. His research interests are focused around visualization of enormous data sets, remote visualization, and the effectiveness of visualization techniques. Moorhead has a PhD in electrical and computer engineering from North Carolina State University. Contact him at rjm@hpc.msstate.edu.

Chris Johnson directs the Scientific Computing and Imaging (SCI) Institute at the University of Utah where he is a Distinguished Professor of Computer Science and holds faculty appointments in the Departments of Physics and Bioengineering. His research interests are in the areas of scientific computing and scientific visualization. Johnson has a PhD in computer science from the

University of Utah. He founded the SCI Research Group in 1992, which has since grown to become the SCI Institute, employing more than 100 faculty, staff and students. Contact him at crj@sci.utah.edu.

Tamara Munzner is an assistant professor of computer science at the University of British Columbia. Her current research interests are information visualization, graph drawing, and dimensionality reduction. Munzner has a PhD in computer science from Stanford University. She was the IEEE Symposium on Information Visualization (InfoVis) Program/Papers cochair in 2003 and 2004. Contact her at tmm@cs.ubc.ca.

Hanspeter Pfister is associate director and senior research scientist at Mitsubishi Electric Research Laboratories in Cambridge, MA. He is the chief architect of VolumePro, Mitsubishi Electric's real-time volume rendering hardware for PCs. His research interests include computer graphics, scientific visualization, and graphics architectures. Pfister has a PhD in Computer Science from the State University of New York, Stony Brook. His work spans a range of topics, including point-based graphics, appearance modeling and acquisition, computational photography, 3D television, and face modeling. Contact him at pfister@merl.com.

Penny Rheingans is an associate professor of computer science at the University of Maryland Baltimore County. Her current research interests include uncertainty in visualization, volume visualization, information visualization, perceptual and illustration issues in graphics and visualization, dynamic and interactive representations and interfaces, and the experimental validation of visualization techniques. Rheingans has a PhD in computer science from the University of North Carolina, Chapel Hill. Contact her at rheingan@cs.umbc.edu.

Terry S. Yoo is head of the Program for 3D Informatics in the Office of High Performance Computing and Communications, National Library of Medicine, NIH. His research explores the processing and visualization of 3D medical data, interactive 3D graphics, and computational geometry. Yoo has a PhD in computer science from the University of North Carolina, Chapel Hill. He is the project officer who conceived and managed the development of ITK, the Insight Toolkit, under the Visible Human Project. Contact him at yoo@nlm.nih.gov.