NIH/NSF Visualization Research Challenges Workshops: Position Statement

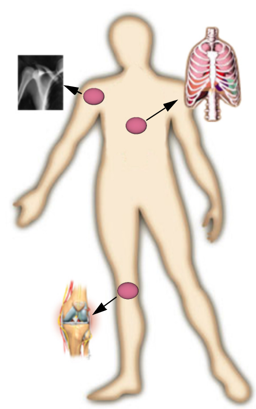
Large Data, Visual Databases, and Software Quality: Research Directions for the Visualization Community

William J. Schroeder
Kitware, Inc.

Our perspective comes from our efforts developing the open-source VTK system, and working closely with researchers at the US National Labs and medical research institutions. We see three important aims for future visualization research: treating large data, cultivating methods for 3D interaction in the form of visual databases, and insuring the quality of the resulting complex software systems.

Large Data. A primary objective of visualization technology is to present large, complex data in perceptual forms that efficiently convey meaningful information. Further, interactive presentation plays a major role in this process. As data sizes increases, whether to improvements in measurement technologies or to increases in computational resolution, visualization systems are often incapable of processing large data, or cannot present the information interactively. Thus visualization systems are at risk of failing to meet one of the primary objectives of the field.

A wide range of approaches to this problem are necessary. Data structures and compression technologies, whether on disk or in memory, are necessary to store and retrieve information efficiently. Multi-resolution algorithms, capable of processing data at multiple levels of detail and in response to spatial/termporal queries, are required. Perceptual studies are needed to determine how to trade off detail versus interactivity. Software and hardware architectures capable of processing data in parallel, or in streaming data in smaller pieces, are also part of the solution.



Visual Databases. Data representational schemes (e.g., data structures or databases) have historically been represented with logical, structured forms. These organizational forms often ignore the inherent spatial/temporal structures that provide context for the data, and that are readily comprehended by the visual system of the human observer. For example, a series of photographs taken on a hiking trip have much more meaning when shown in the context of a map in which the images were acquired. Maintenance records for a complex electro-mechanical system are naturally queried by picking components from a 3D graphics model of the system. Electronic medical records can be represented in visual form—access to patient data is facilitated by selection of data portals placed on a patient-specific map of the body. Further, information can be organized along space-time paths (e.g., the sequence at which photographs were taken) to give additional meaning to data. The point is that the human perceptual system performs natural integration functions that can be leveraged by organizing data into visual representation. These representations also provide natural access to underlying data through a visual interface. A future research direction should include directly incorporating visualization technologies into database design. We believe that this is the natural progression of display technology, much like the evolution of 2D GUI's into the pervasive forms they are today.

Quality: The implementation of visualization technology is by way of large software systems. As these systems become more complex, it is essential to assure the quality of this software since visualization plays a central role in reviewing data. Quality assurance tasks include validating the results of algorithms, and verifying that these results do not change as the software is modified over time. Too often researchers judge the correctness of a visualization technology based on how good the results appear, without deeper investigation into the details of the solution, or relating what is seen to the physics behind the data. Establishing data repositories, with validated results, is vital to researchers who wish to compare and contrast new algorithms. Mandating that software and data utilized in the performance of research are made available in open-source form (especially efforts supported with public funding) is another approach to insure the transparency that the scientific method demands. Investment in software testing and the software process is also important to assure the dependability of complex software systems.