Information Visualization

Nahum Gershon
The Mitre Corporation

Stephen G. Eick
Bell Labs/Lucent Technologies

Visualization links the two most powerful information processing systems known—the human mind and the modern computer. A process, it transforms data, information, and knowledge into a visual form exploiting people's natural strengths in rapid visual pattern recognition. Effective visual interfaces enable us to observe, manipulate, search, navigate, explore, filter, discover, understand, and interact with large volumes of data far more rapidly and far more effectively to discover hidden patterns. In our increasingly information-rich society visualization research and development has fundamentally changed the way we present and understand large complex data sets. The impact of visualization has been widespread and fundamental, leading to new insights and more efficient decision making.

Much previous research in visualization arose from the scientific community's efforts to cope with the huge volumes of scientific data collected by scientific instruments or generated by massive supercomputer simulations.[1] Recently, a new trend has emerged. The explosive growth of the Internet, the overall computerization of the business and defense sectors, and the deployment of data warehouses have created a widespread need—and an emerging appreciation—that visualization techniques are an essential tool for the broader business and technical communities.

Defining information visualization

A new research and development focus has emerged within the visualization community[2-5] to address some of the fundamental problems associated with the new classes of data and their related analysis tasks. This research and development focus, called information visualization,[2] combines aspects of scientific visualization, human-computer interfaces, data mining, imaging, and graphics. In contrast to most scientific visualization, which focuses on data,[1] information visualization focuses on information, which is often abstract. Information in many cases does not automatically map to the physical world (for example, geographical space). This fundamental difference means that many interesting classes of information have no natural and obvious physical representation. A key research problem then is to discover new visual metaphors for representing information and to understand what analysis tasks they support.

Information can come in huge quantities and in fast streams (an information avalanche). Perhaps the biggest information space, the World Wide Web, contains millions of pages. Information visualization must enable users (such as those in the commercial and the defense sectors) to get the information they need, make sense of it, and reach decisions in a relatively short time.

Another key theme for information visualization involves ease of use. In contrast to scientific visualization, which generally serves highly trained scientists, interfaces created for manipulating information might be broadly deployed among a diverse and potentially nontechnical community (see Table 1). The demand for good, effective visualization of information embraces all walks of life and interests. This user community is diverse, with different levels of education, backgrounds, capabilities, and needs. We must enable this diverse group to use visual representations tailored to their specific needs and the problem at hand.

This special issue on information visualization presents the state of the art and trends of this important discipline. It covers two themes: selected state-of-the-

Table 1. Information visualization compared with scientific visualization.

	Audience	Task	Input	Input Quantity
Scientific Visualization	Specialized, highly technical	Deep understanding of scientific phenomena	Physical data, measurements, simulation output	Small to massive
Information Visualization	Diverse, widespread, less technical	Searching, discovering relationships, including action (fast, many times!)	Relationships, nonphysical data, information	Small to massive

A schematic illustration of the visualization process. (Illustration by Elaine Mullen, courtesy of the Mitre Corporation.)

art developments and applications of information visualization to real-world problems.

Some state-of-the-art developments

Over the years the visualization community has developed a diverse array of visualization techniques suitable for specific data and information types. Real-life situations might require visualizations of information and data from diverse sources, while the application might require using different visualization tools. The lack of a common set of operations and ways to integrate information across multiple applications is a long-standing problem in research and applications.

"Information Appliances and Tools in Visage" by Kolojejchick, Roth, and Lucas describes new ways to create a coordinated suite of basic tools and specialized information appliances. This workspace supports specific analysis and reporting tasks for the same data and purpose, and exploratory data analyses done by analysts with varying degrees of skill.

Distortion techniques

Many types of information involve relationships. One common way to visualize structured relationships uses a graph with nodes representing the entities and the links representing the relationships between the entities. Graphs work well for small information sets (tens to hundreds of nodes and links) but are easily overwhelmed and become visually confusing for larger sets. One promising approach for increasing graph information density uses distortion lenses to reveal the detail and the general context. In this issue, "Extending Distortion Viewing from 2D to 3D" by Carpendale, Cowperthwaite, and Fracchia surveys the field and explains their extension of distortion techniques from two to three dimensions.

The World Wide Web

The World Wide Web helps the visualization community provide visualizations to a large audience and has already changed the way visualizations are delivered to users. No longer is the visualization a canned product, cast in stone by its provider. The data could be sent over the network, for example, and the visualization performed on the user (client) side. Not only could this save the time taken to import the visualization product over the network, it also lets users tailor the displayed visualization to suit their specific problems, needs, or capabilities. This can be done by using languages and standards, like Java and VRML (Virtual Reality Modeling Language).

Another advantage in WWW representation is the use of *hyperlinks*, visually connecting points in the visual representation to other objects. "Web-Based Information Visualization" by Rohrer and Swing addresses some of these issues and gives some examples of new and potential applications.

Real-world problems

The importance of information visualization far exceeds academic interest. Its main appeal is its potential to solve real-world problems. Excitingly, this relatively young discipline has already penetrated the commercial market. We dedicated part of this special issue to pointing out where information visualization is being used commercially and where it can go from here.

William Wright (Visible Decisions) describes how information visualization solves problems in the financial market. Ramana Rao (InXight) describes how the ideas developed at Xerox PARC (such as Perspective Wall, Cone Tree, and Wide Widgets) are being used to forge the next-generation user interface, which will be more visual than current ones. Dan Fyock (Lucent Technologies' Visual Insights) outlines using information visualization methods to solve one of the most important software problems ever—the Year 2000 software conversion. And finally, Barry Becker (SGI) describes efforts to use information visualization to represent knowledge extracted from large databases using data mining, thus providing decision support in commercial and other environments.

Applications are not limited to the commercial environment. One important problem, text visualization, arises from the sheer abundance of large corpora of text. This requires users to read many documents in order to understand them and to make sense of the information contained in them. Representing the information in a visual form could allow the user to browse through this ocean of information and to find interesting pieces of text. "Two-Handed Volumetric Document Corpus Management" by Ebert et al. addresses 3D representation of the contents of large collections of documents.

Information visualization is also used in other areas, such as leisure and sports (see "Tennis Viewer: A Browser for Competition Trees," by Liqun Jin and David C. Banks in this issue).

Challenges of information visualization

As a discipline, information visualization is still emerging, tracking the revolution in networking and computing. An emerging discipline progresses through four stages. It starts as a craft, practiced by artisans using heuristics. Later, researchers formulate scientific principles and theories to gain insights about the processes. Eventually, engineers refine these principles and insights into production rules. Finally, the technology becomes widely available. For information visualization, however, these stages are happening in parallel—as the articles in this special issue demonstrate.

Besides the problems faced by all emerging technologies, information visualization faces other challenges, as follows.

Understanding the new media

The media of visual computing and display are quite new, and we do not understand their advantages and disadvantages very well. At present, many developers and users relate to the new medium as if it were a replica of paper—to which we have grown accustomed over the past thousand years. These new technologies, however, allow us to do certain things beyond what is possible with paper.[5] It will take us time to develop sufficient understanding of these new media.

Applications to real-world problems

Using information visualization to solve real-world problems is one of the major challenges of the technology. Much of the research and development done in this area is of academic interest rather than for applicability, thus creating a "bag of tricks" (according to Ken Boff in a private communication). An alternative and more demanding approach is to study the problem first and then look for appropriate solutions, such as using information visualization if appropriate. We need to realize that in many real-world applications, visualization is just one component of a complex system rather than a stand-alone entity. We must understand the system and the user's needs in order to create effective visualizations.

3D versus 2D

With the widespread deployment of 3D graphics chip sets, desktop PCs will soon handle much more sophisticated 3D graphics and animations. The challenge is how best to exploit this forthcoming capability. Currently, we do not always understand when 3D is more effective than 2D and vice versa. As better software makes it easier to produce visualizations, it will be important not to use these new capabilities indiscriminately—only when they are appropriate and convey information effectively.

Human-centered visualization

Tailoring visualization systems based on human capabilities of perception and information processing poses another challenge. We need to better understand how human beings interact with information, how we perceive it visually and nonvisually, how the mind works when searching for both known and unknown information, and how the mind solves problems.

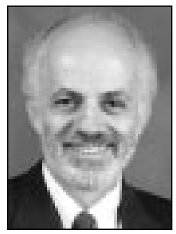
Good human-computer interaction (HCI) is a must, but it is not enough. In designing visualization systems, we also need to better implement what we know about how humans understand and interact with information and the perceptual system. A related challenge involves learning how to create flexible user interfaces, navigation tools, and search methods appropriate for each of the existing types of users, applications, and tasks.

Where do we go from here?

As use of the WWW becomes more sophisticated and more common, it will have a major impact on the way we deliver information visualization. In addition, with the price of hardware and software coming down and their speed rising, more diverse groups of users will use this technology. As visual literacy increases, users will become more comfortable dealing with visuals and will get more information from them. Developers, in turn, will include human considerations when creating increasingly usable visualization systems. We badly need to develop scientific and engineering principles for generating visualizations (given users with diverse needs and capabilities) and a methodology for solving problems with information visualization. ■

References

1. B.H. McCormick, T.A. DeFanti, and M.D. Brown, "Visualization in Scientific Computing," special issue of *Computer Graphics*, Vol. 21, No. 6, Nov. 1987.
2. G.G. Robertson, S.K. Card, and J.D. Mackinlay, "Information Visualization Using 3D Interactive Animation," *Comm. ACM*, Vol. 36, No. 4, 1993, pp. 57-71.
3. *Proc. Information Visualization Symp. 95*, N. Gershon and S.G. Eick, eds., IEEE Computer Society Press, Los Alamitos, Calif., 1995.
4. *Proc. Information Visualization Symp. 96*, S. Card, S.G. Eick, and N. Gershon, eds., IEEE Computer Society Press, Los Alamitos, Calif., 1996.
5. N. Gershon and S.G. Eick, "Visualization's New Tack: Making Sense of Information," *IEEE Spectrum*, Nov. 1995, pp. 38-56.



***Nahum D. Gershon** is a principal scientist at the Mitre Corp. His work concerns data and information visualization, Web browsers, image processing, information and data organization, and analysis of medical, environmental, and other multidimensional data. He pursues research in the use of understanding of the perceptual system in improving the visualization process and dealing with information.*

Gershon co-organized with Steve Eick the Information Visualization Symposia in 1995 and 1996, was a co-chair of the Euro-American Workshop on Visualization of Information and Data 1997 in Paris, and was a co-chair of the IEEE Visualization conferences in 1994 and 1995. He serves as a member of the Advisory Panel of the Earth Observing System Data and Information System (EOSDIS) and the US National Research Council's Codata Committee, and was a member of the Focus Group on Visualization and Presentation of the White House's Globe Program.

***Stephen G. Eick** is the CIO of Bell Laboratories' Visual Insights venture and leads the Interactive Data Visualization Research group at Bell Labs. He and his colleagues have developed a suite of visualizations including tools for displaying geographic and abstract networks, software source code, text corpora, log files, program slices, and relational databases.*

Eick is particularly interested in visualizing databases associated with large software projects and networks, and building high-interaction user interfaces. His educational background includes a BA from Kalamazoo College (1980), an MA from the University of Wisconsin at Madison (1981), and a PhD from the University of Minnesota (1985).

Readers may contact Gershon at The Mitre Corp., 1820 Dolley Madison Blvd., McLean, VA 22102, e-mail gershon@mitre.org and Eick at Bell Laboratories/Lucent Technologies, 1000 E. Warrenville Rd., Room IHC 1G-351, Naperville, IL 60566, e-mail eick@research.bell-labs.com.