The State of Visual Analytics: Views on what visual analytics is and where it is going.

Panel Organizer:

Richard May, Pacific Northwest National Laboratory

Panelists:

Pat Hanrahan, Stanford University
Daniel A. Keim, University of Konstanz
Ben Shneiderman, University of Maryland
Stuart Card, PARC and Stanford University

INTRODUCTION

In the 2005 publication "Illuminating the Path" visual analytics was defined as "the science of analytical reasoning facilitated by interactive visual interfaces." A lot of work has been done in visual analytics over the intervening five years. While visual analytics started in the United States with a focus on security, it is now a worldwide research agenda with a broad range of application domains. This is evidenced by efforts like the European VisMaster program and the upcoming Visual Analytics and Knowledge Discovery (VAKD) workshop, just to name two.

There are still questions concerning where and how visual analytics fits in the large body of research and applications represented by the VisWeek community. This panel will present distinct viewpoints on what visual analytics is and its role in understanding complex information in a complex world. The goal of this panel is to engender discussion from the audience on the emergence and continued advancement of visual analytics and its role relative to fields of related research. Four distinguished panelists will provide their perspective on visual analytics focusing on what it is, what it should be, and thoughts about a development path between these two states. The purpose of the presentations is not to give a critical review of the literature but rather to give a review on the field and to provide a contextual perspective based on the panelists' years of experience and accumulated knowledge.

Each panelist will have at most 15 minutes to present their position in order to establish the context for the audience. The rest of the time will be open for questions. The panel organizer will oversee the panel and ensure there are numerous questions of the panelists.

On August 6, 2010 this world lost Jim Thomas. His influence and impact on our community were clear and profound with conferences, funding programs, research groups, consortia, government agencies, and industry. Through his intellect, energy, and passion Jim was able to create an international movement and shape a new discipline – visual analytics. This discussion on the future of visual analytics is dedicated to Jim and his work. Jim will be greatly missed but his influence and passion will always be felt.

1 POSITION STATEMENTS

1.1 Pat Hanrahan

In the last five years, visual analytics has become a mainstream research topic. Much of this initial research has focused on building demonstration systems that show the potential of coupling visual interfaces to analytical technologies in different domains.

After five years of work, it is worthwhile reflecting on the major accomplishments of the field. What have we learned from these demonstration systems? Has there been progress on the long-term research problems of the field? What progress has been made on the research agenda put forth in "Illuminating the Path"? Additionally, should that research agenda be updated to reflect what has been learned?

The original motivation was to rethink the principles of interactive visualization given the needs of analysis. One insight was to exploit a variety of computational analytics algorithms to create more meaningful visualizations; another goal was to restructure human-computer interaction to directly support complex decision-making tasks. The problems of scale were at the forefront. I'll argue that we have made good progress, but we need to attack the fundamental problems more directly, and to more creatively conceive of ways to combine algorithms and techniques into systems.

1.2 Daniel Keim

Visual Analytics techniques are needed in many applications, including environmental, geo, and social sciences, as well as engineering and economics. Visual analytics techniques are essential to deal with the data sets which are growing fast in size and complexity to gain understanding, to discover patterns, and to optimize and steer complicated processes. Approaches that work either on a purely analytical or on a purely visual level do not help due to the dynamics and complexity of the underlying processes. The multiplicity of research fields and applications, in which visual analysis techniques are urgently needed, demonstrates their scientific relevance and the potential economical utility. A combined automatic and visual analysis is the only chance to capture the complex, changing characteristics of the data and to take suitable measures.

In the future, visual analysis techniques must satisfy a multiplicity of new requirements, which result from the rapid technological development in hardware, software and network infrastructure. In addition to challenges such as high-dimensional data, continuous data streams have additional constraints because evaluation must take place immediately or in the context of a given time frames. Challenges like these pose significant requirements on the data analysis and visualization.

The graphical representation of relevant information from large, complex, and fast growing data streams makes new demands, particularly on the scalability of the techniques. Scalable visual analytics systems should tightly integrate state-of-the-art automatic

data analysis methods with interactive visualization techniques and be integrated smoothly into custom-designed processes.

One of the current challenges in the field of Visual Analytics is that the automatic data analysis community (machine learning, data analysis, statistics, etc.) is not as well represented in the visual analytics research community as is necessary to quickly advance the field. While fully automatic techniques only work if the problem is clearly specified, they are key to the success of the field. The visual analytics research community should ensure that more researchers from the above domains join our efforts to develop visual analytics systems. While a number of successful applications of visual analytics have been developed over the last five years, the development of tightly integrated data analysis and visualization methods is still in the beginning and more research is needed to make progress in this respect.

Due to the immediate demand for visual analytics solutions, visual analytics approaches are being developed within application areas, sometimes without involvement of our visual analytics research community. In such cases, visual analytics is embraced and considered part of the application area, as can be seen by recent visual analytics workshops in the geo and bio domains. It is an open question what impact this development will have and whether visual analytics research will ultimately be embraced by the specific application domains.

As part of the European VisMaster project, a visual analytics roadmap is being developed, describing the state-of-the-art of the field and outlining current research questions and future challenges. The broad European perspective of visual analytics as represented by the roadmap will be covered in the panel.

1.3 Ben Shneiderman

Existing process models for visual analytics can be simple four stage processes, such as Gather Information, Re-represent, Develop Insight, Produce Results. They can also be more elaborate 16-stage sense-making models such as illustrated in the 2005 Book "Illuminating the Path" edited by Jim Thomas and Kristin Cook. These generic process models are a fine start, but different data types (e.g., multi-dimensional, time series, hierarchical, network, textual) require more tailored process models. Further challenges come from dealing with varying information (e.g., real, ordinal, nominal) and domain specific needs (e.g., genomic, financial, social networks). This rich problem space means that careful definitions of inputs, activities, and outcomes will accelerate the development of useful process models, which then can be embedded in generic and domain-specific software tools.

Our experience with multiple tool development projects has given us some experience in developing process models and support software. Overviews of user efforts give a context for work, while showing what has been accomplished and what has not. Then history keeping, undo features, macro making, insight recording, collaboration support, and report export services are part of what is needed in every tool. This panel presentation will concentrate on process models for social network analysis, illustrating our strategies and suggesting new challenges.

1.4 Stuart Card

Visual analytics is an attempt to go beyond the use of Interactive Visualization for gaining insight into data by adding Computational Analysis + a methodology of Analytical Reasoning. But whereas it was possible to layout a more or less coherent framework for codifying visualization as a subdiscipline, this has not yet been possible with visual analytics. The computational analysis part could be anything and an analytical reasoning methodology has only weakly been developed for tying the visual and analytic parts together. There's not enough structure for a coherent subdiscipline. The field may be stuck in some ruts. Many systems have just evolved into groups of different views on the screen + brushing.

I will argue that we should think some more about what this field could be. We might start with defining on the analytic side a more restricted set of analytical technologies to consider combining with visualization. My candidates would be machine learning, network theory, and very large databases. These are techniques that have their own theories to contribute and which concern higher-level abstractions, statistical reasoning, and scale—all useful on the visual side. Then we really do need a methodology of reasoning for individuals and groups that becomes the heart of what we identify with Visual Analytics and coherently ties it together. Finally, we need to understand what we can learn from neighboring disciplines such as data mining.

2 Biographies

2.1 Pat Hanrahan

Pat Hanrahan is the CANON Professor of Computer Science and Electrical Engineering at Stanford University where he teaches computer graphics. His current research involves visualization, image synthesis, virtual worlds, and graphics systems and architectures. Professor Hanrahan has also worked at Pixar where he developed volume rendering software and was the chief architect of the RenderMan(TM) Interface - a protocol that allows modeling programs to describe scenes to high quality rendering programs. In addition to Pixar, he has founded two companies, Tableau Software and PeakStream, and has served on the technical advisory boards of NVIDIA, Exluna, Neoptica, VSee and Procedural. Professor Hanrahan has received three university teaching awards. He has received two Academy Awards for Science and Technology, the Spirit of America Creativity Award, the SIGGRAPH Computer Graphics Achievement Award, the SIGGRAPH Stephen A. Coons Award, and the IEEE Visualization Career Award. He was recently elected to the National Academy of Engineering and the American Academy of Arts and Sciences.

2.2 Daniel Keim

Daniel Keim is Professor and Head of the Information Visualization and Data Analysis Research Group in the Department of Computer Science at the University of Konstanz, Germany. He has been actively involved in information visualization and data analysis research for more than 15 years and developed a number of novel visual analysis techniques for very large data sets. He has been program co-chair of the IEEE InfoVis conference and IEEE VAST Symposium as well as the SIGKDD conference. He is an associate editor of Palgrave's Information Visualization Journal (since 2001) and the Knowledge and Information Systems Journal (since 2006), and has been an associate editor of the IEEE Transactions on Visualization and Computer Graphics (1999 – 2004) and the IEEE Transactions on Knowledge and Data Engineering (2002 – 2007). He is coordinator of the German Strategic Research Initiative (SPP) on Scalable Visual Analytics and the scientific coordinator of the EU VisMaster Coordination Action on Visual Analytics.

2.3 Ben Shneiderman

Ben Shneiderman is a Professor in the Department of Computer Science at the University of Maryland, Founding Director (1983-2000) of the Human-Computer Interaction Laboratory, and Member of the Institute for Advanced Computer Studies at the University of Maryland at College Park. He was made a Fellow of the ACM in 1997, elected a Fellow of the American Association for the Advancement of Science in 2001, and made a member of the National Academy of Engineering in 2010. He received the ACM CHI (Computer Human Interaction) Lifetime Achievement Award in 2001. Since 1991 his major focus has been information visualization, beginning with his dynamic queries and starfield display research that led to the development of Spotfire. Spotfire grew to 200 employees and during Summer 2007 was bought by TIBCO. Dr.

Shneiderman developed the treemap concept in 1991 which continues to inspire research and commercial implementations. Later information visualization work includes the LifeLines project for exploring a patient history, and its successor project, PatternFinder, which enables search across electronic medical records. Three current projects focus on network visualization.

2.4 Stuart Card

Stuart Card works on the theory and design of human machine systems. Until his retirement from PARC, Card was a Senior Research Fellow and the manager of the User Interface Research group. His study of input devices led to the Fitts's Law characterization of the mouse and was a major factor leading to the mouse's commercial introduction by Xerox. His group has developed theoretical characterizations of human-machine interaction, including the Model Human Processor, the GOMS theory of user interaction, the Windows Working Set Theory, Information Foraging Theory, analyses of knowledge creation, and others. The group's work has resulted in a dozen Xerox products as well as contributing to the founding of three software companies. Card holds more than 40 patents and has published more than 90 papers and three books. He is a Fellow of the ACM, a member of the National Academy of Engineering and the recipient the IEEE Visualization Career Award, the ACH SIGCHI Lifetime Achievement Award, and the 2007 Bower Award and Prize from the Franklin Institute. He is currently a Visiting Scientist at PARC and a Consulting Professor at Stanford University.