Guest Editor’s Introduction: Special Section on the IEEE Symposium on Visual Analytics Science and Technology (VAST)

Jarke J. van Wijk

Visual Analytics is the science of analytical reasoning supported by highly interactive visual interfaces. The standard approach for visualization is to aim at interactive visual systems that enable people to obtain insight in large data sets. This is also a central aspect of visual analytics, but here a broader view is taken. If we want to support massive data sets, other data analysis methodologies, such as statistics and machine learning, have to be integrated in order to reduce the data set size and to enable an optimal division of labor between man and machine. In complex, real-world cases the data sets to be analyzed are often not just plain tables with numbers, but are mixed collections of text, multimedia, and relational data. The knowledge discovery process is not just about obtaining insight, but encompasses a series of stages, from data collection via analysis to presentation. Also, large data analysis problems require teams of people, where each brings in his or her own expertise. As a result, visual analytics is an ambitious endeavor, requiring a variety of disciplines, and with many open ends and unanswered questions. Additionally, visual analytics is the focus of a vibrant and active research community.

The field was founded in the 2000s by Jim Thomas, who led a team of scientists and leaders from academia, industry, and government, resulting in the highly influential report “Illuminating the Path,” coedited with Kristin Cook. Jim Thomas passed away in 2010, and the field will miss his vision and leadership. Fortunately, his initiatives have led to a variety of activities, which steadily increase the size and impact of the field.

One important initiative was the IEEE Visual Analytics Science and Technology (VAST) symposium, begun in 2006. Its success has been exceptional. More and more researchers around the world contribute their high-quality results to the meeting each year. As a result, as of 2010, VAST has been promoted from an IEEE symposium to an IEEE conference. IEEE Transactions on Visualization and Computer Graphics (TVCG) has recognized the importance of visual analytics from the start, and has decided to invite authors of the best conference papers to submit archival quality extended versions of their papers to the journal. In this special section, extended versions of the best papers of IEEE VAST 2009, which took place on 12 and 13 October in Atlantic City, New Jersey, are presented. These papers were selected together with the best paper award committee, and went through the standard reviewing process of TVCG. I am happy with the results, not only because of the high quality of the individual papers, but also because they provide a great overview of the many different aspects of visual analytics.

The first paper is “Iterative Integration of Visual Insights during Scalable Patent Search and Analysis” by Steffen Koch, Harald Bosch, Mark Giereth, and Thomas Ertl. The conference version of this paper won the Best Paper award at IEEE VAST 2009. The work is a great example of how visual analytics can contribute to solving a real-world complex task. The paper describes an integrated system to search and find patents. Searching patents is of great interest for many, not only for companies to manage their intellectual property, but also for experts from the finance sector, scientists, and many more. Also, patent search is a difficult task. The number of patents is enormous, on the order of tens of millions, spread over different countries, using different languages and conventions. The keywords used to describe the patent can be unclear or even purposely obfuscated. Hence, this area presents a strong case for visual analytics. The authors introduce PatViz, a new system for interactive analysis of patent information to leverage iterative query refinement. Users are enabled to define complex queries in an interactive, visual way. Via an abstraction layer these queries are translated into specific requests for different retrieval systems. Results are shown to the users in a variety of ways, enabling them to view relations between patents, as well as geographical and temporal aspects, which in turn can be used to further enhance the visual query.

The second paper is “How Can Visual Analytics Assist Investigative Analysis? Design Implication from an Evaluation” by Youn-ah Kang, Carsten Görg, and John Stasko. Evaluation is a notoriously difficult aspect of visualization in general and of visual analytics as well. For some aspects, such as perceptual issues, simple experiments with a short cycle time can be used, with a clear task and measurement of time taken and errors made. To evaluate how a system enables users to obtain insight is a much more difficult task. This paper presents an in-depth study of how users performed a sensemaking exercise, in which a collection of 50 documents had to be analyzed. Sixteen participants...
performed this task under one of four conditions, ranging from just using paper, using a computer and digital documents, or using a reduced and a full version of Jigsaw, the multiview document analysis system developed by the group of John Stasko at Georgia Tech. The work of the participants was carefully monitored and analyzed in great detail, leading to a variety of observations and conclusions. First of all, the group using Jigsaw performed better than the others, showing that such a carefully designed system can be really helpful. But, this was just one result. The authors deduced from their observations that the study participants used four different analytic strategies, each having implications on the results. Also, the authors present observations on the sensemaking process in general, give design implications for such tools, and share new insights on how to evaluate visual analytics tools. In summary, this paper addresses a core issue of visual analytics research in a very thorough way, and I expect it to become an often cited landmark paper.

The third paper is “Automated Analytical Methods to Support Visual Exploration of High-Dimensional Data” by Andrada Tatu, Georgia Albuquerque, Martin Eisemann, Peter Bak, Holger Theisel, Marcus Magnor, and Daniel Keim. Multivariate data are often analyzed by projecting the items onto lower-dimensional representations. If there are many attributes, there is a large choice of possible projections, and to investigate them all requires much time. The authors explore an intriguing idea: Can we automate viewing visualizations, and let the computer come up with a short-list of interesting views? An important step here is to measure the quality of a visualization. The authors present a variety of quality measures for scatterplots and parallel coordinate plots, for classified as well as unclassified data. As an example, the rotating variance measure distinguishes between projections with long, skinny structures that reveal correlations and amorphous, blob-like images. A number of examples are given that show the potential of this approach.

I hope that this special section gives the readers of TVCG an insight into the current state-of-the-art research in visual analytics and encourages them to contribute to this exciting field as well.

Jarke J. van Wijk
Guest Editor