

Introduction to the Special Issue on Visual Analytics

Daniel A. Keim
Databases, Data Mining and Visualization Group
Department of Computer Science
University of Konstanz
Germany
keim@inf.uni-konstanz.de

Joern Schneidewind
Databases, Data Mining and Visualization Group
Department of Computer Science
University of Konstanz
Germany
jschneidewind@acm.org

The information overload is a well-known phenomenon of the information age, since due to the progress in computer power and storage capacity over the last decades, data is produced at an incredible rate, and our ability to collect and store these data is increasing at a faster rate than our ability to analyze it. But, the analysis of these massive, typically messy and inconsistent, volumes of data is crucial in many application domains. For decision makers, analysts or emergency response teams it is an essential task to rapidly extract relevant information from the flood of data.

Today, a selected number of software tools is employed to help analysts to organize their information, generate overviews and explore the information space in order to extract potentially useful information. Most of these data analysis systems still rely on interaction metaphors developed more than a decade ago and it is questionable whether they are able to meet the demands of the ever-increasing mass of information. In fact, huge investments in time and money are often lost, because we still lack the possibilities to properly interact with the databases.

Visual analytics aims at bridging this gap by employing more intelligent means in the analysis process. The basic idea of visual analytics is to visually represent the information, allowing the human to directly interact with the information, to gain insight, to draw conclusions, and to ultimately make better decisions. The visual representation of the information reduces complex cognitive work needed to perform certain tasks. People may use visual analytics tools and techniques to synthesize information and derive insight from massive, dynamic, and often conflicting data by providing timely, defensible, and understandable assessments. This special issue presents articles that address interesting, important and diverse issues in visual analytics research and practice.

Scope of Visual Analytics

The goal of visual analytics research is to turn the information overload into an opportunity. Decision-makers should be enabled to examine this massive, multi-dimensional, multi-source, time-varying information stream to make effective decisions in time-critical situations. For informed decisions, it is indispensable to include humans in the data analysis process to combine their flexibility, creativity, and background knowledge with the enormous storage capacity and

the computational power of today's computers [1].

The specific advantage of Visual Analytics is that decision makers may focus their full cognitive and perceptual capabilities on the analytical process, while allowing them to apply advanced computational capabilities to augment the exploration process [5].

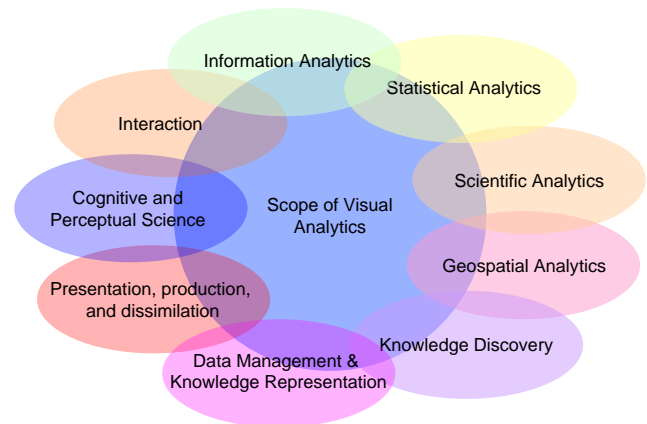


Figure 1: The Scope of Visual Analytics

In general, Visual Analytics can be described as “the science of analytical reasoning facilitated by interactive visual interfaces” [4]. To be more precise, Visual Analytics is an iterative process that involves information gathering, data preprocessing, knowledge representation, interaction and decision making. The ultimate goal is to gain insight in the problem at hand which is described by vast amounts of scientific, forensic or business data from heterogeneous sources. To reach this goal, Visual Analytics combines the strengths of machines with those of humans [1].

On one hand, methods from data mining, statistics, and mathematics are the driving force on the automatic analysis side, while on the other hand human capabilities to perceive, relate, and conclude turn Visual Analytics into a very promising field of research.

The Visual Analytics process, as described in [2], aims at tightly coupling automated analysis methods and interactive visual representations, as described before. Without the support of automated methods, the visual analysis of large data sets will become impossible in the future.

As a consequence, we extended the classical way of visually exploring data sets as defined by the Information Seeking Mantra (“Overview first, Zoom/ Filter, Details on demand”)[3], to the Visual Analytics Mantra [2]:

*“Analyze First -
Show the important -
Zoom, Filter and Analyze Further -
Details on Demand”*

In the age of massive data sets all three steps of the Information Seeking Mantra are difficult to realize. An overview visualization without losing any interesting pattern or subset is difficult to create, since the amount of pixels of modern data display systems do not keep pace with the increasing flood of data. The plenty of information hidden in massive data sets make it very difficult for humans to understand the really interesting or relevant information.

In Visual Analytics it is therefore not sufficient to just retrieve and display the data using a visual metaphor, it is rather necessary to support the analyst by analytically filtering the underlying data by its value of interest, but at the same time providing interaction models which still allow the user to get any detail of the data on demand.

About the articles

This special issue on Visual Analytics attracted a number of high quality submissions from Brazil, France, Germany, Italy, the Netherlands, Switzerland, UK, and the US. Each submission was assigned to at least two domain experts for thorough review and evaluation. We selected five submissions that best represent the current state of the art in the studies and applications of visual analytics.

I. Assent, R. Krieger, E. Mueller, and T. Seidl describe a subspace clustering visualization that allows users to browse the entire subspace clustering, to zoom into individual objects, and to analyze subspace cluster characteristics in-depth. Browsing of the clustering result is possible through a novel distance function that reflects the subspace and the object overlap, respectively. Since the paper shows how visual methods can be tightly coupled with automated data mining techniques it address one of the core issues of visual analytics.

H. Kang, L. Getoor, and L. Singh investigate the use of visual analytics techniques for analyzing dynamic group membership in temporal social networks over time. They present the C-Group tool, that unlike most network visualization tools, which show the group structure within an entire network, or the group membership for a single actor, allows users to focus their analysis on a pair of individuals. The authors show how dynamic patterns can be explored in complex networks, which is a valuable contribution for a number of visual analytics application domains.

The article by D. Yang, Z. Xie, E. A. Rundensteiner, and M. O. Ward presents an innovative visual analytics framework for analysis-guided visual exploration of multivariate data. Their system helps users extract the valuable information (nuggets) hidden in datasets based on their interests. Visualization and interaction techniques are designed to help users observe and organize the extracted nuggets in an intuitive manner and eventually facilitate their sense-making process. Thus the paper provides a promising combination

of automated and interactive techniques for multi-variate data analysis.

The work by T. Schreck, T. Tekusova, D. Fellner, and J. Kohlhammer describes a novel approach for an application of visual analytics techniques in the financial sector, in particular the analysis of financial time-varying indicator data. The system relies on an unsupervised clustering algorithm combined with an appropriately designed movement data visualization technique. Several analytical views on the full market and specific assets are offered for the user to navigate, to explore, and to analyze. The presented tool allows even non-domain experts to quickly get an overview over asset risk-return patterns.

Finally, G. Andrienko, N. Andrienko, and S. Wrobel focus on the use of innovative visual analytics techniques to analyze and visualize geospatial patterns. They explore large volumes of GPS data and present a framework to effectively support human analysts in understanding movement behaviors and mobility patterns. They demonstrate the synergistic use of automated and visual techniques in case studies of two real world datasets. The presented ideas are suited to analyze geo-related information in a number of related visual analytics scenarios.

Conclusion

The articles presented in this special issue represent some of the most important topics for further investigation and exploration, and show that the integration of automated and visual/interactive techniques, as proposed in the context of visual analytics, is of great benefit for many application domains.

Acknowledgments

We are grateful to the reviewers for their timely and constructive reviews. Without their help, this special issue would not be possible. We also thank the authors who responded to the call-for-papers and submitted their articles to this special issue.

REFERENCES

- [1] D. A. Keim. Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 8(1):1–8, January–March 2002.
- [2] D. A. Keim, F. Mansmann, J. Schneidewind, and H. Ziegler. Challenges in visual data analysis. In *IEEE Information Visualization, London, UK*, 2006.
- [3] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *IEEE Visual Languages*, pages 336–343, 1996.
- [4] J. Thomas and K. Cook. *Illuminating the Path: Research and Development Agenda for Visual Analytics*. IEEE-Press, 2005.
- [5] Pak Chung Wong and Jim Thomas. Visual analytics - guest editors’ introduction. *IEEE Transactions on Computer Graphics and Applications*, September/October 2004.