Proceedings
SOFTVIS 2008

4th ACM Symposium on Software Visualization

Herrsching am Ammersee, Germany
September 16 – 17, 2008

Program Chairs

Christopher Hundhausen (Washington State University, USA)
Alexandru Telea (Groningen University, The Netherlands)

General Chair

Rainer Koschke (University of Bremen, Germany)

Proceedings Production Editor

Stephen N. Spencer, University of Washington

Sponsored by ACM SIGGRAPH, ACM SIGSOFT, ACM SIGPLAN and ACM SIGCHI

The Association for Computing Machinery, Inc.
2 Penn Plaza, Suite 701
New York, New York 10121-0701

ACM ISBN: 978-1-60558-112-5
Additional copies may be ordered prepaid from:

ACM Order Department
P.O. Box 11405
Church Street Station
New York, NY 10286-1405

Phone: 1-800-342-6626
(USA and Canada)
+1-212-626-0500
(All other countries)
Fax: +1-212-944-1318
E-mail: acmhelp@acm.org

ACM Order Number: 102082

Printed in the USA

Table of Contents

Table of Contents

Table of Contents

Posters

Preface

Welcome to SOFTVIS 2008, the 4th ACM Symposium on Software Visualization. This symposium focuses specifically on visualization techniques that draw on aspects of software maintenance, program comprehension, reverse engineering, and reengineering. The high-level research question is "How visualization can help programmers understand and analyze large-scale software systems?"

We are very pleased to be part of the "Visual Week" held from September 15th to 19th 2008 in Herrsching am Ammersee, Germany. This event combines three conferences (IEEE Symposium on Visual Languages and Human-Centric Computing, Diagrams, and SOFTVIS), together with two workshops (Layout of Engineering Diagrams and Sketch Tools for Diagramming). "Visual Week" promises to be a veritable feast for the connoisseur of visual computing application, with a program that highlights the multiple uses of visual and interactive techniques in different areas of computing, including software engineering.

This year's edition of SOFTVIS continues the tradition of co-location with the VL/HCC conference started at the previous edition in 2006. We have a joint technical paper session where representative papers from both events will be presented. Also, we share a joint keynote address delivered by John Stasko, and joint coffee breaks with the main conference of the Visual Week and a poster session and reception with the Graduate Consortium event. As with the previous edition, discounts for joint registrations were figured out by all the General Chairs of the Visual Week events. Mark Minas, the General Chair of VL/HCC, kindly helped us with managing the joint registration and logistics for all the involved events.

This year's edition of SOFTVIS is marked by a number of innovations. First, we introduced the short paper category. Papers accepted in the short category were those that the program committee felt contained important contributions for the symposium, but that would require further technical work or maturity to warrant a full paper publication. Short papers were given 4 pages in the proceedings, as opposed to 10 pages for full papers. Second, for the first time in the history of SOFTVIS, we have full-color proceedings, which does proper justice to the highly visual content of the accepted papers.

This year there were 38 technical papers submitted. Each submission was reviewed by at least three members of the international program committee. On the basis of these reviews, the program committee accepted 16 full papers (42 percent acceptance rate) and 8 short papers for inclusion in the technical program and publication in the proceedings. Additionally, we continued the poster track established in previous editions of SOFTVIS. In all 9 poster papers were accepted for publication in the proceedings; these were presented in a special poster session held as part of the opening reception.

We would like to thank the program committee members and the additional reviewers for their diligence and attention to reviewing the submissions. Their input was extremely invaluable in making the exciting program we have this year. We also thank the SOFTVIS steering committee for their help and advice with the organization of the conference. We thank Jochen Quante for his support in assisting us with EasyChair, the paper review system we used. We also wish to acknowledge the excellent support provided by Stephan Diehl with the symposium's web page, and Stephen Spencer from ACM SIGGRAPH for managing the publication of the proceedings.

Last but not least, we greatly thank the General Chair of VL/HCC 2008, Mark Minas, for helping us with the co-location of SOFTVIS and for his help in organizing the local arrangements. We wish all participants an excellent Visual Week!

General Chair
Rainer Koschke

Program Co-Chairs
Alexandru Telea
Chris Hundhausen

Keynote Presentation

Visualization for Information Exploration and Analysis

John Stasko
School of Interactive Computing & GVU Center
Georgia Institute of Technology, Atlanta, GA 30332-0760
stasko@cc.gatech.edu

Abstract

Making sense of data becomes more challenging as the data grows larger and becomes more complex. If a picture truly can be worth a thousand words, then clever visualizations of data should hold promise in helping people with sense-making tasks. I firmly believe that visual representations of data can help people to better explore, analyze, and understand it, thus transforming the data into information. In this talk, I will explain how visualization and visual analytics help people make sense of data and I will provide many such examples. I also will describe my present research into visualization for investigative analysis. This project explores how visual analytics can help investigators examine a large document collection in order to discover embedded stories and narratives scattered across the documents in the collection.

1. Visualization for Understanding

The amount of electronic data available to people continues to grow and grow. Berkeley's "How Much Information" project estimated that 5 exabytes (10^{18} bytes) of information was produced in 2002 [1]. International Data Corporation estimated the digital universe to be 281 exabytes in 2007 [2]. But taken in its raw form, data in the digital universe is just that: raw data (text and numbers) that may or may not be useful to people. Thus, the goal of researchers such as myself is to help people transform the raw data into *information* that carries meaning, that assists decision-making, and that generates knowledge and wisdom.

Computational techniques and tools already have been very useful in the process of transforming data into information. For instance, spreadsheets provide structure and order on data that help us to examine and think about the data better. Computational algorithms for finding clusters and correlations in data provide us with useful meta-information about the characteristics of the data. Natural language parsing and understanding systems take raw text and extract useful features for further analysis. Data mining algorithms can identify specific data cases with particular characteristics that are difficult for us to determine in other ways.

My research focus, visualization, is another computational technique for transforming data into information. By visually depicting data in meaningful, expressive representations, we can translate it into a form that facilitates human exploration, analysis, and understanding.

The notion of visual imagery helping people to think is not new. Common clichés such as "Seeing is believing," "I see what you're saying," or "A picture's worth a thousand words" address this idea. More scientifically, visualization serves as an external cognition aid [3] and augments our memory as we perform cognitive operations [5].

But with other computational techniques and tools available, such as those mentioned earlier, when is visualization particularly useful? I believe that visual representations of data are particularly germane in situations involving unfamiliar data when a person is performing exploratory data analysis. In such situations, the person may not have a specific analytic goal in mind and may not have particular questions to be answered. Instead, the person is exploring the data, hoping to learn more about it and to discover "interesting" characteristics of the data. Throughout the exploration process, specific questions may arise and ideally the visualization will help to answer them, but the data analysis expedition likely did not initiate with all these questions in mind. Together with colleagues J.-D. Fekete, J. van Wijk, and C. North, I explore this "value of visualization" more fully in a recent book chapter [4].

In the academic community, the growing research areas of information visualization (infovis) and visual analytics address these very issues. Visual analytics augments ongoing infovis research on visual representations (and interactions) of data with studies of data transformations and analytical reasoning [9].

SOFTVIS 2008, Herrsching am Ammersee, Germany, September 16–17, 2008.

2. Investigative Analysis on Text

My recent research has focused on exploration and analysis of unstructured text document collections. Of the 281 exabytes of data estimated to exist by IDC in 2007, over 95% was said to be unstructured data [2]. Of course, much of this may be images, audio and multimedia, but an important percentage is likely text.

Together with my students at Georgia Tech, I have developed the Jigsaw system for assisting investigative analysis on text documents [7,8]. Jigsaw helps analysts with sense-making [6] activities, in particular, weaving together different chunks of information scattered across a variety of documents into a coherent story or narrative. Essentially, different documents will provide individual pieces of evidence that on their own may not be particularly illuminating. But when these individual pieces are considered together, they reveal a broader story and set of connected events. Such a capability could be useful for different types of analysts such as police investigators or news reporters.

Jigsaw provides multiple visualizations of the documents in a collection and the entities within the documents, such as people, places, organizations, phone numbers, etc. The system helps an investigator uncover and explore connections between entities across documents and it points the investigator to the documents to read next in a large collection. As such, it facilitates exploratory data analysis and sense-making activities on text documents.

3. References

[1] http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/. Accessed June 26, 2008.

[2] http://www.emc.com/digital_universe. Accessed June 26, 2008.

[3] Card, S., J. Mackinlay, and B. Shneiderman, *Readings in Information Visualization, Using Visualization to Think*, Morgan Kaufmann, 1999.

[4] Fekete, J.-D., J. van Wijk, J. Stasko, and C. North, "The Value of Information Visualization," in A. Kerren, J. Stasko, J.-D. Fekete, and C. North (Eds.), *Information Visualization – Human-Centered Issues and Perspectives*, Volume 4950 of LNCS State-of-the-Art Survey, Springer, 2008.

[5] Norman, D., "Visual Representations", in *Things that Make us Smart: Defending Human Attributes in the Age of the Machine*, Addison-Wesley, 1994.

[6] Russell, D.M., M.J. Stefik, P. Pirolli, and S.K. Card, "The Cost Structure of Sensemaking," Proceedings of INTERCHI '93, Amsterdam, April 1993, pp. 269-276.

[7] Stasko, J., C. Görg, Z. Liu and K. Singhal, "Jigsaw: Supporting Investigative Analysis through Interactive Visualization", Proceedings of IEEE VAST '07, Sacramento, CA, October 2007, pp. 131-138.

[8] Stasko, J., C. Görg, and Z. Liu, "Jigsaw: Supporting Investigative Analysis through Interactive Visualization", *Information Visualization*, Vol. 7, No. 2, Summer 2008, pp. 118-132.

[9] Thomas, J.J. and K. A. Cook, *Illuminating the Path*. IEEE Computer Society, 2005.