Lecture Notes in Computer Science 1183
Edited by G. Goos, J. Hartmanis and J. van Leeuwen

Advisory Board: W. Brauer D. Gries J. Stoer

Andreas Wierse Georges G. Grinstein
Ulrich Lang (Eds.)

Database Issues for Data Visualization

IEEE Visualization '95 Workshop
Atlanta, Georgia, USA, October 28, 1995
Proceedings

Springer

Series Editors

Gerhard Goos, Karlsruhe University, Germany

Juris Hartmanis, Cornell University, NY, USA

Jan van Leeuwen, Utrecht University, The Netherlands

Volume Editors

Andreas Wierse
Ulrich Lang
Universität Stuttgart, Rechenzentrum
Allmandring 30, D-70550 Stuttgart, Germany
E-mail: (wierse/lang)@rus.uni-stuttgart.de

Georges G. Grinstein
University of Massachusetts Lowell
Institute of Visualization and Perception Research
1 University Avenue, Lowell, MA 01854, USA
E-mail: grinstein@cs.uml.edu

Cataloging-in-Publication data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Database issues for data visualization : proceedings / IEEE Visualization '95 Workshop, Atlanta, Georgia, USA, October 28, 1995. Andreas Wierse ... (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Budapest ; Hong Kong ; London ; Milan ; Paris ; Santa Clara ; Singapore ; Tokyo : Springer, 1996
(Lecture notes in computer science ; Vol. 1183)
ISBN 3-540-62221-7
NE: Wierse, Andreas [Hrsg.]; Visualization Workshop <2, 1995, Atlanta, Ga.>; Institute of Electrical and Electronics Engineers; GT

CR Subject Classification (1991): H.2, H.5, I.6.8, E.1

ISSN 0302-9743
ISBN 3-540-62221-7 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer - Verlag. Violations are liable for prosecution under the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1996
Printed in Germany

Typesetting: Camera-ready by author
SPIN 10550015 06/3142 – 5 4 3 2 1 0 Printed on acid-free paper

Preface

Database Issues for Data Visualization: the word data occurs twice in this title and thus clearly shows the emphasis of these proceedings. Databases and visualizations both deal with data. But each of these fields typically deals with its data independently from the other. Visualization features found in today's database management systems are far from what are provided by visualization systems. And today's visualization systems provide very primitive data access and manipulation tools compared to those that a well developed database management system offers.

It is obvious that both groups will benefit significantly from close cooperation. The ever increasing amount of data stored in databases demands new, comfortable ways of access and manipulation. The amount of data that is produced daily by satellites and sent back to Earth to be stored in huge databases require interactive visual support to provide more efficient data management and other tasks. Visual exploration of data stored in databases offers new ways of accessing and interpreting this data.

There are many other applications providing examples for the need to integrate databases and visualization. The data that is currently dealt with and visualized in supercomputing environments has grown in step with the ever increasing size of the numerical problems that can currently be computed and solved. The results of time-dependent numerical simulations can easily reach gigabytes in size. To allow a useful interpretation of these results it is important that the data can be accessed and managed efficiently, and thus visualized.

On October 28, 1995, the second workshop on Database Issues for Data Visualization was held during the IEEE Visualization '95 conference in Atlanta, Georgia. Two years after the first and very successful workshop (proceedings of which are available in Lecture Notes in Computer Science Volume 871), researchers from the areas of databases and visualization met again to discuss and share their problems, needs, and goals. The second workshop topics included data modeling and access, object modeling, user interface construction, dataflow and program module storage and retrieval, the composition and manipulation of graphical representations, and the integration of knowledge bases and rule-based systems, all in a visualization context.

These proceedings offer a snapshot of current research in this field and together with the proceedings of the first workshop offer a way to see the progress of this emerging area. The proceedings also provide a survey of the problems that must be addressed now and in the future towards the integration of database management systems and data visualization.

With these proceedings, the reader is presented with a treatment of a wide range of issues, top to bottom, of the research areas and problems facing the integration of database and visualization systems. We hope to further stimulate research activity in this field, and look forward to the realization of truly integrated systems that accommodate end-user requirements in terms of models, services, displays, and interaction capabilities. We also hope the reader will find the reports and papers as invigorating as the discussion sessions during the workshop.

October 1996, Stuttgart and Lowell

Andreas Wierse
Georges G. Grinstein
Ulrich Lang

Table of Contents

Workshop Description IX

Workshop Participants XIII

Workshop Subgroup Reports

Scientific Data Modeling 3
William Hibbard, David T. Kao and Andreas Wierse

Supporting Interactive Database Exploration 12
Daniel Keim, John Peter Lee, Bhavani Thuraisingham and Craig Wittenbrink

Visualization Related Metadata 26
Ulrich Lang, Georges Grinstein and R. Daniel Bergeron

Workshop Papers

Views, Visualization and Databases 37
John Peter Lee

Data Mining and Data Visualization 54
Georges Grinstein and Bhavani Thuraisingham

Requirements and Architecture of an Information Visualization Tool 57
Sougata Mukherjea and James D. Foley

Mathematical Structures of Data and Their Implications for Visualization 76
William Hibbard

Semantics and Mathematics of Scientific Data Sampling 86
David T. Kao, R. Daniel Bergeron, Michael J. Cullinane and Ted M. Sparr

Enhancing the Visual Clustering of Query-Dependent Database Visualization Techniques Using Screen-Filling Curves 101
Daniel A. Keim

Realtime Database Support for Environmental Visualization 111
Craig M. Wittenbrink, Eric Rosen, Alex Pang, Suresh K. Lodha and Patrick Mantey

Metalevel Database Programming and Visualization with POETView 131
Matthias Neugebauer

Large Scale Data Analysis Using AVS5 and POSTGRES 145
Ray E. Flanery Jr. and June M. Donato

FAN - An Array-Oriented Query Language 155
Harvey Davies

LadMan - A Large Data Management System 171
Walter Schmeing

The Tioga-2 Database Visualization Environment 181
Alexander Aiken, Jolly Chen, Mark Lin, Mybrid Spalding, Michael Stonebraker and Allison Woodruff

Collaborative Visualization Based on Distributed Data Objects 208
Andreas Wierse

Workshop Description

The second workshop on database issues for data visualization took place in Atlanta in 1995. Two years after the first workshop (1993 in San Jose) it gave the opportunity to analyze and discuss the current situation in this field and the progress achieved in the meantime.

The basic problem is still that current data visualization environments offer minimal support for database management and data interrogation methods. Although in the research area a lot of work has been done, the results of this work are not yet available for a broader range of users. The visual presentation of data still predominates over all other system functions. While the graphical modeling and rendering issues are addressed sufficiently by these systems, the underlying problems of data modeling, data access, and dialog design are not yet handled as efficiently.

In contrast to the first workshop,however,the submissions show a much larger share of descriptions of existing systems that try to address some of these problems. These systems have also reached a very high level of functionality, giving hope that these achievements will soon be found in commercially available systems.

This development confirms our opinion that the application of database management system (DBMS) technology to data visualization for scientific discovery brings powerful data access methods in addition to the standardization of data models and access interfaces. The ever increasing amount of data (produced by satellites or the new generations of parallel supercomputers) makes it mandatory to have an efficient data management.

Related to this is a field that crystallized during the workshop: data mining. While the submitted papers only marginally touch this topic, the discussions held during the workshop quickly showed that the problems related to data mining are very similar to those encountered in bringing visualization and databases together. In fact they are basically the same. For the next workshop (to be held parallel to Visualization '97) data mining will be an important part of the topics to be discussed.

1 Workshop Goals

As in the first workshop, the primary goal was to provide an open forum for interested researchers to discuss the important issues concerning the integration of database systems and data visualization systems. Of course the developments during the two years since the last workshop were of special interest. The progress was evaluated and new fields of interest were established, based upon this evaluation.

The exchange of information and discussion between the researchers from different fields was very important, with the aim of establishing a communication that survives the end of the workshop. Although this cross-fertilization was better than two years ago, it can still be improved in the future.

This book is another goal of the workshop, collecting research papers from scientists working in these fields and giving them the opportunity to publish their latest results as well as outlining the problems of current interest and guiding researchers into new activities.

2 Workshop Format

The papers were solicited in the Call for Participation. The working subgroups during the workshop were chosen accordingly. After a short introduction each participant had the opportunity to sketch his work, so that each participant could get an impression about the others. Following these presentations a discussion was held to collect topics to be discussed in the subgroups. These topics were distributed into three subgroups: Scientific Data Modeling, Supporting Interactive Database Exploration, and Metadata, and the participants worked on these topics. The key research issues that were worked out can be found in the subgroup reports that form the first part of this book. The contributions of the participants describing their work can be found in the second part.

3 Contents Overview

The three workshop subgroup reports summarize the discussion sessions held by the participants. The content of these reports was determined by the subgroup participants based upon the topics agreed upon by all workshop participants. Naturally the style of these reports depends on the subgroup members, there was no general guideline for their appearance. It is also clear that the fields pointed out by the subgroups are not completely distinct, but overlap to a certain degree. So the three reports should not be viewed as independent articles but as a whole intended to describe the current research field of database issues for data visualization.

In contrast to the first book the articles are not grouped according to the subgroups. Since the subgroups were formed during the workshop and not based on the submissions it is not possible to assign every article to an appropriate subgroup. Instead the articles are ordered in a way that leads the reader from a general overview via some theoretical papers to case studies of existing systems.

The first three papers (Lee, Grinstein, and Mukherjea) are more general overviews of the field of visualization and databases. Some concepts are introduced and explained and interesting directions are examined. The papers from Hibbard and Kao take a look at the mathematical basics of data and its structure. Beginning with Wittenbrink's paper the section of system description starts. All these articles present a software system and an application. The authors describe their systems, their advantages and disadvantages and why they are written the way they are.

4 The Editors

Andreas Wierse is a doctoral candidate at the University of Stuttgart. He is employed at the Institute for Computer Applications, Department for Computer Simulation and Visualization. He is fully integrated into the Visualization Department of the Computing Center at the University of Stuttgart. Here he is engaged in research on collaborative work, distributed visualization systems, and the data management of these systems. He received his Diploma in Mathematics in 1991 at Bonn University. He co-organized the second Database Issues for Data Visualization Workshop in Atlanta in 1995.

His research interests include distributed visualization environments, distributed data management, and computational steering.

Georges Grinstein is a full time Professor of Computer Science at University of Massachusetts Lowell, Director of the Graphics Research Laboratory, and Director of the Institute for Visualization and Perception Research. He is also a Principal Engineer with the MITRE Corporation. He received his B.S. from the City College of N.Y. in 1967, his M.S. from the Courant Institute of Mathematical Sciences of New York University in 1969 and his Ph.D. in Mathematics from the University of Rochester in 1978. Dr. Grinstein is a member of IEEE's Technical Committee on Computer Graphics and on the editorial board of several journals including Computers and Graphics and the Eurographics Society's Computer Graphics Forum. He was vice-chair of the executive board of IFIP WG 5.10 (Computer Graphics) and was co-chair of the IFIP Conference on Experimental Workstations held in Boston in 1989. He was panels co-chair of Visualization '90 (San Francisco), program co-chair for Visualization '91 (San Diego), conference co-chair for Visualization '92 (Boston) and for Visualization '93 (San Jose), co-chair of the IFIP 1993 Workshop on Cognitive and Psychological Issues in Visualization, and co-chair for the Database and Visualization Issues Workshop. He has chaired several committees for the American National Standards Institute (ANSI) and the International Standards Organization (ISO). He is co-chair for the SPIE '95 Visual Data Exploration and Analysis Conference.

His areas of research include graphics, imaging, sonification, virtual environments, user interfaces, data mining, and interaction, with a very strong interest in the visualization of complex systems.

Ulrich Lang is head of the Visualization Department at the University of Stuttgart Computing Center, which is the First German Academic Supercomputing Center. He received his University Diploma (Dipl.-Ing.) in 1979 and his Ph.D. (Dr.-Ing.) in Mechanical Engineering in 1988, both from the University of Stuttgart. He was a Visiting Scientist at the Center for High Performance Computing of the University of Texas System in 1989. Dr. Lang is a member of the steering committee of the German Computer Society Group on Imaging and Visualization Techniques, member of the organization committee of Eurographics Workshops on Visualization in Scientific Computing, member of the program committee of the Eurographics Workshops on Parallel Visualization, International Liaison Co-Chair of IEEE Visualization conferences, and reviewer for international visualization journals. He is project leader of multiple European and national research, development, and industrial cooperation projects. He is heading the development of a distributed collaborative visualization environment (COVISE) specifically oriented towards High Performance Computer and Network usage.

His areas of activities include distributed software environments, visualization, virtual environments, and interaction, with special emphasis on the handling of large volumes of complex data.

Acknowledgements

We would like to thank all those who participated in the workshop and assisted us by providing timely submissions of extended papers. Special thanks go to David Kao and John Peter Lee for leading the subgroup report writing activity. We also wish to thank Alfred Hofmann, Editor, Springer-Verlag, for his continued support in making these im-portant proceedings available to the research community within the Lecture Notes in Computer Science series.

October 1996

Andreas Wierse
Georges G. Grinstein
Ulrich Lang

Workshop Participants

The following people participated in the workshop:

R. Daniel Bergeron, University of New Hampshire, Durham
rdb@cs.unh.edu

Harvey Davies, CSIRO, Melbourne
hld@dar.csiro.au

June M. Donato, Oak Ridge National Laboratory
donato@msr.epm.ornl.gov

Eve Edelson, University of California, Berkeley
eve@garnet.berkeley.edu

Georges Grinstein, University of Massachusetts, Lowell
grinstein@cs.uml.edu

Bill Hibbard, University of Wisconsin-Madison
whibbard@macc.wisc.edu

David T. Kao, University of New Hampshire
dtk@cs.unh.edu

Daniel Keim, University of Munich
keim@informatik.uni-muenchen.de

Ulrich Lang, University of Stuttgart
lang@rus.uni-stuttgart.de

John Peter Lee, University of Massachusetts, Lowell
jlee@cs.uml.edu

Sougata Mukherjea, Georgia Tech Station, Atlanta
sougata@cc.gatech.edu

Bernice Rogowitz, IBM T.J. Watson Research Center, Hawthorne
rogowtz@watson.ibm.com

Walter Schmeing, VISTEC Software GmbH, Berlin
walt@vistec.FTA-Berlin.de

Bhavani Thuraisingham, MITRE Corporation, Bedford
thura@mitre.org

Aravindan Veerasamy, Georgia Tech Station, Atlanta
veerasam@cc.gatech.edu

Andreas Wierse, University of Stuttgart
wierse@rus.uni-stuttgart.de

Craig M. Wittenbrink, Hewlett-Packard Laboratories
craig_wittenbrink@hpl.hp.com