Databases and Visualization

Daniel A. Keim
Institute for Computer Science, University of Munich
keim@informatik.uni-muenchen.de

1 Description

Over the last decades, the areas of database systems and data visualization developed almost independently. While research on databases deals with the efficient storage and retrieval of large amounts of data, research on data visualization deals with the effective portrayal of data with the goal towards insight about the data. There is, however, a great need for combining the benefits from both areas. On one side, visualization systems deal with very large amounts of data and therefore need database systems as backends; and on the other side, there are very large databases containing a treasure of undiscovered information which could be unveiled using visualization technology. Until recently, however, there has not been much interaction between the two areas. With the exceptions of visualizing database schemas for database design (cf. semantic data models such as the ER-model) and visual querying of the database (cf. visual query systems such as QBE), visualization has been only of minor importance in the database area, especially the aspect of visualizing the data itself. Since size and dimensionality of databases is growing rapidly, the visualization of the data itself however is becoming increasingly important.

The goal of the tutorial is to show the potentials of visualization technology for databases. The tutorial provides an overview of the state-of-the-art in data visualization, classifying the existing data visualization techniques into eight groups: Geometric, Icon-based, Graph-based, Pixel- oriented, Hierarchical, 3D, Dynamic, and Hybrid Techniques. Besides describing each of these classes, the tutorial focuses on new developments in data visualization, which are relevant to the area of database systems. In particular, we describe a wide range of recently developed techniques for visualizing large amounts of arbitrary multi-attribute data which does not have any two- or three-dimensional semantics and therefore does not lend itself to an easy display. A detailed comparison shows the strengths and weaknesses of the existing techniques and reveals potentials for further improvements. Several examples demonstrate the benefits of visualization techniques for real databases applications. The tutorial concludes with an overview of existing database visualization systems, including research prototypes as well as commercial products.

SIGMOD '96 6/96 Montreal, Canada
© 1996 ACM 0-89791-794-4/96/0006...$3.50

2 Outline

1. Introduction
2. Data Preprocessing Techniques
3. Data Visualization Techniques
   - Geometric Techniques
   - Icon-based Techniques
   - Pixel-Oriented Techniques
   - Hierarchical Techniques
   - Graph-based Techniques
   - 3D-Techniques
   - Dynamic Techniques
   - Hybrid Techniques
4. Comparison of the Techniques
5. Database Exploration and Visualization Systems
6. Summary and Conclusion

3 Target Audience

The target audience of the tutorial are researchers and practitioners interested in the state-of-the art of data visualization and its potential application to databases. The tutorial addresses in particular people from academia who want to integrate database and visualization systems for the purpose of data exploration, and people from industry who want to apply visualization technology in exploring large databases.