VisDB: A System for Visualizing Large Databases

Daniel A. Keim, Hans-Peter Kriegel

Institute for Computer Science, University of Munich
Leopoldstr. 11B, D-80802 München, Germany
e-mail: {keim,kriegel}@informatik.uni-muenchen.de

1 Introduction
The VisDB system developed at the University of Munich is a sophisticated tool for visualizing and analyzing large databases. The key idea of the VisDB system is to support the exploration of large databases by using the phenomenal abilities of the human vision system which is able to analyze visualizations of mid-size to large amounts of data very efficiently. The goal of the VisDB system is to provide visualizations of large portions of the database, allowing properties of the data and structure in the data to become perceptually apparent.

2 Visualization Techniques
Two aspects distinguish our visualization techniques from other approaches: First, the number of data items which can be visualized is much higher than in other approaches. In general, our techniques use each pixel of the display to represent one data value. This means that the number of data values which can be visualized at one point of time is only limited by the number of pixels of the display. A second unique feature of our techniques is that the generated visualizations are query-dependent. Query-dependency means that not only the data items fulfilling the query are visualized, but also a number of data items that only approximately fulfill the query. The approximate results are determined by calculating distances for each of the attributes and by combining them into an overall distance [KKS 94].

The basic idea for visualizing the data is to map the distances to colors and represent each data value by one or multiple colored pixels. The colorscale maps correct answers to yellow and approximate results to colors ranging from green, blue, and red to almost black (decreasing lightness denotes an increasing distance from the correct answers). For presenting the data on the screen, five different visualization techniques are available. Three of the techniques (Spiral, Axes-, Grouping Technique [Kei 94, KK 94]) have been developed at the University of Munich; the remaining techniques (Parallel Coordinates [ID 90] and Stick Figures [PG 88]) have been adapted for the purpose of visualizing large databases. The possibility to switch between multiple visualization techniques makes the VisDB system ideally suited for comparing and evaluating the different visualization techniques.

In interfacing with database systems, interesting new requirements on the underlying database servers emerge. First, the VisDB system requires fast access to approximate results, which must be supported by adequate multidimensional data structures; and second, adequate support for incrementally changing queries is necessary.

3 Interactive Graphical User Interface
The graphical user interface of the VisDB system allows a direct interaction with the data. Users may, for example, interactively change the query ranges, highlight corresponding pixels in different windows, use a projection of the visual representation to specific color ranges, or access the attribute values that correspond to some specific pixels. For details of the interface see [Kei 94].

4 Applications
The VisDB system has been used in our molecular biology project to find possible docking regions by identifying sets of surface points with distinct characteristics. Currently, we explore several other data sets including a large database of geographical data, a large environmental database, and a NASA earth observation database.

Acknowledgements
We would like to thank Thomas Seidl and Juraj Porada for their contributions to the design and implementation of the VisDB system.

References