A Spreadsheet Approach to Information Visualization

Ed Huai-hsin Chi, Joseph Konstan, Phillip Barry, John Riedl
Department of Computer Science, University of Minnesota
4-192 EE/CS Building, 200 Union St. SE, Minneapolis, MN 55455
{echi,konstan,barry,riedl}@cs.umn.edu

ABSTRACT

In information visualization, as the volume and complexity of the data increases, researchers require more powerful visualization tools that allow them to more effectively explore multi-dimensional datasets. In this paper, we show a novel new visualization framework built upon the spreadsheet metaphor, where each cell can contain an entire dataset. Just as a numerical spreadsheet enables exploration of numbers, a visualization spreadsheet enables exploration of visualizations of data. Our prototype spreadsheets enabled users to compare visualizations in cells using the tabular layout. Users can use the spreadsheet to display, manipulate, and explore multiple visual representation techniques for their data. By applying different operations to the cells, we showed how visualization spreadsheets afford the construction of 'what-if' scenarios. The possible set of operations that users can apply consists of animation, filtering, and algebraic operators.

KEYWORDS: Visualization, Information Visualization, Interactive Graphics, Spreadsheet

INTRODUCTION

Visualization research spans a remarkable range of scientific disciplines and corresponding visualization techniques. Visualization researchers have discovered that certain operations are needed across this entire range. These operations include comparing visualizations of two different datasets, as well as performing algebraic operations between two or more visualizations, such as visualizing the differences between two datasets. Furthermore, the need to explore multiple visual representations simultaneously arises especially in information visualization, because different techniques often extract different visual features and the complexity of the data. A visualization spreadsheet is an ideal way to address these issues that involve multiple visualizations.

Spreadsheets have proven to be highly successful tools for interacting with numerical data by applying algebraic operations, manipulating rows or columns, and exploring "what–if" scenarios. Spreadsheet techniques have recently been extended from numeric domains to other domains [3, 2, 5]. Here

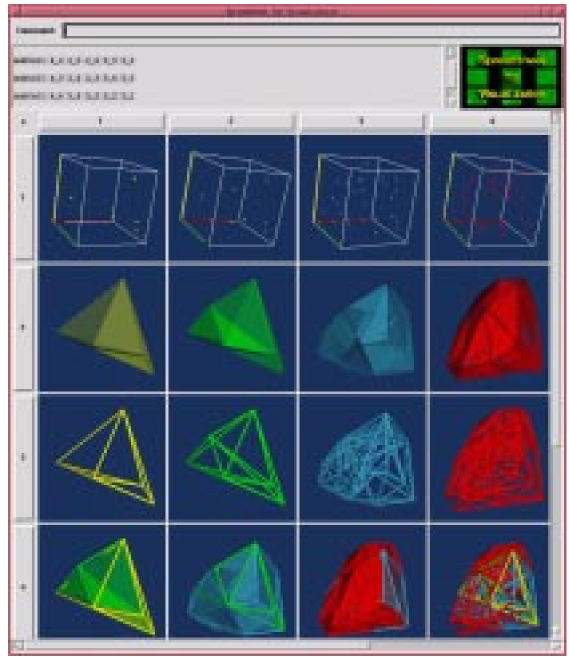

Figure 1: Visualization of 3D random point generation and Delaunay triangulation of the resulting point set. The columns visualize the outcome of the algorithm after 5, 6, 25, and 50 steps, respectively. The last row shows the result of several addition operations. For example, Cell (4,4) shows the result of adding all of the cells in Row 3 together, resulting in a visualization that shows the progress of the algorithm by representing the steps using colored tube edges.

we show a novel spreadsheet framework designed to support cells containing complex datasets, viewed through powerful information visualizations, with constraints between cells linking both data and view attributes.

By studying a spreadsheet approach to the display and exploration of scientific and information visualizations, we can build upon current interaction methods by consolidating all the operations under one paradigm. In this demonstration, we present two prototype systems to illustrate some of the important capabilities offered by the spreadsheet paradigm. We discuss how spreadsheet techniques can provide a structured, intuitive, and powerful interface for investigating information visualizations of abstract multidimensional datasets.

VISUALIZATION SPREADSHEET

In this demonstration, we will demonstrate the utility of the spreadsheet paradigm using two prototype spreadsheets. One prototype visualization spreadsheet is built on top of VTK [4] and Tcl/Tk, and has several advanced capabilities such as user-defined operators and script language support. This prototype, which we call the Spreadsheet for Information Visualization (SIV), is shown in Figure 1.

Here we show how a spreadsheet can be used to visualize several steps of the 3D Delaunay triangulation algorithm. This example shows how algorithm visualization can be supported in our visualization spreadsheet. The columns show the results of the algorithm after 5, 6, 25, and 50 steps, from left to right respectively. The layout is accomplished by loading in a 25 line script that describes the content of each cell and how they are inter-related.

We can utilize multiple visual representations to enhance the comprehensibility of the visualization. Even though each cell in Column 1 represents the same point set, each cell shows a different view of the data, and extracts different visual features. Row 1 shows the point set using 3D scatter plots. Row 2 shows the same data using transparent tetrahedra after 3D Delaunay triangulation has been performed on the point sets. Row 3 represents the tetrahedra using edges between vertices. The last row aggregates several cells together to form new visualizations. For example, cell (4,1)=(3,2)+(3,1)+(2,2)+(2,1).

The combination of geometries from several cells results in visualizations that show differences between successive steps of the algorithm. The aggregation is performed using an addition operator that adds the geometric content in several cells. Subtraction between cells is also available.

Another prototype, shown in Figure 2, is built for viewing biological sequence similarity reports. The cells are loaded with similarity data for genetic sequences. Each comb glyph within a cell represents an alignment, which is region of similarity between the input sequence and a sequence from the database [1]. The prototype implements domain-specific subtraction and addition operators to enable users to compare and contrast between different datasets.

The user can explore the data further by such means as interactively rotating, translating or scaling the representation, following a hyperlink to the textual report, mapping the data into a different geometric representation, animating the information over a variable, and filtering the data. The animation tool provides accumulative, or sliced animation over any variable [1], and adds an extra dimension to the spreadsheet. A synchronized animation can be performed on a group of cells simultaneously. The filtering tool enables the user to explore subsets of the data. When the user interactively adjusts sliders controlling each variable, the view is updated in real-time. The mapping tool enables the user to select the variable that she wishes to be represented on the scatter plot axes. Having these capabilities inside the spreadsheet environment enables a wide range of different exploration tasks.

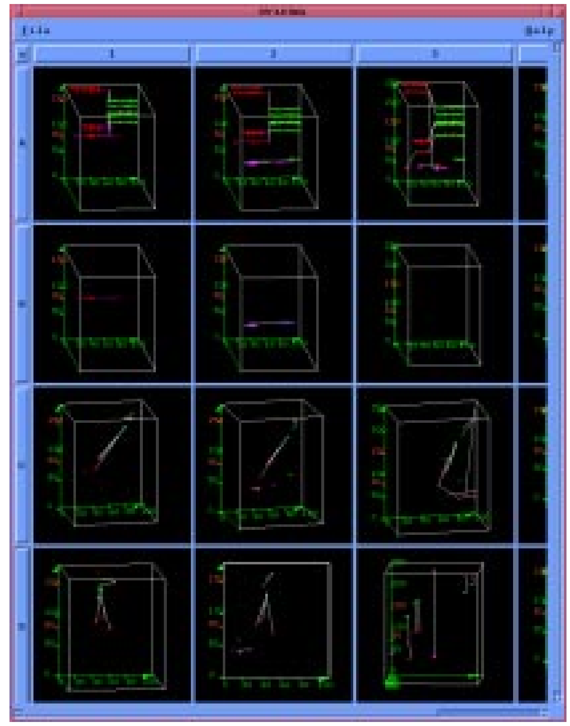

Figure 2: A screen snapshot of the first system (SSR) after performing three operations. (Step 1) Initially, we loaded each column with a slightly different, but related, dataset (A1=B1=C1=D1, A2=B2=C2=D2, A3=B3=C3=D3). (Step 2) We selected Row B, and then subtracted cell A3 from it (B1=B1-A3, B2=B2-A3, B3=B3-A3). Cell B3 contains the empty set as expected. (Step 3) We changed Row C and D to show different views of Row A. The views show different sets of variables using a different representation, thus increasing our ability to see other dimensions of the multivariate datasets simultaneously.

Acknowledgments

This work has been supported in part by the National Science Foundation under grants BIR 940-2380 and CDA 9414015.

REFERENCES

1. E. H. Chi, J. Riedl, E. Shoop, J. V. Carlis, E. Retzel, and P. Barry. Flexible information visualization of multivariate data from biological sequence similarity searches. In *IEEE Visualization '96*, pages 133–140, 477. IEEE CS Press, 1996.
2. A. F. Hasler, K. Palaniappan, and M. Manyin. A high performance interactive image spreadsheet (IISS). *Computers in Physics*, 8(3):325–342, May/June 1994.
3. M. Levoy. Spreadsheet for images. In *Computer Graphics (SIGGRAPH '94 Proceedings)*, volume 28, pages 139–146. SIGGRAPH, ACM Press, 1994.
4. W. J. Schroeder, K. M. Martin, and W. E. Lorensen. The design and implementation of an object-oriented toolkit for 3d graphics and visualization. In R. Yagel and G. M. Nielson, editors, *IEEE Visualization '96*, pages 93–100. IEEE CS Press, 1996.
5. A. Varshney and A. Kaufman. FINESSE: A financial information spreadsheet. In *IEEE Information Visualization Symposium*, pages 70–71, 125, 1996.