Scientific Visualization versus Information Visualization

Henrik R. Nagel

Norwegian University of Science and Technology
Department of Computer and Information Science
`hrn@idi.ntnu.no`

Abstract. While Scientific Visualization techniques are used for the clarification of well-known phenomena, Information Visualization techniques are used for searching for interesting phenomena. There are therefore important differences between the two fields of research. The two fields of research are compared, but since this paper is meant for Scientific Visualization researchers, focus will be on explaining Information Visualization techniques. This paper list the main characteristics of the two fields of research, describes common Information Visualization techniques and discusses differences and similarities in the software that commonly is used in the two fields of research. . . .

1 Introduction

Some researchers mainly use visualization for finding interesting phenomenas in completely unknown data, whiles others use visualization for the confirmation or rejection of hypotheses. These two scientific communities live separate lives, with very little sharing of knowledge between them. While there naturally are important differences between the two kinds of visualization, it is the hypothesis of this paper that an understanding of both of these two ways of using visualization potentially could lead to new ways of visualizing data for the benefit of both communities.

This article therefore contains a summarization of the most important and irreconcilably differences between the ways of thinking. However, since this article is presented at a Scientific Visualization symposium, the focus will here be on Information Visualization ideas and techniques.

2 Principal Differences

Card et al. [1] defines the two forms of visualization as:

Scientific Visualization: the use of interactive visual representations of scientific data, typically physically based, to amplify cognition.

Information Visualization: the use of interactive visual representations of abstract, non-physically based data to amplify cognition.

While Scientific Visualization covers accurate visualizations of the real world, Information Visualization covers visualization of concepts that often are abstract in nature. The purpose of Information Visualization is to make it possible for analysts of data to obtain internal mental models of the information content in datasets; models which subsequently can be used for characterization, prediction, and/or decision making.

Keim et al. [2] classifies visualizations into the categories:

Exploratory Analysis: the search for hypotheses

Confirmatory Analysis: the confirmation or rejection of hypotheses

Presentation: the presentation of facts that are fixed a priori

2.1 Information Visualization and Exploratory Analysis

This example is completely imaginary.

A grocery store sells Feta-cheese for use in salads. It notices that it is loosing money on this product and therefore decides to stop selling it. However, after having stopped selling this product, the grocery store notices that it now looses far more money than it did before. The reason for this is that the few customers who bought this product were wealthy and also bought many other products at the grocery store. When the store stopped selling the product, the customers went to other stores where they could buy it and, at the same time, purchased all other products they needed there.

The grocery store has modern cash-registers and records all purchases in a large database, so the information about the customers' purchase patterns were, in principle, available before the store made the erroneous decision of stopping to sell Feta-cheese. An Information Visualization analyst studying the grocery store's database, in the state that it was before the erroneous decision was made, should by performing an Exploratory Analysis, in principle, have been able to foresee this, without having been told anything about the data in advance.

2.2 Scientific Visualization and Confirmatory Analysis

This example is from the IEEE Visualization 2006 conference, where a visualization contest will take place. The subject of the contest is the TeraShake 2.1 earthquake simulation data set and the task is, through visualizations, to answer five questions, such as "Which regions produce wave reflections?". There is thus, in this case, a clear hypothesis that some of the regions will produce wave reflections and this hypothesis must either be rejected or confirmed by finding these regions. Using the terminology of Keim, the task is therefore to perform a "Confirmatory Analysis" of the data.

3 Common Visual Data Exploration Techniques

In the following is listed a few examples of the techniques that are used in Information Visualization to explore datasets that initially are completely unknown.

3.1 Icons and Glyphs

Markers are featureless objects for representing positions in scatter plots, but multiple, clearly distinguishable markers are sometimes used in one and the same scatter plot to visualize important information. While markers usually only have positional attributes, there are objects that are designed to represent many pieces of information simultaneously. Such objects are called *icons* or *glyphs* ([3, 4]) and have attributes, such as position, size, shape, color, and orientation, to which data is mapped. The hope is that users, by visual exploration, may be able to find patterns among the objects.

3.2 Brushing & Linking

Brushing & Linking is a widely-used visual data exploration technique, with which multiple different visualizations of a dataset are viewed simultaneously. This technique is often used to combine the advantages of multiple forms of data visualization techniques, but can also be used to e.g. allow simultaneously brushing of markers in multiple scatter plots, organized in a scatter plot matrix. The advantage is that the interaction involved is both unambiguous and simple to perform and therefore seem familiar and natural to most human beings.

In the beginning of a visual data exploration process, all markers have a common color. The process consist of

1. Selecting one of multiple views.
2. Brushing selected markers in this view with a color.
3. Inspecting the other (linked) views to see the effect.

3.3 The Grand Tour

In "The Grand Tour" technique [5–7] high-dimensional data are considered as forming a shape in a high-dimensional space. and the method therefore consists of moving a 2D or 3D "window" around this shape, so that a low-dimensional projection of it becomes visible. Significant features of the high-dimensional shape may thereby become distinguishable, and, if one is lucky, one may even be able to guess what kind of high-dimensional shape the data form.

This technique is often combined with "Brushing & Linking", so that subsets of the markers at any time may be brushed with a color to make them distinguishable from the other markers.

4 Software

There are also differences in the software that typically is used for Scientific Visualization and Information Visualization. Consider for example the Visualization Toolkit (VTK) [8], which is an open source visualization system that dates back to 1993. It was designed to be robust, understandable, extensible, modular and maintainable. For that purpose, it offerers scientists a consistent

object-oriented programming model that in 1993 was considered sophisticated, where visualizations are created by linking together a series of objects into a pipeline.

Seen from an Information Visualization perspective, it is clear problem that the visualization pipeline is designed to be "single-pass", where only the resulting visualizations can be interacted with and where interaction usually is limited to rotating the visualization. Information Visualization techniques, such as Brushing and Linking require continuous interaction with the a database containing the original dataset. This requirement is not listed above, which makes VTK unsuitable for Information Visualization.

However, many Scientific Visualization analysts study "Computational Steering," with which feedback can be given to running simulations, based on visualizations of early results from the simulations. The requirements for these kinds of software therefore approach the requirements that Information Visualization scientists place on the software they use.

5 Conclusion

This paper presented Scientific Visualization and Information Visualization with emphasis on the latter. It describe a number of Information Visualization visual data exploration techniques and finally compared the requirements for the software used in both disciplines.

References

1. Card, S. K., Mackinlay, J. D., Shneiderman, B.: Readings in Information Visualization: Using Vision to Think Morgan Kaufmann (1999)
2. Keim, D. A., Ankerst, M.: Visual Data Mining and Exploration of Large Databases Tutorial, Conf. on Principles of Knowlege Discovery and Data Mining Freiburg, Germany (2001)
3. Carr, D. B., Nicholson, W. L.: Evaluation of Graphical Techniques for Data in dimensions 3 to 5: Scatterplot Matrix, Glyph, and Stereo Examples Proceedings of the Section on Statistical Computing (1985) 229–235
4. Pickett, R. M., Grinstein, G.: Iconographic Displays for Visualizing Multidimensional Data Proceedings of the IEEE Conference on Systems, (1988) 514–519
5. Asimov, D.: The grand tour: A tool for viewing multidimensional data SIAM. Journal of Science and Statistical Computing **6** (1985) 128–143
6. Buja, A. and Asimov, D.: Grand tour methods: an outline. Computing Science and Statistics: Proceedings of the Seventeenth Symposium on the Interface (1985) 63–67
7. Wegman, E. J.: The grand tour in k-dimensions Computing Science and Statistics: Proceedings of the 22nd Symposium on the Interface (1991), 127–136
8. Schroeder, W.J., Martin, K., Loresen, W.: The visualization Toolkit: An Object Oriented Approach to Computer Graphics, Third Edition Kitware, Inc. (2004)