INTERACTIVE VISUALIZATION OF ENVIRONMENTAL DATA

Eva-Maria Stephan

University of Zurich
Winterthurerstrasse 190
CH - 8057 Zurich (Switzerland)
email: stephan@gis.geogr.unizh.ch

Abstract

This paper is concentrating on interactive visualization techniques for viewing multivariate environmental data. Based on common visualization techniques, extended techniques, which link thematic mapping of multivariate data to interactive computer graphics, are introduced. Examples are given to emphasize the necessity for *interactive* manipulation of data display, for visual data exploration in the context of environmental data modelling.

1 Introduction

Environmental sciences involve the combination and analysis of many different data sets. Often, data are measured and uncertain, they are multivariate and representing local and global characteristics. In integrated analysis based on some model, measured data often are combined and interpolated to obtain continuous and/or contiguous data. Techniques of interactive computer graphics and scientific visualization [Mac94],[HM90] introduce new ways for exploring such data.

The paper names some common visualization techniques for spatial data visualization and extends these and static cartography, which is one of the oldest sciences using data visualization to present spatial structures in maps.

2 Spatial data and computer graphics

Common graphic representations of spatial data range from *realistic* display of data, which show spatial phenomena in 2- and 3-dimensional (2D and 3D) images, to *symbolic* depictions, which abstract and analyze data, e.g., in statistical plots. Referring to Foley [FvFH90] visualization techniques can be classified by type of object and representation:

Spatial objects (2D): spatial attribute data given as regular grid, vector or point data, can be displayed as color coded image (2D realistic) in, e.g., a vegetation map, or as graphs (1D symbolic), e.g., histogram.

Time series of spatial objects: series of spatial attribute data, e.g., given as sequence of 2D grids, point or vector data, can be shown as animation of images (3D realistic) or graphs, e.g., an animated sequence of 12 curves showing the monthly precipitation during one year in a histogram.

3D spatial objects: spatial attribute information in combination with the digital elevation model (DEM) can be displayed as a perspective view, i.e., as rendered surface with attribute data mapped onto the DEM by image warping.
Data may also be represented in symbolic 2D plots, e.g., histogram or scatterplot, with one axis representing the attribute value and the other the height, or frequency. Often information in rendered surfaces is not visible due to hidden line or hidden surface removal. This may be overcome by allowing interaction with the display or the object, e.g., interactive change of perspective.

Time series of 3D spatial objects: multiple time-dependent data of a 2D object in combination with the DEM may be displayed as *animation* of an image warped terrain. The disadvantage is a fix or predefined change of the viewing angle.

2.1 *Environmental data are multivariate and uncertain*

Different to most spatial objects, environmental data are representing continuous and contiguous natural processes. Their formal and graphical representation goes beyond representing discrete spatial and temporal values [Vck95], [Ste95]. Environmental data in particular are *multivariate* and *uncertain.*

In order to visualize multivariate and uncertain spatial data, an extension of the above mentioned visualization techniques is introduced in 3. Techniques are designed for two main purposes, to support the analyst in the modelling process and to extend common graphics techniques for interactively viewing environmental data.

3 Environmental modelling and interactive visualization techniques

As mentioned, environmental modelling involves the combination of multiple data sets. Intergrated analysis of such data requires the exploration of local and global characteristics, uncertainties, trends and correlations between data sets. Data visualization techniques need to be designed along these requirements, to support explorative data analysis and the modelling process. Next, multidimensional dynamic visualization techniques for investigating and exploring multivariate spatial data are presented. They go beyond predefined data display, encouraging - even requiring - the analyst to interact with displayed information and to look *behind* color images. Basic interaction techniques include changing viewing angle, zooming, adjusting color, shading, visibility and individual scaling of objects[1]

3.1 *Exploring multivariate data*

Representing multiple data sets simultaneously requires the *combination* of symbolic and realistic visualization techniques and the dynamic *linking* of data displays in multiple windows. Techniques for comparing and exploring correlations between data sets are:

Multiple symbolic representations: Multiple 2D plots of different attribute layers, or a temporal sequence of one attribute information can be shown parallel in one or multiple windows. If multiple plots are overlaid in one display, they can be distinguished by different colors or plotting symbols. Problems of different value ranges of attribute data can be overcome be scaling all value ranges to the size of the plotting window. An interactive selection of one curve by mouse click results in adjusting the annotation of the y-axes to the selected curve. If plots are displayed in separate plotting windows, the selection of values in one window should result in simultaneous high-lighting of the corresponding values in the other plots. Dynamic graphics techniques for data analysis were introduced by Becker and Cleveland [RBW87], an example is given figure 3 II.

Multiple realistic representations: Different realistic visualization techniques may be combined in one window. One possibility is to show a color coded image of one data set as base image from a 45 degree angle, and overlay a wire-frame of the DEM and a contour plot of a third layer. Interactive change of perspective, transparency and interactive cutting provide

[1] Throughout the paper the following terminology is used: Raster data sets referring to the same area (not necessarily the same spatial reference frame) are called *layers*.

simultaneous viewing of three different attribute layers. If several images or surfaces in different windows are shown, those may be linked by selecting a specific area in one display, resulting in a simultaneous high-lighting of the corresponding area in the other display. An example is shown in figure 3 I.

Pseudo realistic representations: Instead of warping attribute information onto the DEM, it may be warped to any attribute data. This technique enables 3D thematic mapping for any spatial attribute data and allows direct comparison of two attribute layers (see figure 3 III.).

Stacking multiple pseudo surfaces above one-another, with individual control of exaggeration, position and transparency, enables to compare many layers directly in one view.

Linked symbolic and realistic representations: Dynamic linking of displays offers nearly unlimited possibilities, and is "only" limited by computing power. Statistical views, like histograms, scatterplots etc. may be linked with image or surface representations in different windows so that, data sub-selection in one view is simultaneously high-lighted in all displays. For example, sweeping a horizontal cut through the DEM will dynamically display 2D curves of attribute data along that line in separate windows. Another possibility is to select data dynamically by rectangular or self-defined masks as the mouse is moved over an image and get local analysis of the sowith selected data values, e.g., the arithmetic mean or a profile plot. Several examples of cutting planes and profile plots are given in figure 3 I linked with II and V linked with VI.

3.2 *Exploring uncertain data*

As mentioned, data uncertainty is a mayor issue in environmental modelling. Uncertainty results from interpolation of sampled data and/or modelling (combination, integrated analysis) of multivariate data. Cartographic representation of uncertainty was discussed by Buttenfield [But93], Mac Eachren [MD92], and Beard [BM93] and concentrated n static displays. The following visualization techniques showed to support the interactive analysis of data uncertainty:

- dynamic high-lighting sampling points helps the analyst to locate where *original data* was sampled, e.g., measured from (see figure 1)
- dynamic adjusting of certainty levels, e.g., a image dynamically alters as threshold values of uncertainty are changed by a slider bar. Less uncertainty can be shown by darkening colors, transparency, or fading pixels [Fis92]. This techniques allows to draw objects with smooth boundaries, which represents well the *uncertain* data.
- overlaying contour lines on images and surfaces, with interactive control of contour levels, may be used for representing certainty ranges.

In addition, data models can be validated and different simulation results compared and investigated with the previously outlined techniques for exploring multiple data layers. It must be mentioned, that the techniques for depicting uncertainty require the complying information to be given with the data sets[2].

4 Examples for exploring environmental data

Next, examples of interactive visualization follow. Screenshots were taken from the prototype-system *DataScaping*[3], which was implemented with IDL [RSI93] for exploring environmental

[2] Problems of adequate storage of data quality is not subject of this paper. If treating certainty information as separate data layer or additional vector data, any visualization technique for multiple data sets is applicable.

[3] The term was derived from the term *Landscape-Viewing*, but rather than photo-realistic images of landscapes, *DataScaping* presents any data layers as rendered scenes.

data[4]. The system provides loading of data, e.g., generated with a Geographic Information System (GIS), and various visualization facilities. All displays provide necessary controls, e.g., buttons, slider-bars, menus, for adjusting viewing angle, zooming, object scaling, color mapping, and empirical statistical analysis of data sub-selections.

4.1 *Sampled Data*

Figure 1 shows a daily temperature distribution in a 1 km resolution. Sampling points and the interpolated regular grid of the daily mean temperature are shown in the left image. As the cursor is moved over the image, additional information can be interrogated and displayed in a separate window. In this case the daily temperature distribution of *sampling station no. 5* is shown in the center image. The plotting window adapts its curve as the cursor is moved over a different sampling station.

The image to the right - showing a rendered surface of the DEM in 1 km resolution - was linked to the temperature image and indicates the location of sampling points.

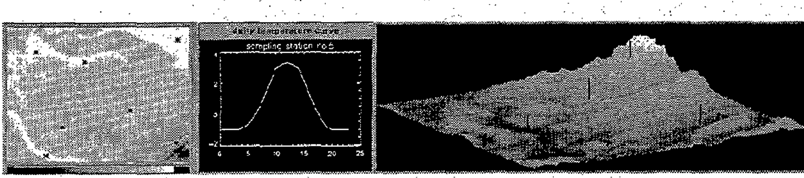

Fig. 1. Temperature, daily curve and DEM with Sampling stations

4.2 *Exploring correlation between temperature and the DEM*

Figure 3 gives a an impression of a typical *DataScaping* session. The screenshot was taken during the analysis of a 1 km raster temperature data set - different area as in the previous example - which is compared to the DEM. Figure 2 gives an overview of the arrangement of different windows and their controls, which are described next.

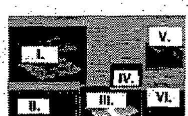

Fig. 2. Overview of plots in figure 3

Window V is an *information window* about the DEM data. It is linked with VI. Controls provide interactive horizontal, vertical and self-defined cuts with profile plots in window VI. Furthermore, the color-coded image may be queried by mouse for single or multiple neighbor data values.

[4] Please note, that interactive and highly dynamic processes cannot sufficiently be narrated in a static paper.

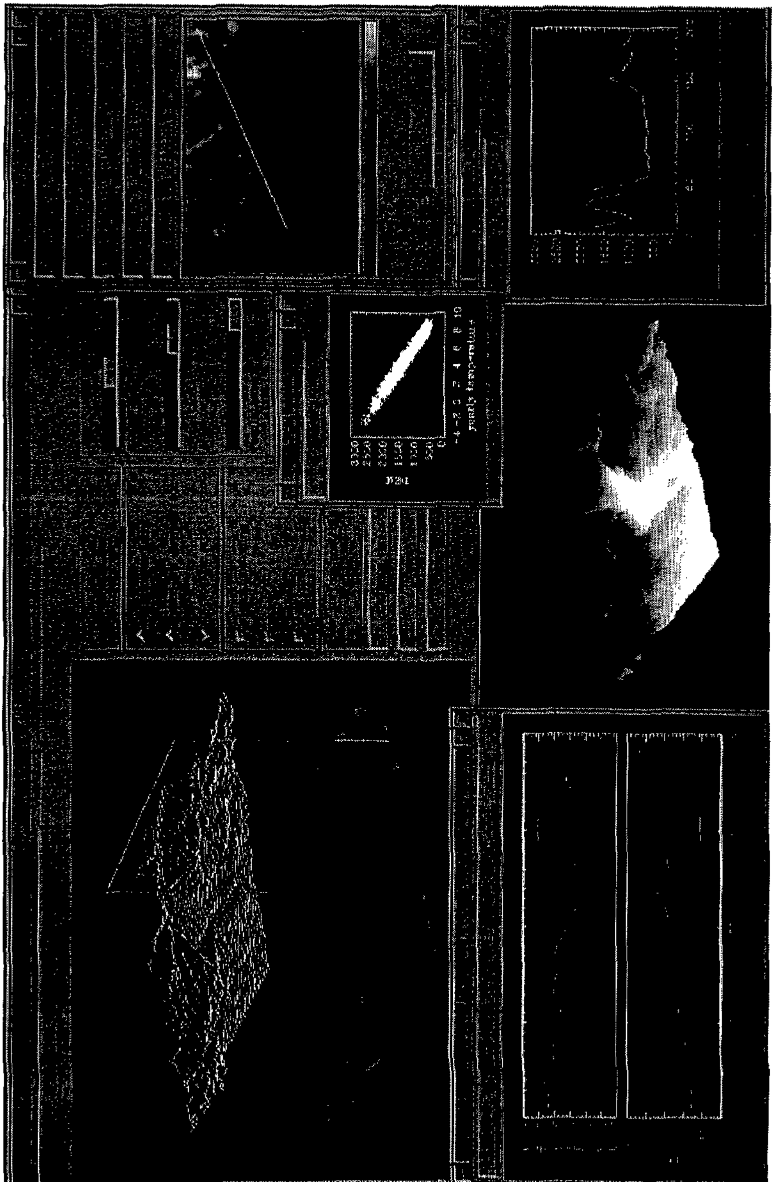

Fig. 3. Example of an interactive data exploration

Image III and IV show a *combined realistic and a symbolic view* of both data layers. III is the DEM surface, which is color coded by the temperature data. This technique does not necessarily require both data sets to have the same resolution, whereas the symbolic scatterplot in IV only maps values with the same reference. Both images prove a clear correlation between the two data layers.

Window I shows both data sets as stacked surfaces and some controls of the *two-surface viewer window*. Surfaces may also be shown as wire-frames or rendered surfaces; each surface can individually be positioned and scaled in exaggeration and resolution. The option *interactive profile plot* was turned on and the frame n I indicates where the profile plot in I refers to. The curves in II adapt smoothly as the mouse is moving over the above image. On mouse click, vertical cuts in the east-west-direction are enabled.

5 Conclusion

Today's geographic systems do not include high-level interactive graphical procedures for exploring the relationships between spatially configured objects and the multivariate properties of the attributes associated with those objects. The paper pointed out specific requirements for interactive exploration of multivariate environmental data and gave visualization examples. Those showed that interacting with multidimensional views of data requires techniques which meet special charcateristics of environmental data. The presented visualization techniques for *multivariate* and *uncertain* data, offer new ways for exploring environmental data and their interrelationships. Allowing data analysts to experiment with different views of data, encourages to go beyond the image and learn about underlying data characteristics.

Future work will concentrate on the refinement of techniques and visualization metaphors, and on formal and graphical representation of uncertainty in data.

Acknowledgements : This research project has been funded by the Swiss National Science Foundation under contract No. 50-35036.92.

References

[BM93] Kate Beard and Wiliam Mackaness. Visual access to data quality in geographic information systems. *Cartographica*, 30(2):37–45, 1993.

[But93] Barbara P. Buttenfield. Representing data quality. *Cartographica*, 30(2):1–7, 1993.

[Fis92] Peter F. Fischer. Real-time randomization for the visualization of uncertain spatial information. In *Proc. of the 5th Int. Symp. on Spatial Data Handling*, pages 491–494, 1992.

[FvFH90] Foley, vanDam, Feiner, and Hughes. *Computer Graphics - Principles and Practice*. Addison Wesley, 2 edition, 1990.

[HM90] Robert B. Haber and David A. McNabb. Visualization idioms: A conceptual model for scientific visualization systems. In Shriver Nielson and Rosenblum, editors, *Visualization and Scientific Computing*, pages 74–92. IEEE Computer Society Press, 1990.

[Mac94] A.M. MacEacheren. Visualization in modern cartography: Setting the agenda. In A.M. MacEacheren and D.R. Taylor, editors, *Visualization in Modern Cartography*, pages 1–12. elsevier Sciences Ltd, 1994.

[MD92] A.M. MacEacheren and D.W. Dibiase. Visualizing uncertain information. *Cartographic Perspectives*, Fall(13):10–19, 1992.

[RBW87] William S.Cleveland R. Becker and Allen R. Wilks. Dynamic graphics for data analysis. *Statistical Science*, pages 355–395, 1987.

[RSI93] RSI. *IDL - Interactive Data Language*. Research Systems, Inc., 1993.

[Ste95] Eva-M. Stephan. Exploratory data visualization for reliable integration. In *Proceedings of the AutoCarto 12*, 1995.

[Vck95] Andrej Vckovski. Representation of continuous fields. In *Proceedings of the AutoCarto 12*, 1995.