INTERACTIVE MAPS FOR VISUAL EXPLORATION OF GRID DATA

Gennady Andrienko[1], Natalia Andrienko[1], Valery Gitis[2], and Ivan Denissovitch[2]

[1]GMD - German National Research Center for Information Technology
AiS.KD - Autonomous Intelligent Systems Institute, Knowledge Discovery Team
Schloss Birlinghoven, Sankt-Augustin, D-53754 Germany
http://borneo.gmd.de/and/
e-mail: gennady.andrienko@gmd.de
tel: +49-2241-142329, +49-2241-142486, fax +49-2241-142072

[2]IITP – Institute for Information Transmission Problems RAS
http://www.iitp.ru/
e-mail: gitis@iitp.ru

KEY WORDS: Grid-based data, visualization, interactive maps

EXTENDED ABSTRACT

Grid-based data format with attribute values specified in nodes of a regular rectangular grid is most often used for conveying information about continuous spatial phenomena, for example, elevation or density of earthquakes. Visual representation of grid-based data is necessary for supporting exploratory analysis of such phenomena, i.e. attempts to reveal characteristic features of their spatial variation and relationships between different phenomena. We propose a set of interactive techniques intended to support visual exploration of numeric grid data on the basis of their representation on a map. Some of them are extensions of techniques we developed earlier for attribute data referring to discrete spatial objects represented as points or polygons (Andrienko and Andrienko 1999a).

Grid data are displayed on maps as raster images where colors of pixels encode attribute values. Attribute values and, respectively, colors for pixels falling between grid nodes are determined through interpolation. A similar procedure of color encoding of numeric values is applied in producing choropleth maps used to represent values referring to units of territory division.

There are two opportunities for encoding numeric attribute values by colors. One is to match the number interval to be represented on a map with a continuous color scale where shades change gradually. The other variant is to divide the interval into subintervals and assign a distinct color to each subinterval. Hence, attribute values fitting in the same subinterval are represented identically. The latter procedure is known as classification, and the subintervals are called classes. The first variant of encoding is often referred to as "unclassed".

Both in raster representation of grid data and in choropleth maps various color scales can be used (Brewer 1994), in particular, a diverging color scale. Such scale uses two different color hues for values below and above certain number, the so-called midpoint. The distance of a value to the midpoint (i.e. the difference between the two numbers) is encoded by the degree of darkness. In an unclassed variant the degree of darkness is proportional to the distance. In a classed variant the parts of the attribute value range below and above the midpoint are subdivided into classes, and different degrees of darkness are assigned to the classes according to their distance to the midpoint.

In our system Descartes, we have implemented an interactive tool that accompanies an unclassed choropleth map with a diverging color scale. This tool allows the user to move the midpoint of the scale within the value range of the attribute and immediately observe the effect of her/his actions on the map (Andrienko and

Andrienko 1999a). This operation, called "visual comparison", can make spatial patterns of value distribution more prominent. We found that the technique of "visual comparison" can be effectively applied also to grid-based data in order to support exploration of spatial variation of continuous phenomena. In the recent version of the system we have implemented a "visual comparison" tool for grid data.

Classification, being implemented as interactive, dynamic procedure, can also be a valuable instrument of exploratory analysis of spatial data (Andrienko et al. 2001b). For classed choropleth maps representing a single attribute we have developed an interactive tool that allows the user to add, move, and remove class breaks, with any slight change being immediately reflected on the map. Recently we have developed a similar tool for grid data.

Classification may be also used to summarize values of two attributes (cross-classification) or even more. In Descartes we had a tool for interactive cross-classification in application to choropleth maps (Andrienko and Andrienko 1999a) enabling the same operations as the tool for one-dimensional classification. For more than two attributes the system can build a choropleth map on the basis of classification by dominance between attribute values (Andrienko and Andrienko 2001a). Analogous techniques can be used for combining data from two or more grids.

One more useful technique applicable to grid data is focusing. In Descartes focusing may be used in conjunction with many methods of cartographic representation of thematic data available in the system (Andrienko and Andrienko 1999a). It allows the user to remove some part of the value range of an attribute from view (in particular, outliers, i.e. very high or very low values, in comparison to the rest) and represents the remaining part with the maximum possible expressiveness. Thus, in a case of color encoding, the whole color scale is matched with a shorter number interval. This makes differences between close values more noticeable.

In application to grid data focusing plays also an additional role: removing of fragments of a raster image from view uncovers the parts of the map beneath. This may help in investigation of relationships between the phenomenon represented by the raster and other phenomena, which may be represented either by vector or raster data. For the same purpose semi-transparent display of a raster may serve. A recently implemented in Descartes tool makes it possible to vary the degree of transparency of a raster image representing grid data. The user may specify any transparency value from 0% (fully opaque) to 100% (fully transparent, i.e. invisible).

Figures 1 to 3 demonstrate the result of application of the described tools to grid data, using a digital elevation model of the area of Turkey as an example. Figure 1 illustrates variation of image transparency. Figure 2 shows the effects of focusing and visual comparison. In both images in Figure 3 elevation values are classified into 3 classes. The second image is produced from the first by means of movement of the boundary between the second and the third class.
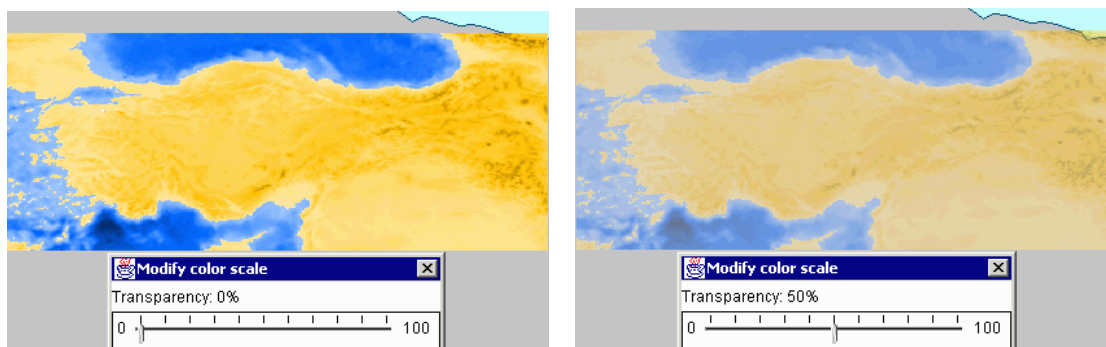

Figure 1. A raster image is shown in opaque (left) and 50% transparent (right) modes.
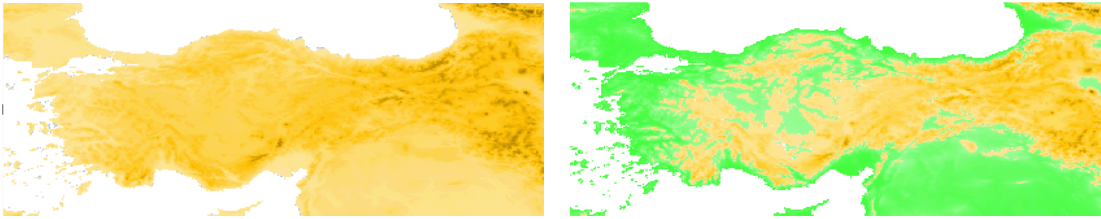
Figure 2. Left: a result of focusing of the map from Fig.1 with only positive elevations being shown. Right: additionally to the focusing, a diverging color scale with the midpoint value +1000 is applied.
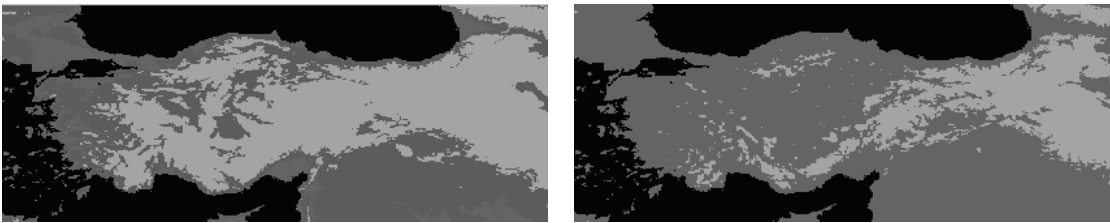

Figure 3. An effect of movement of a class boundary upon a classification-based raster image. On the left: the class breaks are 0 and 1000m. On the right: the second break is changed to 1500m.

Currently following operations over grid data are available in Descartes:

- Selection of a color scale for encoding attribute values;

- Selection of the variant of encoding: classed or unclassed;

- Interactive change of the midpoint of a diverging color scale;

- Interactive change of class intervals for classed encoding;

- Variation of the degree of transparency;

- Focusing on value subinterval.

In the nearest future we plan to investigate the possibilities of combining data from several grids by means of cross-classification and classification by dominant attribute. We would also like to extend the analytical facilities of the system with respect to grid data by applying to them non-cartographic visualization techniques. Thus, a very useful tool for revealing correlation between attributes is the dot plot widely used in statistics. We expect that this display can be used to grid data, specifically, to two grid-related attributes. Of course, it is impossible to show on a scatter plot values referring to all locations since a grid corresponds to a continuous fragment of territory with an infinite number of locations. It is possible, however, to represent values for a finite number of sample locations. The sample locations can be selected automatically with regular spatial intervals between them that can be specified by the user. Varying the intervals, the user changes the level of detail of the derived discrete data represented on the plot. The user may also interactively select locations to be added to an automatically built scatter plot. This procedure of data discretization can also be used to enable application of data mining methods to grid data. In (Andrienko and Andrienko 1999b) we described how exploratory analysis of spatial data can benefit from combination of cartographic visualization with methods of data mining, using one specific method as an example. We foresee that a similar synergic effect of the two approaches can be achieved for grid-based data.

REFERENCES

Andrienko, G. and Andrienko, N., 1999a. Interactive Maps for Visual Data Exploration. *International Journal Geographical Information Science*, Special Issue on Visualization for Exploration of Spatial Data, Vol. **13** (4), pp.355-374.

Andrienko, G. and Andrienko, N., 1999b. Data Mining with C4.5 and Cartographic Visualization. In, N.W.Paton and T.Griffiths (Eds.) *User Interfaces to Data Intensive Systems*, IEEE Computer Society, Los Alamos California, pp. 162-165.

Andrienko, G. and Andrienko, N., 2001a. Exploring Spatial Data with Dominant Attribute Map and Parallel Coordinates. *Computers, Environment and Urban Systems*, Special Issue on GIS Research UK - 2000, Vol. **25** (1), pp.5-15.

Andrienko, G., Andrienko, N., and Savinov, A, 2001b. Choropleth Maps – Classification Revisited. In, *Proceedings ICA 2001* (accepted).

Brewer, C.A., 1994. Color Use Guidelines for Mapping and Visualization. In, *Visualization in Modern Cartography* (NY: Elsevier Science Inc.), pp.123-147.

Gitis, V., Osher, B., Dovgyallo, A., and Gergely, T., 1999. GeoNet: An Information Technology for Space-Time WWW On-line Intelligent Geodata Analysis. In, R.J. Peckham (Ed.), *Proceedings of the 4th EC-GIS Workshop*, JRC, Ispra, pp.124-135.