A GIS-based Module for Training and Visualization of Self-Organizing Maps

Martin Lacayo
Department of Geography
San Diego State University
San Diego, California, 92182-4493
USA
lacayo@rohan.sdsu.edu

André Skupin
Department of Geography
San Diego State University
San Diego, California, 92182-4493
USA
skupin@mail.sdsu.edu

ABSTRACT

A crucial element of efficient geovisual analytics is an integration of traditional geographic information technology with various computational and visual methods developed in other domains. Some of those methods have shown great promise in being able to shed light on complex *n*-dimensional relationships, but their adoption is being hampered by a lack of integrative approaches. One example is the self-organizing map (SOM) method (also known as Kohonen map), which has been applied to a wide array of data, from population attributes of geographic objects to the semantic content of text documents. While the promise of the SOM method has been widely acknowledged, its adoption in geographic contexts has been hampered by a lack of easy-to-use tools. On one hand, there are a number of existing out-of-the box SOM solutions, like *SOM_PAK* or the *SOM Toolbox for Matlab*, but their use for geographic analysis has been characterized by extremely loose coupling, and as a result one observes a certain disconnect with other means of geographic analysis. On the other hand, SOM software components are included in such development systems as *GeoVista Studio*, where tight integration with other components in the design of custom analytics approaches has been demonstrated. However, in either case, the tremendous power of existing GIS software for simultaneously dealing with large databases and complex geometries has remained largely untapped. Existing approaches also severely limit how much control an analyst can exert over the geometric and visual transformations of a self-organizing map, compared to the design choices offered in a modern GIS environment.

For all of these reasons, this paper presents an implementation of the SOM method in the context of GIS. The main goal of this GIS-based module is to fill the need for tools that facilitate the training and display of a trained SOM and the mapping of other high-dimensional data onto the SOM within a standard GIS environment. Analysts would thus be able to remain within the framework of a familiar interface, while tightly integrating geographic visualization with spatialization towards enhancing the explanatory power of geographic databases. The module wraps standard SOM training software from within GIS and then outputs the trained SOM as a Shape file, where neurons of the artificial neural network are given two-dimensional coordinates, to which *n*-dimensional attributes are attached. Other elements included allow the mapping of *n*-dimensional data onto the SOM, which is demonstrated for a multi-temporal data set of population census data, with geographic objects becoming represented as *n*-dimensional trajectories. All of the geometric and attribute elements created are immediately and completely usable in off-the-shelf GIS, thereby enabling geometric transformation (e.g., interpolation of component plane surfaces), visual transformation (e.g., advanced label placement), and simultaneous exploration of geographic and attribute space. The paper discusses the methodology and implementation of this module, the use and significance of specific features, and enumerates some areas requiring further development.