GeoVISTA

Visual Inquiry Toolkit:

Integrating Visual, Computational, and Cartographic Methods

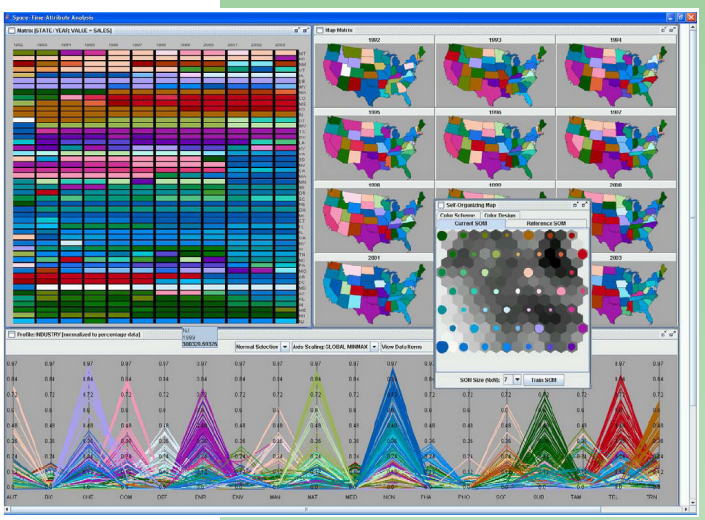
Jin Chen, Diansheng Guo, and Alan MacEachren

http://www.geovista.psu.edu

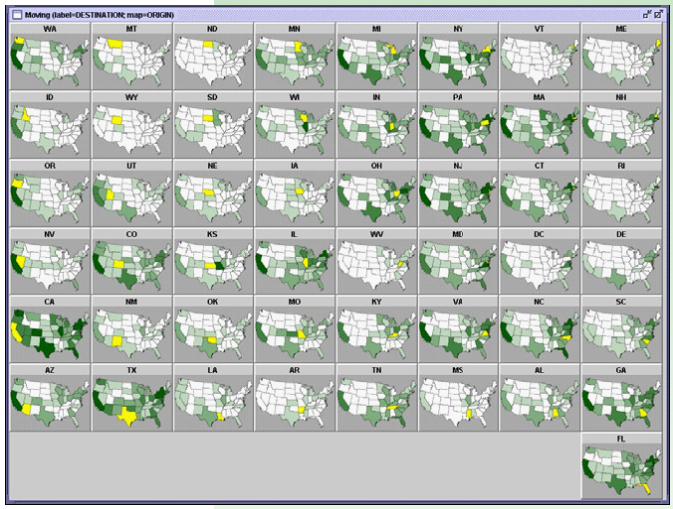
The Visual Inquiry Toolkit (VIT) integrates visual, computational, and cartographic methods to enable human knowledge and judgment to be coupled productively with computational methods for incrementally searching patterns. The VIT was initially developed, entered, and took first place in the 2005 IEEE Information Visualization Contest (www.infovis.org), held annually to promote advances in information visualization. Contestants must develop interesting and insightful ways of visualizing and analyzing a complex benchmark data set provided by the contest organizers. The 2005 data set contained geographically-referenced statistics on employment, sales, and company relocations within selected U.S. industries over a twelve year period. The GeoVISTA Center's winning submission demonstrated geo-visual analytic strategies for detecting and exploring multivariate, spatio-temporal patterns. By leveraging GeoVISTA *Studio* as a component-sharing and application-building environment, the entry integrates information visualization and cartographic methods with computational methods for sorting and finding clusters in large data sets. The integrated VIT approach to geo-visual analytic methods and tools is able to:

- perform multivariate analysis (including time series analysis) with the Self-Organizing Map (SOM)
- encode the SOM result with colors derived from the Color-BrewerPlus component, which produces a 2D diverging-diverging color scheme
- visualize the data in a hierarchical data matrix view
- visualize the multivariate patterns with a modified Parallel Coordinate Plot (PCP) display and a map matrix
- support human interactions to explore and examine patterns

This research shows that such integrated methods (computational and visual) can mitigate each other's weakness and collaboratively support discovery and analysis of complex space-time attribute patterns, in an effective and efficient way. We are now generalizing the VIT for application to cancer incidence and covariate data.



Using a map matrix, data matrix, parallel coordinate plot (PCP), and SOM, the image above represents sales in 18 US industries over time. Colors are shared across views, and similar colors depict similar mulitvariate data items. Even in this static image, we can easily spot states (across years) that have similar industry makeup and investigate how that makeup varies over time and space. Dynamic linking among views provides added analytical power.



To explore company relocations from state to state, we developed a novel map matrix (Map^2 above). The overall view is a schematic "map" made up of multiple component maps. Each component map represents company moves into one destination state from all origin states. The destination state for each map is labeled and signified in yellow in the component map. Color coding signifies relocations from each origin state (darker green = more). The overall matrix of maps (arranged in a roughly geographic pattern) provides a national overview of relocation patterns.

This material is based upon work supported by the National Institutes of Health under Grant # R01 CA95949-01 Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Institutes of Health or the National Cancer Institute.

PENNSTATE 1855