外国語要旨

“Composite Time-varying Data Visualization Applying Level-of-detail Control”

Sayaka Yagi

“Big data” is a common term of the information technology industry in recent years. Various computer systems store huge information from various datasets including computer system logs, sensor measurements, and transactions. Since these datasets easily get huge due to the time-by-time accumulation, it gets more complicated in analyzing and understanding characteristics of such dataset directly on a screen. Moreover, such time-varying values are often associated with other information; for example, temperatures can be associated with weather attributes and/or geo-locations, stock prices can be associated with social or economic incidents, and Web access logs can be associated with link structure on web pages. This thesis focuses on such “composite time-varying data” and proposes visualization techniques for simultaneously observing numerical time-varying values and associated information. This information visualization approach is useful and effective for understanding, analyzing, and monitoring of composite time-varying datasets. It is also efficient for supporting active decision making in business. Although information visualization techniques are often used to analyze composite time-varying data, many techniques have focused on representing multi-variable data and have not been specialized to observe time-varying values. This thesis aims to effectively utilize associated information of time-varying data by contriving data mapping methods and applying level-of-detail control techniques, on the basis of standard visualization principle; "Overview first, zoom and filter, then details-on-demand" .

The thesis firstly proposes a visualization tool to support analysis and knowledge discovery including geographic problems, especially for flood control operations. It is important to geographically represent the distribution of numerical information related to the weather and water level, so that users can realize combinational analysis, monitoring, and alert of the danger of floods. The tool adequately and quickly places rectangles which represents numerical information onto maps. It provides two types of representation for the numerical information; one is better to quickly look over the distribution of the values and find danger regions, the other is adequate for locally observing the focused regions. Users can switch the two types according to their preferences, usages and characteristics of datasets. The tool also enables level-of-detail by means of interactively zooming in and out to broaden or narrow down the region of interests.

This thesis then proposed two versatile visualization methods for a larger-scale composite time-varying data; polyline-based and storyline-based visualization techniques. Polyline-based visualization is popular and helpful to understand detailed numerical time-varying values. On the other hand, storyline-based visualization is effective to observe associative features among elements over time. The polyline-based visualization method supposes that

associated information consist of a set of predefined term and are assigned to the values of each time step. The technique maps colors to the associated information and draws assigned colored polylines in a single display space to compare the time variation of multiple values. However, if hundreds or thousands of polylines are rendered in a single space, it may result in severe cluttering, and this may cause difficulty in reading data values and following data trends. Thus, the technique realizes smooth level-of-detail control by interactively controlling the number of polylines to be displayed. It clusters polylines based on their shapes and associated information, and then selects representative polylines from the clusters. The technique also features click and sketch interfaces so that users can interactively select particular polylines which are associated with the user-interested terms. The storyline-based visualization method aims to simultaneously observe short-term features of time-varying values and cluster transitions. The method first measures similarity of elements in each time-step, and divides the elements into clusters. It then defines the cluster layout by matching corresponding clusters between two adjacent time-steps, and draws similar elements as proximity storyline. Reflecting transparency on storyline as a visual variable, the technique also emphasizes the amount of line changes. Moreover, the technique provides a user interface so that users can interactively select interesting parts on storyline, and explore the numerical values by observing a polyline-based visualization.

Adaptation to the concept of Visual Analytics, “Analyze First - Show the Important - Zoom, Filter and Analyze Further - Details on Demand”, is one of the final goals of this study. The Visual Analytics process would be possible by coordinating methods proposed in this thesis. The storyline-based visualization technique could help analysts to observe relevance among composed time-varying data elements and to narrow down the elements for further investigation. Moreover, adopting similar methods to the proposed flood data visualization method could help analysts to understand the focused spatial information or graph structure directly on a display space. Repeating the process and making discoveries into knowledge would lead future prediction of composed time-varying data.