A Similarity-Based 3D Time-Varying Data Visualization Technique

Maiko Imoto*
Ochanomizu University

Takayuki Itoh†
Ochanomizu University

1 Introduction

Time-varying data visualization is an active research topic in these years. We are currently focusing on visualization of various time-varying data, including weather, gene expression, and atomic plant measurement data.

Polyline chart is the most common time-varying data visualization technique. We think famous Information Seeking Mantra [2], which suggests important operations for information visualization "Overview, Zoom, Filter, and Detail on demand", is also important for polyline-based time-varying data visualization. Several interactive techniques [1][4] have realized the filtering of unfocused portions of the time-varying data. We think that still we need to improve the readability of initial representation of time-varying data, so that we can easily discover interesting phenomena.

The poster presents a 3D time-varying data visualization technique, for all-in-one display and extraction of similar values. The technique firstly orders the time-varying data based on their similarity, and places the polylines in the order. It then provides two views, while one view briefly represents the whole time-varying data, and the other represents the partial data extracted by click operations. Users can look over the time-varying data, extract groups of interested similar values, and carefully observe the extracted values, by using this technique.

2 Related Work

We often deal with large time-varying data containing hundreds or even thousands of time-varying values. It is a serious problem that it may be very difficult to understand if we draw such large time-varying data in one polyline chart.

There have been several interactive techniques to easily extract interested polylines from the large data [1][4]. For example, Hochheiser et al. presented a user interface [1] to specify polylines by placing rectangles on the display space.

We are developing two approaches to discover interesting phenomena from initial visualization results of large scale time-varying data. Uchida et al. reported a technique as one of our approaches, which clusters the polylines and renders only representative polylines for each cluster [3]. The technique effectively improves the readability by reducing the number of rendered polylines without missing interesting features of the data. As the other our approach, the technique presented in this poster provides a user interface to easily extract groups of similar polylines.

3 Time-Varying Data Visualization

Most of existing polyline-based time-varying data visualization place sets of polylines onto 2D (xy-place) display spaces. On the other hand, the technique presented in this poster places the polylines onto a 3D space (xyz-space). Figure 1 shows the overview of data placement by the technique. It firstly orders the polylines based on the similarity of their feature values, and then assigns z-coordinate values to the polylines according to the order. The placement is useful to extract groups of similar polylines. It then provides two views, while "Viewpoint 1" has a ray perpendicular to xz-plane, and "Viewpoint 2" has a ray perpendicular to xy-plane. The first view provides an overview of the whole time-varying data and a user interface to extract interested polylines, and the second view provides zoomed display of the extracted polylines.

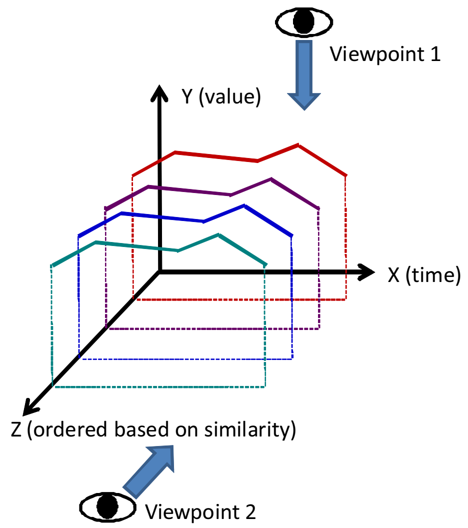

Figure 1: 3D placement of polylines. The technique orders the polylines based on their similarity, and assigns z-coordinate values according to the order.

The technique consists of the following three components: 1) ordering of polylines, 2) overview of the whole time-varying data in the left-side of the window, 3) display of the extracted polylines in the right-side of the window.

3.1 Ordering of polylines

This section describes the j-th value at the i-th time as a_{ij}, and the j-th time series values as $A_j = \{a_{i1}, .., a_{in}\}$, while the data has n time steps. Here, the technique calculates distances between arbitrary pairs of polylines, and constructs a dendrogram of the polyline to order them. Current our implementation prepares several distance calculation schemes so that users can select one of them according to their use scenario. The poster introduces a scheme that calculates distances based on positions of polyline-vertices. The scheme calculates the distance between the j-th and k-th polylines $D(j,k)$ by the following equation:

$$D(j,k) = \sum_{i=1}^{n} \sqrt{(a_{ij} - a_{ik})^2} \quad (1)$$

After calculating the distances between all the possible pairs of polylines, the scheme constructs a dendrogram. Initially the scheme treats a polyline as a cluster, and merges the closest clusters into a single cluster. The merge process is repeated until all the polylines are connected by a single dendrogram. Finally, the scheme orders all the polylines by using the dendrogram.

*e-mail: maiko_i@itolab.is.ocha.ac.jp
†e-mail:itot@is.ocha.ac.jp

3.2 Overview

As shown in Figure 1, the technique prepares two viewpoints. These are used to global and local visualization of the time-varying data. The technique firstly displays the whole data from ”Viewpoint 1”, in the left side of the window, as shown in Figure 2(1). Here all the polylines are drawn as ribbons, while they are shaded after setting several light sources. The polylines are colored according to y-coordinate values of the polyline-vertices: current our implementation assigns warm colors (e.g. red) to higher y-coordinate values, and cool colors (e.g. blue) to lower y-coordinate values. The technique initially adjusts ”Viewpoint 1” to display the whole data in one space; however, users can control the viewpoint and freely focus on specific parts of the data.

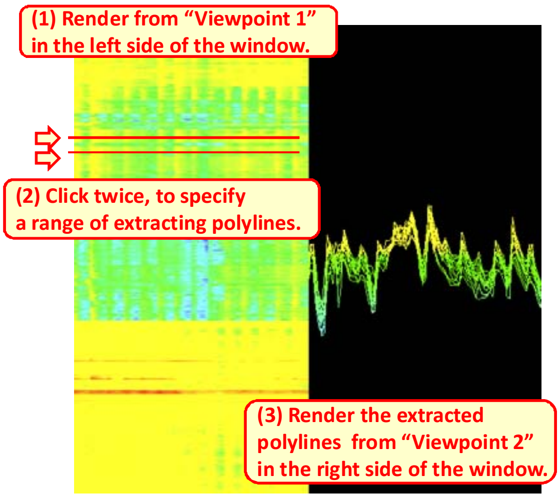

Figure 2: Processing flow of the presented technique.

3.3 Similar polyline extraction

The technique also provides a user interface to extract specific similar polylines on the left side of the window. Our implementation supports two-click operation to specify a range of extracting polylines, as shown in Figure 2(2). When a user specifies a range, the technique renders the extracted polylines from ”Viewpoint 2” and displays on the right side of the window, as shown in Figure 2(3). Again, the technique initially adjusts the viewpoint to display the whole part of the extracted polylines; however, users can freely control the viewpoint to focus on specific parts of the extracted polylines.

As above mentioned, the technique supports dual display of the time-varying data, so that users can initially overview the whole data, extract interesting similar polylines, and focus on the extracted polylines to carefully observe their coherency or differences.

It is often effective to assign another color in the right side, by calculating from other values or metadata of the polylines. We support the coloring according to the gradients of polylines as one of the variety of color calculation. It can effectively emphasize suddenly changing values.

4 Experiment

Figure 3 shows an example of temperature data recorded by Japanese weather measurement system AMeDAS. We applied the data recorded in January 2006 at 913 points in Japan. We clicked a part of the data in the left side, as marked by a red rectangle. We focused this part because its color changed between yellow and green very frequently. We then observed the specified polylines in the right side. The temperatures were totally less-varying between daytime and night, but actually changed between yellow and green almost every day. Most of the polyline looked totally very similar; however, they were very different in a few days, as marked by a sky blue circle. It was difficult to recognize such local difference in the left side, but we found it by carefully observing in the right side. The example denotes that the mechanism of our technique is effective, to overview the whole data first, and then extracts parts of the data to carefully observe. Another interesting point is that the extracted polylines corresponded to temperatures at different regions of Japan, including Ibaraki prefecture (along Pacific Ocean), Ishikawa prefecture (along Japan Sea), and Nara prefecture (inland). Usually we do not think such different regions had similar variation of temperature. The example also denotes that we found interesting and unexpected phenomena by the presented technique.

5 Conclusion and Future Work

The poster presented a 3D time-varying data visualization technique, which places polylines in the similarity-based order. Our potential future works include validation of the polyline ordering algorithms, scalability tests, evaluations with experimental users, and experiments with other time-varying data.

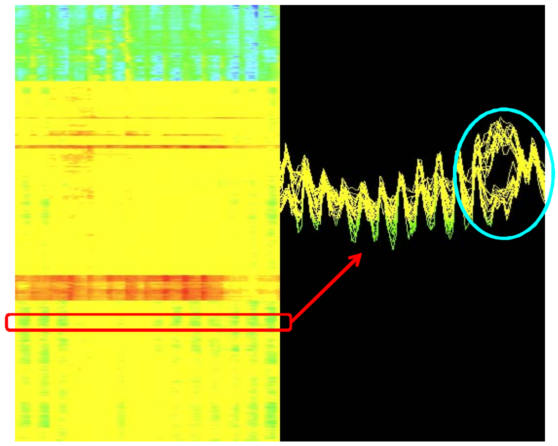

Figure 3: Visualization of temperature measurement data in Japan. A user clicked the region marked by red rectangle. The extracted polylines were very similar, but partially different as marked by a sky blue circle.

References

[1] H. Hochheiser, B. Shneiderman, Dynamic Query Tools for Time Series Data Sets: Timebox Widgets for Interactive Exploration, *Information Visualization*, 3(1), 1-18, 2004.

[2] B. Shneiderman, Dynamic Queries for Visual Information Seeking, *IEEE Software*, 11(6), 70-77, 1994.

[3] Y. Uchida, T. Itoh, A Visualization and Level-of-Detail Control Technique for Large Scale Time Series Data, *IEEE Information Visualization, Poster Session*, 2008.

[4] M. Wattenberg, D. Jones, Sketching a Graph to Query a Time-Series Database, *SIGCHI Conference on Human Factors in Computing Systems Extended Abstract (CHI2001)*, 381-382, 2001.