A 3D Visualization of Multiple Time Series on Maps

Sidharth Thakur, Andrew J. Hanson
Renaissance Computing Institute, Indiana University Bloomington
sthakur@renci.org, hansona@cs.indiana.edu

Abstract
In the analysis of spatially-referenced time-dependent data, gaining an understanding of the spatio-temporal distributions and relationships among the attributes in the data can be quite difficult. We present a visualization technique that addresses some of the challenges involved in visually exploring and analyzing the distributions of geo-spatial time-varying data. We have developed a pictorial representation that is based on the standard space-time cube metaphor and provides in a single display the overview and details of a large number of time-varying quantities. Our approach involves three-dimensional graphical widgets that intuitively represent profiles of the time-varying quantities and can be plotted on a geographic map to expose interesting spatio-temporal distributions of the data. We show how combining our visualization technique with standard data exploration features can assist in the exploration of salient patterns in a data set. The visualization approach described here supports expeditious exploration of multiple time-varying quantities having a geo-spatial component. The main components in our visualization include a 3D graphical representation that we have developed to intuitively represent the time-varying data and an adaption of the standard space-time cube system [15] to generate a holistic display of the spatio-temporal distributions of the data on a geographic map. In addition, we employ standard visual-analytical tools such as interactive data mapping and filtering techniques to support exploratory analysis of multiple time series. These methods assist an analyst to quickly identify important patterns in the spatio-temporal distributions in the data and allow expeditious exploration of multiple numeric quantities.

1 Introduction
In many domains that deal with information and scientific data, a frequently occurring and important attribute of the data is time. Time is often also considered significant because the temporal domain has unique characteristics (e.g., inherent levels of granularity such as day, month and year) compared to some of the other types of fundamental data quantities, such as space and population [1].

A rich variety of visual-analytical methods have been developed to assist in the exploration of time-varying data (e.g., see [2]). However, effective exploratory analysis of geo-spatial data that involve multiple time-varying quantities can be quite challenging.

The analysis of geo-spatial time-dependent data typically requires an understanding of the distributions of the patterns of changes and relationships among data attributes over both space and time [4]. These spatio-temporal relationships can be especially complicated when the data involve a large number of individual time-varying data series that correspond to multiple variables [25].

Figure 1: A dense plot showing temporal changes in unemployment rates in North Carolina’s (USA) 100 counties.

A standard example of data containing multiple time-indexed numeric quantities is census data. Besides containing demographic data, census records often contain a number of socio-economic factors that indicate growth in one or more geographic regions (e.g., per capita income, unemployment rates, and poverty). Standard representations such as the line graph shown in Figure 1 are useful for exploring temporal patterns in a large number of time series. However, it is challenging to graphically expose both the overview and details of such spatial-temporal distributions in a single line graph.

In this paper, we present a three-dimensional (3D) visualization approach to address some of the challenges in effective visual exploration of multiple time-varying quantities having a geo-spatial component. The main components in our visualization include a 3D graphical representation that we have developed to intuitively represent the time-varying data and an adoption of the standard space-time cube system [15] to generate a holistic display of the spatio-temporal distributions of the data on a geographic map. In addition, we employ standard visual-analytical tools such as interactive data mapping and filtering techniques to support exploratory analysis of multiple time series. These methods assist an analyst to quickly identify important patterns in the spatio-temporal distributions in the data and allow expeditious exploration of multiple numeric quantities.
We begin with a review of the problem and discuss related work in Section 2. We present our visualization approach in Section 3 and discuss features and design considerations in our method and how they apply to visualization issues. Section 4 deals with various analytical techniques and data exploration tools used in our approach for exploring spatio-temporal data. Finally, we conclude in Section 5 with a summary of our techniques.

2 The Problem and Related Work

Many map-based visual-analytical techniques have been developed for visualizing numeric time-varying data having a geo-spatial context; some of the standard approaches involve animated two- and three-dimensional maps that use color schemes (e.g., choropleths) or graphical icons (e.g., ellipses) to encode and display data values [20]. Other methods involve graphical icons that represent profiles of time-varying quantities and are plotted directly on a map (e.g., bar charts [4] and theme-river icons [21].

However, it is still challenging to visualize holistically the spatial and temporal distributions in geo-spatial time-varying data. For instance, some of the approaches discussed earlier can be limited either because they can handle only a limited number of time steps (e.g., animated displays) or because they cannot support effective display and exploration of complicated spatio-temporal patterns (e.g., bar graphs plotted on a map).

A visualization approach that is capable of showing overview and details facilitates exploring many types of spatio-temporal data. For example, such an approach can allow expeditious exploration of the spatial and temporal distributions in geo-referenced data involving multiple time-varying quantities (e.g., socio-economic indicators in census records). The readily apparent distributions can help in formulating initial hypotheses about the patterns and even stimulate further inquiry by allowing an analyst to pose interesting questions about the data.

Some methods have been developed that can provide overview and details for some types of geo-spatial time varying data. A classic example is the standard space-time cube (STC) metaphor [15] used to display the positions and evolutions of locations of objects as a function of time. The STC approach consists of a 3D plot in which the spatial component (e.g., a geographic map) is plotted in the $X−Y$ plane and the temporal domain is represented on the vertical axis. The STC method has been applied to visualize the spatio-temporal information in a variety of event-related data; two representative examples are the visualization of movement of crowds and traffic [18, 22] and the visual analysis of events [9, 12].

Although the STC approach is primarily applicable for visualizing event-based data, some approaches have explored its use for visualizing time-varying data that are associated with a continuous, linear temporal domain. For example, an interesting application utilizes 3D graphical icons plotted on a geographic map to display multi-variate time-varying data [25]. In another example [8], the time-varying profiles of a temporally evolving financial portfolio are visualized using cylindrical 3D icons whose positions are determined using a network graph.

In our work, the space-time cube provides a basic framework within which we display geo-spatial temporal data. For visualizing the profiles of time-varying quantities, we rely on a two-dimensional (2D) technique called Data Vases [23]. The 2D approach generates graphical widgets (or glyphs) whose shapes represent the profiles of numeric time varying quantities. Figure 2 shows an application of this 2D approach for visualizing a census-related quantity. In the figure, the time scale is vertical, increasing from the bottom to the top, and data values are encoded.
Figure 3: Visualizations of the spatial and temporal distributions of food stamps issued in the 100 counties in North Carolina through a nation-wide program of the federal government. In the standard small multiples [26] display of the data (top) some interesting patterns are visible; for example, higher food stamp claims occur in the north-eastern part of the state. Our 3D visualization on the bottom provides enhanced access to the spatio-temporal distributions of the food stamp claims and selected details of changes in the claims in individual counties, all in a single view.

by the width of the 2D glyphs. Optionally, color is used in this technique to redundantly encode data values and/or statistical quantities such as percent change and moving averages. Another related method for displaying multiple time-varying quantities is the Horizon Graph [16], which employs an effective layering and color coding technique to compactly display data such as stock prices. However, a Horizon Graph is primarily suited for visualizing non-spatial data and is therefore not considered here.

The 2D plot in Figure 2 provides a compact representation of the spatio-temporal relationships in the data. However, the approach provides limited information about the spatial distributions since the horizontal arrangement of the glyphs is restricted to showing gross regional clusters.

3 Visualization Approach

We present here a visualization method that is capable of displaying a large number of time-varying data series (time series) on maps. Our approach involves 3D glyphs based on the 2D data vases technique and adapts the space-time cube system to generate a display that allows effective representation of multiple time series.

3.1 Construction of the 3D Glyphs

Our visualization approach employs 3D glyphs whose shapes represent the profiles of time-varying quantities corresponding to various geographic regions in a data set. Our method involves a straightforward construction in which polygonal disks, one for each time step, are stacked along a vertical temporal axis. The size (diameter) of the disk at each time step is scaled according to normalized data values and based on a pre-determined maximum disk size. The thickness of the disk is usually set to some constant value, though a user can interactively adjust the maximum width and thickness of the glyphs. Finally, we use the colors of the disks to encode data values and statistical quantities.

In the next step, the 3D glyphs representing the time-varying quantities are generated for each of the geographic regions in the data set; each glyph is positioned at the corresponding region’s centroid on a geographic map in the X − Y “ground” plane of the space-time cube. Figure 3 shows a visualization of a time-varying data set generated using our 3D approach and compares our method with a standard approach that involves multiple thumbnail charts.

3.2 Design Considerations and Issues

We next discuss some features and design considerations in our visualization approach for generating informative views of geo-spatial time-varying data.

3D vs 2D. Our primary motivation in choosing a 3D approach based on the space-time cube metaphor is to enable a holistic display and exploration of interesting patterns in space and time rather than presenting the related information as separate spatial and temporal “slices” [12]. Another motivation is based on the usefulness of the space-time cube approach for supporting exploration of spatio-temporal data; for example, a recent study [19] found the space-time cube system to be faster for performing selected spatio-temporal exploration tasks compared to a baseline 2D method. Although a study to evaluate the representations and data in our case still lies in future, we discuss in Section 4 some of the exploratory tasks supported by our approach. Finally, using our 3D approach we can easily display data sets having a total of up to a few tens of thousands of time steps, which may not be possible using
approaches involving 2D graphical icons (refer to Section 2). Figure 4 shows an example visualization of a dense data set using our technique.

Scene simplification. In 3D visualizations a standard problem affecting readability of graphical elements is occlusion caused by overlapping glyphs. Another general factor that affects the complexity of the display is data density; depending on the type of visualization method used, a dense data set can result in overplotting and occlusion of graphical elements. Our technique also suffers from these issues, as shown in the dense visualization in Figure 4.

We deal with the problem of complexity in our 3D visualizations using two general approaches, namely (a) manipulation of display properties, and (b) transformation of data (e.g., data aggregation). Our first approach involves interactive manipulation and transformation of the 3D view (e.g., rotation, translation, and scaling); this straightforward approach is used for reducing occlusion in a scene and for closer inspection of the profiles of the 3D glyphs. Another general technique that we use is filtering or information hiding [25], which employs transparency and hiding of irrelevant information to present a simplified view to an observer; see the inset in Figure 4 for an example. Other methods have been developed that use sophisticated exploratory tools to navigate complex 3D data visualizations [10], but are yet to be explored in our approach.

Our second approach for simplifying scenes is based on standard data transformation methods (e.g., filtering and aggregation) and is discussed in Section 4.

Smooth vs. discrete profiles. The 3D glyphs in our visualizations can be generated using either line graph (smooth) profiles or bar graph (stepped) profiles depending on the sampling in the temporal domain. Figure 5 shows a closeup of some 3D glyphs drawn using the two profiles. Line graph profiles are typically used to display time-dependent data that vary continuously between the time intervals. However, since data values need not change smoothly between adjacent time steps, for example, when time intervals are relatively long (e.g., months and years), the glyphs can be generated using a stepped profile.

Color mapping. Color is used as a redundant graphical attribute of our 3D glyphs (i.e., in addition to the width of the glyphs) to encode data values and statistical quantities. In fact, color is extremely useful for exposing unusual data values such as outliers and abrupt changes, especially when the data set is dense. We use standard available tools and guidelines for generating suitable color encoding schemes [5, 24]. Another color-encoding technique used with our 3D glyphs is binning, which is useful for displaying changes in data values or for highlighting clusters of similar data values. In binning, the data values are categorized according to user-defined intervals and the attributes of the corresponding graphical elements (e.g., colors of the disks) are determined based on a discrete color scheme. Examples of this approach are shown in Figures 3 and 4, where data values have been binned into six intervals.
Figure 6: Visualizations of the distribution of the numbers of alcohol-related accidents reported in the 100 counties of North Carolina generated using distinct data mapping schemes: (left) absolute values, and (right) per-capita values.

Orthographic vs perspective projection. The choice of projection for the 3D glyphs in our visualization is an important consideration because our method depends on the visual comparison of the shape and size of the 3D glyphs. We choose an orthographic (parallel) projection to render the 3D scene because it preserves the shapes of the glyphs in various rotationally-transformed views.

Variation on bubble charts. Our 3D glyphs can alternatively be imagined as a stack of bubbles (i.e., ellipses used in traditional 2D map-based plots to encode numeric data). However, in our initial experiments, stacks of individual 2D bubbles were found to be useful in only limited 3D rotated views for observing the profiles of time-varying data. We have therefore chosen solid shapes for our 3D glyphs for representing time-varying data.

Miscellaneous considerations. Although we primarily show visualizations involving geographic maps, our 3D glyphs can be used for displaying time-varying data that have a non-geographic spatial context, for example, abstract layouts such as network graphs and tree maps. In addition, it is sometimes useful to display missing or unreported data values, which are easily displayed in our approach by using distinct colors or graphical markers on the 3D glyphs. Finally, the 3D glyphs in our approach are similar to the cylindrical widgets in [8]; however, our method has evolved directly from the 2D data vase approach.

4 Exploratory Navigation In Our Approach

The compact pictorial overviews of spatial-temporal data that are generated using our approach allow a user to quickly browse through a large number of time series in multiple data sets. However, our objective is not limited to (just) creating interesting displays but also supports various exploratory tools that can enable an analyst to pose interesting questions and hypotheses relating to his or her data. In this section we discuss some visual-analytical tools and exploratory tasks supported by our approach.

4.1 Exploratory Tools and Features

We employ some standard data manipulation methods as the basis of interactive exploratory tools in our approach; the following is a discussion on these techniques.

Data mapping. A fundamental operation during the construction of any visualization is the mapping of data values to various attributes of graphical icons or glyphs (e.g., size and color). The options for mapping different data-related quantities can itself serve as an effective visual-analytical tool [11]. For example, one option in our method is to map onto the 3D glyphs either the absolute (raw) data values or data values that have been transformed, that is, normalized by some quantity such as population of a region. This straightforward option to switch between absolute and per-capita values allows exploration of interesting spatio-temporal patterns in a given data set; an example of these mappings is shown in Figure 6. Another useful data mapping option is to encode statistical quantities using the size and/or color of the 3D glyphs. For example, visualizing the rate of change of a time-varying quantity (percent change) and moving average can provide useful information during the visual data exploration process.

Data transformation. Data transformation in our context refers to a set of general data processing and manipulation methods that are useful for highlighting salient patterns in a data set. Some of the standard data transformation methods include aggregation, clustering, summarization, and computing correlations [4]. We briefly discuss some of these techniques and their applicability in our approach.

Aggregation and Data abstraction. Aggregation is a common data transformation method in which multiple data points are replaced by a few data points or aggregates that represent the original data set; the aggregate can be derived using a variety of statistical formulations such as sums, averages, and variances [13]. Additionally, data abstraction methods such as summarization employ standard analytical techniques such as dimensionality-reduction to
extract and display salient properties of attributes of multidimensional data [7].

In our approach, data aggregation and summarization serve as a useful tool for generating overviews of the data over space and time [3]. For instance, spatial aggregation is obtained by creating glyphs based on data aggregated over regional clusters. On the other hand, temporal aggregation is achieved by collapsing data values over the multiple levels of granularity inherent in the temporal domain (i.e., collapse monthly values to yearly averages). These aggregation schemes support our overall goal of providing quick overviews and greatly simplify the visualizations of dense data sets. Figure 7 shows an example of these spatio-temporal aggregation schemes.

Clustering. Clustering is a computational technique for identifying subsets that contain similar items based on some similarity measure. A specialized application of this approach is clustering of time series, which is used to determine salient patterns in either multiple distinct time series (whole clustering) or on parts of a single, streaming time series (subsequence clustering) [6]. In our approach, whole clustering can be used to determine spatial distributions of salient patterns in geo-spatial time series data. The profiles of the representative patterns of the time series can be displayed using our 3D glyphs to depict similarity relationships over space and time. Another option is to support visual exploration of patterns generated by a user based on standard querying tools for exploring temporal data [17].

Correlation analysis. A standard analytical technique that is used for understanding relationships in a multivariate data set is based on examining correlations among the multiple attributes in the data. In many visualization approaches, features such as correlation matrices and node-link diagrams can be exploited to display and explore the inter-relationships. In our approach, however, the visual representation is limited to the display of multiple time series corresponding to at most a single attribute. This limitation can be overcome to some extent by displaying correlations between pairs of time-varying attributes. For example, we can pre-compute temporal correlations among a set of data attributes and display the correlation values using our 3D glyphs. An example of a comparative analysis is shown in Figure 8, where the differences between the values of two numeric time-varying quantities are plotted using our 3D glyphs; the visualization has been annotated to highlight features of interest.

Data filtering. Data filtering is a useful technique that can assist not only in the visual-analytical exploration process but can also reduce complexity in a 3D scene by reducing the data points displayed, thereby making it easier to observe patterns in the data. Our visualization approach involves a variety of standard interactive tools to filter data based on user-specified ranges or intervals; for example, (a) selecting a pre-determined interval by clicking on the discrete ranges in the color palette, (b) using a range slider to select a continuous interval, and (c) selecting some arbitrary maximum and rescaling the width and color of the glyphs based on the new maximum value. Figure 8 summarizes the results of applying these filtering techniques.

4.2 Analytical Tasks Supported

Visual-analytical tools are often designed to support two general types of analytical tasks: directed exploration and exploratory navigation [11]. The tasks are differentiated based on whether the exploration involves prior knowledge or hypotheses about the relationships in the data, or whether the tasks are purely exploratory.

Our visualization approach has been conceived to support the latter, exploratory type of interactive data exploration. Our methods can also be used as a data exploration tool to complement some of the existing standard visual-analytical methods for investigating geo-spatial temporal data (e.g., see the methods in [3, 14, 27]).

We next discuss various data exploratory tasks supported by our approach; our discussion is based on a comprehensive data-task framework [4] that describes exploratory/analytical tasks that are useful for analyzing.
spatio-temporal data. For details beyond the brief discussion given here, refer to [24] for an overview of the task typology of the framework.

According to the data-task framework in [4], exploratory tasks can be broadly categorized as elementary tasks and synoptic tasks. Elementary tasks are basic and are usually concerned with determining the values of unknown quantities (i.e., dependent variables) when one or more known quantities (independent variables) have been specified (i.e., lookup of values). Synoptic tasks are more complicated and are related to finding general patterns and relationships among a set of multiple dependent and independent variables. Since the general behavior of a system or phenomenon is usually represented by relations among groups of entities (e.g., system parameters and data attributes), synoptic tasks are considered more important than elementary tasks. However, elementary tasks are still quite useful, for example, to obtain details about individual data points (e.g., data values at selected time steps).

Our visualization approach supports a number of elementary tasks and some synoptic tasks. For example, a user can readily look up where and when certain data values occur. Our method also provides quick overviews of distributions of numeric quantities over space and time. Other, detailed information on spatio-temporal patterns can be obtained by using some of the data mapping and filtering tools discussed earlier in this section. Finally, our method supports some basic synoptic tasks: for example, a straightforward visually-guided task in our approach is the comparison of the time-varying profiles of a quantity across multiple geographic regions. Some other, more complex synoptic tasks that are supported include identification of salient temporal patterns using some of the clustering and correlation techniques discussed earlier.

However, our visualization approach has certain limitations. For instance, it does not currently provide a means to directly display and compare multi-variate time-varying data. In the visualization of Figure 8, the changes in the political affiliations could be due to changes in demographics or other factors; however, these relationships are not shown in the display. Another general limitation pertains to the perceptual issues relating to some of the dense data visualizations generated using our approach; for example, a comparison of multiple 3D glyphs can be difficult, especially when the profiles of time-varying quantities are complicated. A thorough evaluation of the analytical capabilities of our methods will require a comprehensive user study: such an evaluation has been planned as future work.

5 Conclusion

We have presented a three-dimensional visualization approach for displaying profiles of multiple time-varying data series on geographic maps. Our method adapts the space-time cube metaphor and employs intuitive 3D graphical icons to generate a visualization that provides overviews and details of spatial-temporal distributions of time-varying quantities, all in a single view. We present some analytical exploration features and tools that can be used in our visualization approach for supporting various exploratory tasks. Our visualization approach is intended to assist an analyst to quickly explore multiple time-varying quantities on a map and extract useful information that can be used for building hypotheses about the data, as well as stimulating further explorations.
Acknowledgments

This work was carried out at the Renaissance Computing Institute’s (Renci) engagement center at North Carolina State University (NCSU). We would like to thank our colleagues at Renci and NCSU for providing valuable feedback and interesting data sets to work with.

References


