Web Visualization of Geo-Spatial Data using SVG and VRML/X3D

Jianghui Ying
Virginia Tech
Dept. of Computer Science
Falls Church, VA 22043, USA
jying@vt.edu

Denis Gračanin
Virginia Tech
Dept. of Computer Science
Blacksburg, VA 24061, USA
gracanin@vt.edu

Chang-Tien Lu
Virginia Tech
Dept. of Computer Science
Falls Church, VA 22043, USA
ctlu@vt.edu

Abstract

Data visualization is an important technique that helps in understanding and analysis of complex data. Most of current web-based geo-spatial data visualization products use two-dimensional visualization. A tool that uses the combination of two- and three-dimensional visualization can provide extra spatial dimension as well as flexible usability. Current Web technologies like Scalable Vector Graphics (SVG), Virtual Reality Modeling Language (VRML), and eXtensible 3D (X3D) provide a necessary foundation for such a tool. In this paper, we present a general framework for effective integration and web-based presentation of complex heterogeneous spatial-temporal data sets. This framework facilitates interpretation, exploration and analysis of large volume of data with significant geo-spatial and temporal characteristics. Advantages of this approach include improved visualization of geo-spatial and temporal raw data; better navigation and selection of data; and intuitive user interface. A web-based road traffic visualization system illustrates this approach using the Twin Cities road traffic data.

1. Introduction

Data visualization is an important technique that helps in understanding and analysis of complex data. It is an important area of data analysis, where the data collected is summarized and presented in visual form to aid in decision-making and in grasping the minute details and relationships of data sets. Advances in visualization technology have revolutionized fields such as architecture, medicine, etc. However, the ability to visualize complex geospatial systems, like regional infrastructure systems, introduces fundamentally new challenges to theory and application of visualization. The key reason that these systems have not reached their full potential has been an inability to communicate either current system state or expected system state information to decision makers in a meaningful format. The system state description is derived from many data sources so to fuse this data and concisely define the state of the system is a difficult task. Currently, only simplistic graphic maps are used for this purpose. Such maps fail to effectively capture temporal information or any level of spatial resolution. Such systems like road transportation are characterized by presence of rich information. That information has significant geo-spatial and temporal characteristics and consists of large quantities of data from a variety of data sources. Information should be stored efficiently and then integrated in a common framework. A mapping between the information elements and their spatial and temporal locations may be used to provide service information in a meaningful form. The provided service information is, in general, complex, multi-dimensional, physical or abstract information that is intrinsically difficult to represent and manipulate.

2. Web-based geo-spatial data visualization framework

Visualization techniques can be classified by whether data visualization is 2D or 3D based. 2D representations of geographic data have been widely used in many systems. There is an increasing interest in 3D representations of geographic data, especially for universal web access of such data. Both 2D and 3D visualization methods have their own strengths and weakness. When large amount of data has to be presented in a limited space, 2D visualization introduces a cognitive overload and thus heavily reduce usability. Instead, 3D visualization uses an extra spatial dimension to create a virtual world where information is presented. However, delivering 3D content over Internet is more technically challenging. The difficulty of 3D content distribution and navigation over the Internet is the major obstacle preventing the popularity of 3D web site. Detailed 3D models generation is time consuming and may require a lot of “manual” work. Moreover, user interaction in 3D is not always intuitive and may difficult to learn.

The objective of this visualization framework is to provide a good foundation to integrate various data sources and automatically create 2D (maps) and 3D (virtual worlds) visualizations. At the same time, we aim to provide highly interactive and intuitive graphical user interface for users. We try to combine both 2D and 3D to improve usability of the user interface to the largest extent. Specifications such as Scalable Vector Graphics (SVG) [4], Virtual Reality Modeling Language (VRML) [5], and eXtensible 3D (X3D) [5], defined by the World Wide Web Consortium (W3C) are used to display geo-spatial data on web.

The framework for web-based geo-spatial data visualization is shown in Figure 1. It consists of four parts: data source layer, data storage layer, data process layer, and visualization layer.

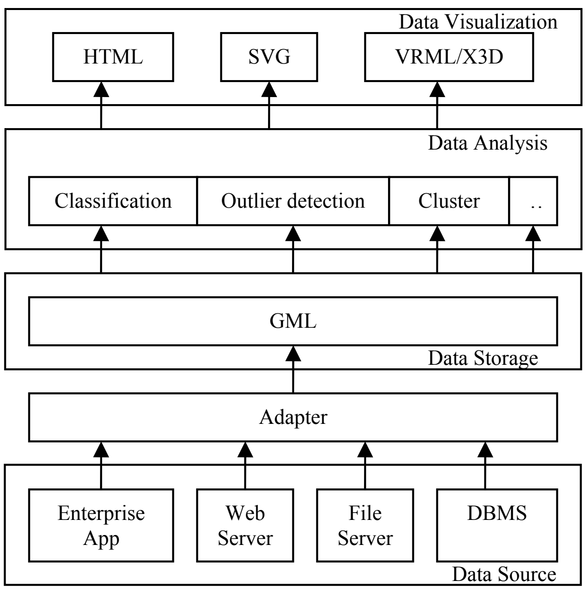

Figure 1. Web-based geo-spatial data visualization framework

Data Source Layer: The main challenge of an geo-spatial data visualization tool is to integrate and present multiple data sources and types in a comprehensible fashion, so that user can focus on extracting meaning from data without being required to explicitly manage the heterogeneous data. The data sources include enterprise information systems, file system servers, relational or object-oriented database systems, etc.

Data Storage Layer: Data from different sources have different formats. Different adapters are applied to them to convert them into a global and consistent format. Extensible Markup Language (XML) provides a common medium of data description and display for diverse systems to understand each other. Geography Markup Language (GML) has been defined as an XML encoding for geographic information. GML has been designed to be used as a mechanism for information discovery, retrieval and exchange [2]. This encoding helps in the storage, exchange and modeling of geographic information containing both spatial and non-spatial attributes of data.

Data Analysis Layer: Geo-spatial data is retrieved from the data storage layer and can be further analyzed based on user requirements and using data mining techniques.

Visualization Layer: The analysis results are presented in various visual forms. The data described and stored in GML can be extracted and styled to suitable graphical representation.

2.1. Geographic data representation using GML

GML is an XML encoding for storing geographic information containing both spatial and non-spatial attributes of data. It has been developed by the OpenGIS as medium of uniform geographic data storage and exchange among diverse applications, especially in a wide-area Internet context [2]. Since spatial databases are also used to store geographical information, the concepts of GML can be used for the storage and exchange of spatial databases. A spatial database system has been defined as a database system supporting spatial data types. Spatial data type objects not only have a non-spatial description such as name and population but also have spatial attributes associated with them such as location, geometry and neighborhood properties- distance and borders etc. A spatial database system has to provide various functionalities including- input, storage, retrieval, selection, analysis and display of the information [3]. Although these features are provided by traditional databases, such databases do not store information in a uniform format. GML representation of information is unique but the use of information can be different and meaning can vary depending on the context. This makes the data very flexible and portable.

2.1. Visualization using SVG and VRML/X3D

It is easy to apply style sheet (XSLT) to the geographical content (GML) and generate 2D (SVG) and 3D (VRML/X3D) geographical presentation. We use XSLT to map the road traffic data to the corresponding region in the virtual world. Mapping data to a 3D virtual world requires converting 2D sensor coordinates into 3D coordinates in the virtual world.

3. Web Visualization of Geo-spatial Data

In this section, we will discuss in detail how to apply this framework to build a web-based visualization tool for geo-spatial data. Specifically, we aim to build an interactive traffic visualization system. It explores various possibilities and ways to represent traffic data in a virtual world with a goal to help both expert and non-expert users to understand traffic data more intuitively.

3.1. Data integration

The main challenge of an interactive 3D visualization tool is to integrate and present multiple data source and types in a comprehensible fashion, so that user can focus on extracting meaning from data without being required to explicitly manage the heterogeneous data. For traffic visualization, typically two types of data sources are needed: map data and traffic data.

To provide an intuitive and easy to use virtual world, it is helpful for user to have an environment built based on geographical data because of the geo-spatial characteristics of traffic information. Such an environment can be constructed using publicly available data, for example, US Census Tiger system, United State Geological Survey (USGS) web site and its Earth Resources Observations Systems (EROS) Data Center. From these websites, we can get a wealth of information that can be used to create and customize the virtual world. Some of the data layers include country/state/county boundary, road atlas, road labels, elevation data, etc.

Road traffic information is provided by the sensor network that may include thousands of stations, each of which contains several loop detectors, depending on the number of lanes. Each road is represented as a sequence of stations. Each station has several sensors, each monitoring one lane at that section of the road. The station is described by its identification number and the coordinate points of the corresponding polygon vertices. Each polygon is specified by a set of vertices and each vertex is specified by longitude and latitude coordinates. These data are converted into GML format for further processing.

3.2. Data analysis

The subjects of analysis in a multidimensional data model are a set of numeric measures and each of them is determined by a set of dimensions. For traffic data, the measures would be volume and occupancy, and the dimensions be time and space. Both dimensions have hierarchical element as time can be grouped by week, month, season, or year, and space by station, county, freeway, or region. CubeView [1] provides a cube view as a specialization of general data cube operator as shown in Figure 2. In this cube view, S means Station, TTD means Time of Day, TDW means Day of Week, and TMY means Month of Year. Each node is a data cube operation and a view of the data. Because of the nature of data cube, this cube view can analyze any traffic data [1].

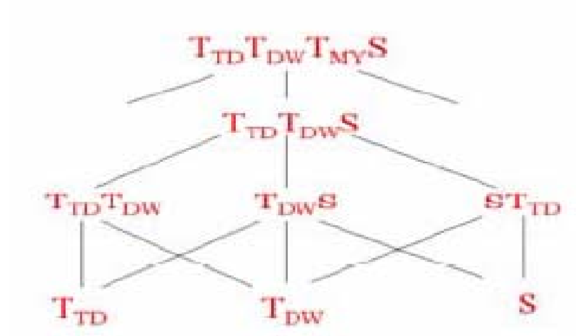

Figure 2. Cube view for traffic data [1]

3.2. Visualization and integration

Data visualization is important for complex data analysis. In complex data, the structure, features, patterns, trends, anomalies and relationships are not easily detectable. Visualization supports the extraction of interests by presenting the data in various visual forms with different interactions. Visualization not only provides a qualitative overview of large and complex data sets, it also can assist in identifying regions of interest and parameters for more focused quantitative analysis.

We create SVG or X3D graphics based on the characteristics of queried data by cube operations. We show maps in 3D and 2D. The 2D map gives user a big picture and overview of the whole world, while 3D virtual world present geographical objects in more detail and with a more realistic way. Both 2D and 3D environment can be presented at different levels of details or at different scales, giving user a choice of the granularity of presented data. In most of the cases, the value of the selected data type (e.g. traffic volume) is used to create a visual cue in the 3D virtual world. The visual cues include both geometry and appearance. The preferred choice in this case is the vertical extrusion of the polygon specifying the region of the sensor. The amount of the extrusion is proportional to the value of the sensor data. In addition, a color coding is used to indicate different levels of data. The time dimension of the data is presented by animating the visual cues with appropriate scaling.

Due to the spatial and temporal characteristics of the data, several selection mechanisms are available for interaction. A user uses a ”traditional” web form to select the data type for a given road or a set of roads, as well as the time interval of interest. The user can also “pick” in the 2D and 3D virtual world a region of interest to select sensor data mapped to that region.

4. Case Study

A traffic data visualization system based on the framework is used to visualize traffic situation for Twin Cities metropolitan. The system uses map data and traffic data from the Traffic Management Center of Minnesota Department of Transportation. The sensor network monitors the occupancy and volume of traffic on the road. Figure 3 shows the result of visualization.

The left panel is a 3D virtual world in X3D, and the right panel is a 2D map in SVG. The 2D map provides an overview of the traffic network in twin cities. It also provides navigation for the 3D virtual world. If the user moves a mouse in the map, the viewpoint in the 3D virtual world will change accordingly. For example, the user moves the mouse over the icon I-394 in the 2D map while in the 3D virtual world the user’s viewpoint moves close to the highway I-394. The user can also use the mouse or a key to freely control the 3D viewpoint. The navigation facilities include zoom in, zoom out, walk, fly, pan, rotate, etc.

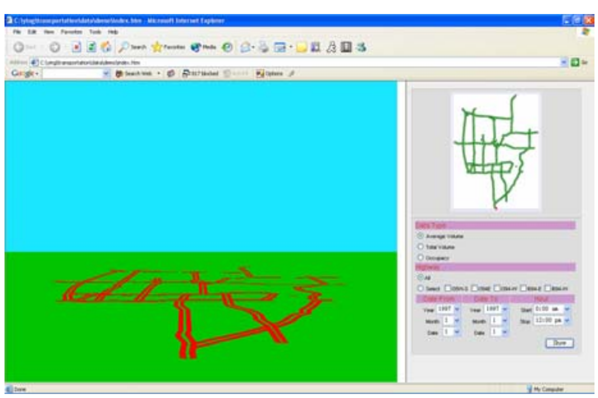

Figure 3. Twin cities highway map

User can choose date to analyze the traffic. Figure 4 shows a snapshot of dynamic traffic volume from 0:00AM to 12:00PM in 01/01/1997.

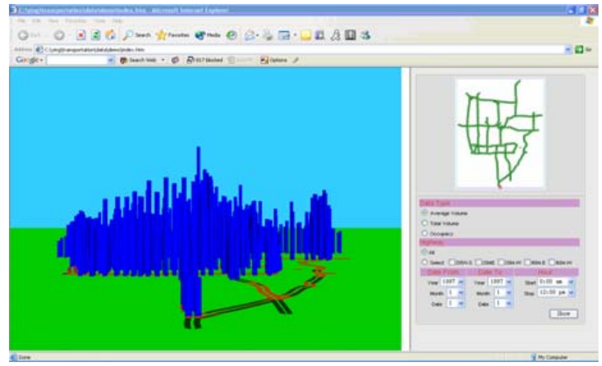

Figure 4. Traffic volume map

5. Conclusion

This paper introduces a general framework for effective integration and presentation of complex heterogeneous spatial-temporal data sets on web. This framework can be used to facilitate the interpretation, exploration and analysis of large volume of data with significant geo-spatial and temporal characteristics. Advantages of this approach include improved visualization of geo-spatial and temporal raw data; better navigation and selection of data; and intuitive user interface. Current work focuses on analyzing and reviewing performances of the system for very large data sets as well as on better integration with various web based data sources. Future work will focus on data analysis and user interface performance evaluation. Data mining and data warehousing techniques will be used for complex data analysis.

6. References

[1] S. Shekhar, C. Lu, P. Zhang, and R. Liu, “Data Mining for Selective Visualization of Large Spatial Datasets”, *Proceedings of the 14th IEEE International Conference on Tools with Artificial Intelligence*, 2002.

[2] Z.R. Peng and M.-H. Tsou, *Internet GIS Distributed Geographic information Services for the Internet and Wireless Networks*, John Wiley & Sons, Inc., 2003.

[3] P. Rigaux, M. Scholl, and A. Voisard, *Spatial Databases with Application to GIS*, Morgan Kaufmann Publishers, London NW1 7BY, UK, 2002.

[4] J. Boye, “SVG Brings Fast Vector Graphics to Web”, http://tech.irt.org/articles/js176/#2.

[5] A. E. Walsh and M. Bourges-Sévenier, *Core Web3D*, Prentice Hall PTR, Upper Saddle River, NJ, USA, 2001.