British Journal of Applied Science & Technology
17(1): 1-6, 2016, Article no.BJAST.27685
ISSN: 2231-0843, NLM ID: 101664541

SCIENCEDOMAIN *international*
www.sciencedomain.org

Visualization of Traffic Data Using View Fusion Techniques

Ren Li Jing[1*] and Zhao Zheng Xu[1]

[1]*School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang 054300, China.*

Authors' contributions

This work was carried out in collaboration between both authors. Author RLJ conceived and designed the work that led to the submission and wrote the first draft of the manuscript. Author ZZX managed the analyses of the study and literature searches. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJAST/2016/27685
Editor(s):
(1) Wei Wu, Applied Mathematics Department, Dalian University of Technology, China.
Reviewers:
(1) Yusuf Lateef Oladimeji, Federal University of Agriculture, Abeokuta, Nigeria.
(2) K. Jayakumar, Kamaraj College of Engineering and Technology, Tamilnadu, India.
(3) Amaia Aguirregoitia Martínez, Universidad del País Vasco, Spain.
Complete Peer review History: http://www.sciencedomain.org/review-history/15653

Original Research Article

Received 14th June 2016
Accepted 28th July 2016
Published 4th August 2016

ABSTRACT

The traditional data visualization technology concentrates on the chart analysis and the summary of calculation analysis, and ignore the spatial distribution characteristics of statistical information. Currently, data visualization technology and its applications suffer from the issues such as low-accurate visualization, limited representation solutions and blurred inherent information. In order to overcome the shortages caused by these problems, we propose a mixed visual analysis scheme combined with geographic information data to better display spatial statistical data. The proposed system combines a variety of visualization techniques, such as geographic maps, bubble chart, and Treemap visualization method to enhance the visual results and ensure that users can observe various data indicators from different angles. In order to further analyze the data of interest, we design and achieve a variety of visualization techniques coordinated with multiple views. The experimental results show that the proposed fusion visualization techniques can help users to discover the regularity of statistical data more easily and offer a fast and convenient visualization tool for statistical data analysis.

**Corresponding author: E-mail: 310368284@qq.com;*

Keywords: Information visualization; view fusion; Treemap; 3d maps; bubble chart.

1. INTRODUCTION

In recent years, the dramatic growth in the processing power and storage capacity of computers has been witnessed and information that people need to process increases accordingly. According to statistics [1], there are currently 4 billion shallow web pages and 50 billion deep web pages. These webpages comprise of up to 220 languages and text information. Digital information resources have become mankind's greatest resources. How to collect, analyze and find rules behind hidden data and its information pattern has become a major issue in digital age. Information visualization can solve this problem effectively by helping us to understand the relationships between large and complex data and how to visualize them. Compared with other senses, people can obtain more information through vision. The research shows that the method of processing information through visualization techniques features high-speed processing and a large memory capacity in parallel [2], which also implies that interpreting the massive data through visualization provides a reasonable approach.

Information visualization makes use of computer to display abstract data interactively, which enhances people's perception of abstract information [3]. It is a visual interface which connects people and information. The theory and methods in scientific visualization, human-computer interaction, data mining, image processing, graphics and many disciplines are combined to present abstract information in an intuitive way thus people can take advantage of visual perception to process information and discover relationships and the hidden patterns between the information [4]. Information visualization is widely used in every walk of life including meteorological and geological field. Information visualization could reflect the state of things truly. Visualization of multidimensional data such as financial data and statistical data is an effective method to understand the massive and complex data. Visualization of network information can dig the value and wealth behind business data [5].

With the emergence of the information age, statistical data has become the most important information source of a country, which is a true reflection of various social conditions. The effective analysis of statistical data has become a challenging job. The statistics data collected by the national statistics departments at all levels is archived in the form of tables and texts through Excel, Access, Oracle and other types of databases, which is not conducive to discover in-depth information and rules [6]. Visualization has become trend to preserve and express information. Combining single visual images as a whole to show statistical data through the horizontal, vertical and comprehensive way has been proved to be an effective way to enhance visual effects and improve visual perception. So users discover the patterns and trends of information data more directly.

2. VIEW FUSION

Large-scale statistical data exists related to geographic distribution, and the method to present and analyze these data has come to be hot issues in the visualization research field. Types of visualization technology are growing nowadays. Fusion visualization which applies multiple views to display data from various angles provides an effective method for visualization and user interaction. In this paper, a framework of fusion visualization is proposed to visualize and analyze statistical data. The system flow chart is shown in Fig. 1.

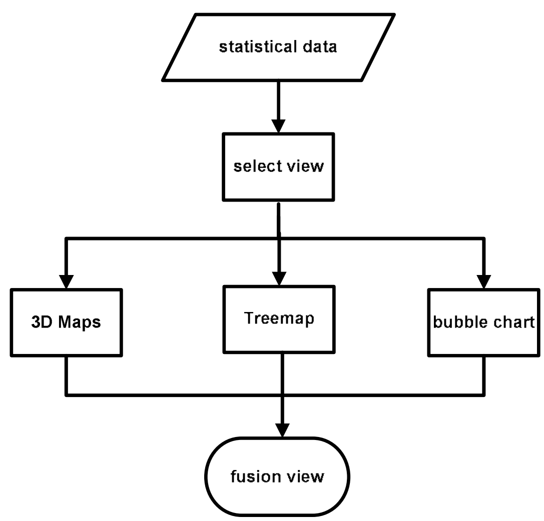

Fig. 1. The framework of view fusion system

The system includes a plurality of view interfaces, that is, 3Dmaps to display space information, Bubble Chart to represent the movement focus over time and Treemap to represent the hierarchical relationship between the two sets of data. Same data structure and color matching scheme are adopted in the system.

Multi-dimensional spatial data can be analyzed in multi angles by view fusion. On the other hand, each view can be used to highlight the characteristics of the data and detect trends and rules of complex data. Users can explore data using a variety of visualization methods through the integration of view fusion techniques.

2.1 3D Maps Method

3D Maps method sets a different color to represent each region based on the information data and display statistical data on the map with a certain proportion. For example, the spatial position of each province can be displayed by the pie chart, bar charts, etc. The pie chart reveals the data distribution relationship, while histogram shows the height corresponding to the data size. Studies have shown that yellow, red, orange and other warm color are very cheerful colors for the live colors, and the luminosity of the sun makes people feel warmth. Meanwhile, the color brightness of warm color is relatively high, which is very attractive and visual sensitive to people [7]. The adjacent color contrast method is adopted in the system to unify vision and enhance expression effect. What is most important is that it can arrest user's attention.

2.2 Treemap Method

Treemap method is the main method of visualizing hierarchical data. It is used in the system to divide the screen space according to the hierarchical structure of the data in a rectangular subspace. Each rectangle can be divided by corresponding color matching and necessary instructions can also be added to the diagram. Different optimization algorithms can be used to get rational distribution in the Treemap [8]. The squarified layout method is chosen to avoid elongated rectangle in the system. The concept of entropy is introduced to represent data distribution. Let V_i be the quantity of sample i. The ratio of V_i can be got as follows:

$$\mathrm{P}_i = \frac{V_i}{\sum_{i=1}^{n} V_i} \tag{1}$$

So entropy H can be formulated as follows:

$$H = -\sum_{i=1}^{n} P_i ln P_i \tag{2}$$

The subspace size is decided by the entropy in order to make use of the display space effectively.

2.3 Bubble Chart Method

The bubble chart is another visualization method which usually used to display and compare data, such as statistical data, engineering data and scientific data. It can help users find outliers, trends among data and relationships. Suitable function can be chosen to fit the data points. As shown in Fig. 2, the horizontal axis represents provinces, and the longitudinal axis and the bubble size represents number of accessed websites for each province in a sample survey. It can be clearly seen from the figure that economically developed provinces with large population density such as Guangdong province and Hainan province experience more-trafficked sites. The developing provinces with sparse population density such as Qinghai province, Ningxia province and Xinjiang Uygur Autonomous Region only consists of a dozen web sites in the survey.

From Fig. 2, it was observed that the number of websites in four municipalities reaches 1,414,697, which accounts for the largest percentage in the country. The color contrast method is applied to compare the number of websites, which shows strongly expressed effect. The axes, grid, border, background is in blue tone, while the data is relatively bright color in order to make data become the focus of attention. The method allows users to better analyze the data and discovers potential data trends.

3. CASE STUDIES

On account of the proposed view fusion framework, a prototype visualization system was implemented in C++ program language on the Windows 7 operating system. A snapshot of the system is shown in Fig. 3. Two case studies are offered to show the efficacy of the proposed framework based on websites traffic sample data from different provinces and cities in China [9]. Sample sites for cities and provinces were firstly selected, and then the traffic of each site was collected in June-September 2015, at last average daily visits was got. It is well known that the economic level of China's coastal provinces such as Guangdong province and Hainan province are more developed than the inland provinces such as Henan province and Gansu province. The system provides a new perspective to analyze the relationship among geography, websites traffic and economic development.

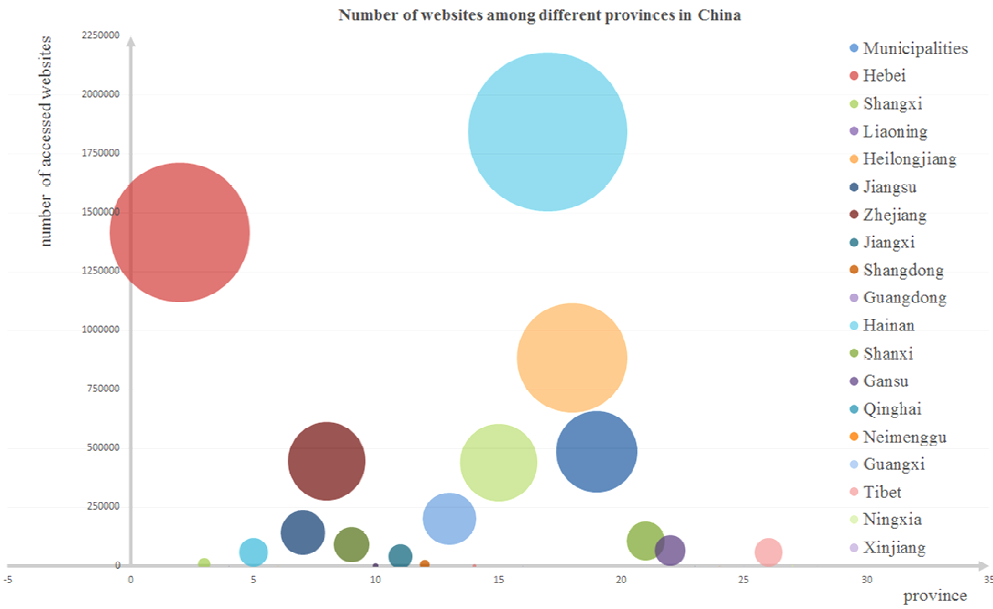

Fig. 2. Number of accessed websites among different provinces in China

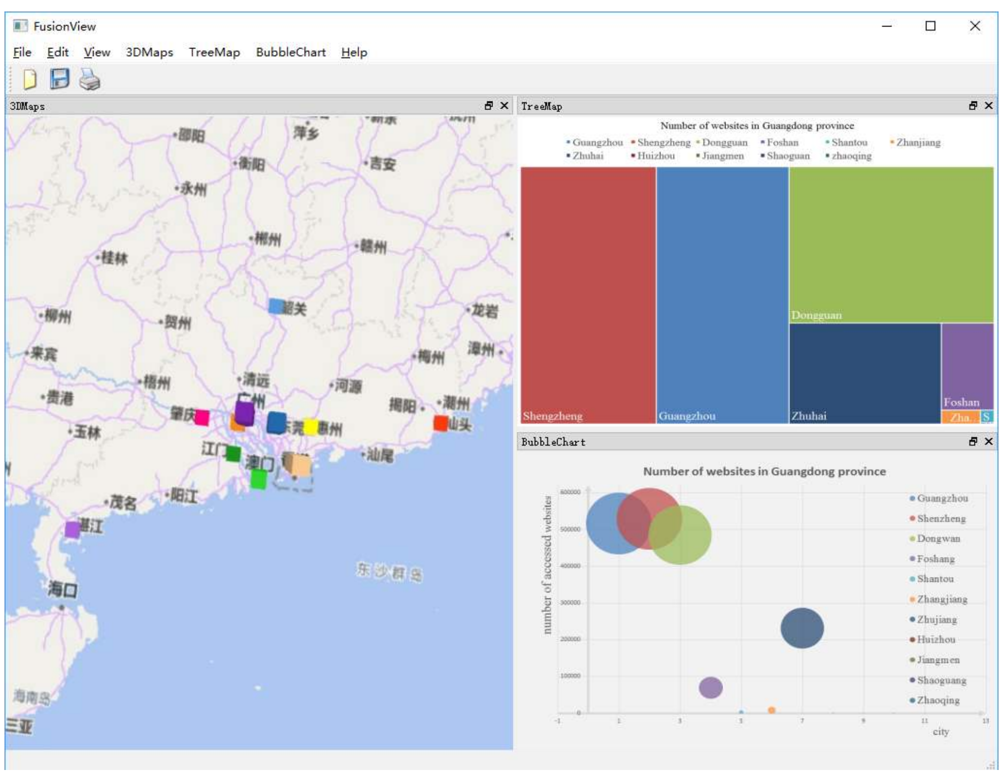

Fig. 3. Guangdong province website samples data visualization

3.1 Case Study I: Sites Traffic in a Province

Fig. 3 shows a sample of websites in Guangdong Province. It can be seen that the view fusion system shows sample data from all angles.

The sample cities of Guangdong Province where transportation is convenient and economic development is rapid are located in coastal. It can be seen from the column height of the 3D maps figure in Fig. 3 there is a big gap between cities in the website construction. The traffic of Guangzhou city and Shenzhen city is significantly more than that of Foshan city, Zhaoqing city and other cities. From the Bubble Chart it can be seen that the number of websites reaches 50 million in developed cities such as Guangzhou city, Shenzhen city and Dongguan city. There are only 10 to 100 sites in less developed cities such as Jiangmen city, Shaoguan city and Zhaoqing city.

3.2 Case Study II: Sites Traffic in Different Provinces

With development of economy, the Internet is playing more and more important role and the traffic is becoming bigger and bigger. Because obvious income difference exists in provinces, the amount of websites and the traffic provide a good measure to understand economic development status in provinces.

As shown in Fig. 4, sites traffic in Guangdong province, Hainan province and other coastal provinces are much higher than that in Henan province and Gansu province in 3D maps. In the bubble chart, the ordinate represents the number of sites and the bubble size represents the number of website visitors. Users can easily find the difference of the number of the websites and sites visitors among provinces and cities, so underlying causes of the difference can be discovered.

The status of the network development in China is just a microcosm of global network development. Although the development of information technology is not subject to geographical restrictions; the number of sites and the number of Internet users showed explosive growth. Different geographical location leads to gaps in the levels of economic development which in turn would lead to the gap among the regional network construction and the number of websites [10]. More attention should be paid to find the underlying cause for the unbalanced development among regions.

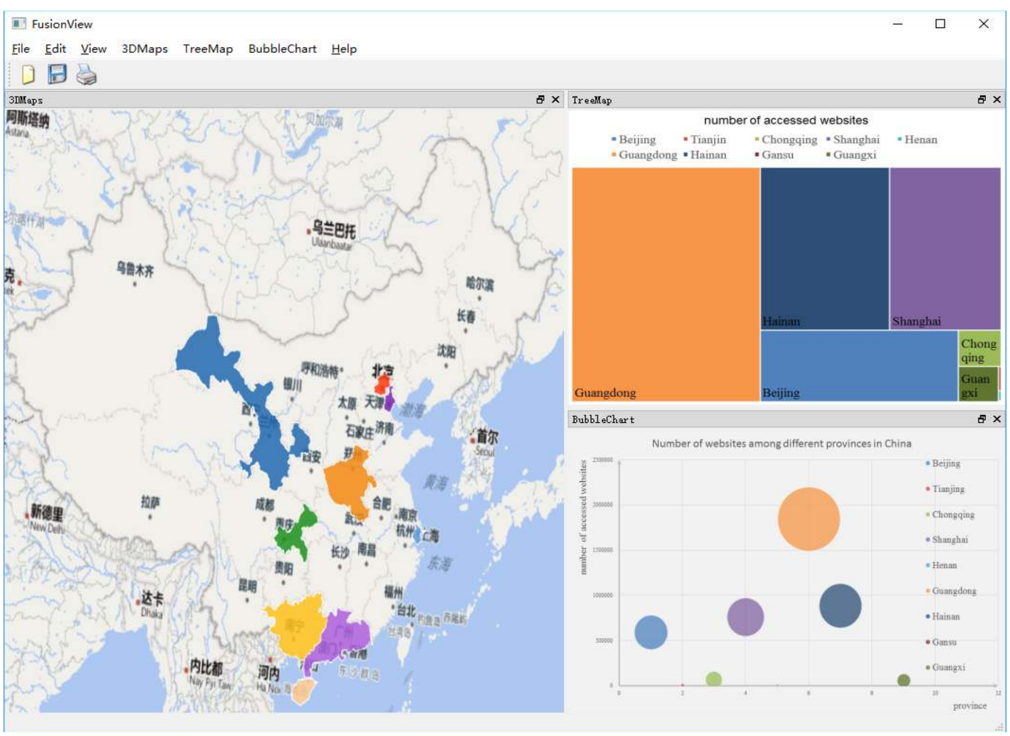

Fig. 4. Multiple provinces website samples data visualization

It can be seen in the case study that the traffic of coastal provinces is larger than the Northwest Territories. Northwest provinces locate in China's inland where science technology, foreign economic and network construction are relatively backward. Since the reform and opening up, the government vigorously supports the northwest, and introduces appropriate policies to promote the development of the Northwest Territories which reduces disparities between Southeast and Northwest region. Although the economy is relatively backward, with the continuous progress of science and technology, great potential exists in the development of network construction in Northwest Territories.

4. CONCLUSION

With the explosive growth of data, information visualization technology will be more challenging. Previous works focus on new visualization methods. The significant of our work is the integration of a host of visualization techniques to visualize processed information. A fusion visualization scheme is also proposed to display massive statistics. The scheme integrates 3D maps, bubble charts and Treemap methods to assist users in understanding the various attributes of data from different angles. The plug-in mechanism will be used for the development of the system in the future. So new visualization methods can be added and data can be presented by the most suitable visualization method. On the other hand, the visualization scheme is limited by the hardware device. Our future work will examine the display of massive, complex data on existing visualization environment.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Barlowe S, Liu Y, Yang J, Livesay DR, Jacobs DJ, Mottonen J, et al. Wavemap: Interactively discovering features from protein flexibility matrices using wavelet-based visual analytics. Computer Graphics Forum. 2011;1001–10.
2. Telea AC. Data visualization: Principles and practice, Second Edition; 2014.
3. Ward MO, Grinstein G, Keim D. Interactive data visualization: Foundations, techniques and applications; 2010.
4. Chen IX, Yang CZ. Visualization of social networks. Handbook of Social Network Technologies and Applications. 2010;585–610.
5. Zhao Z, Zhao LZ. Small-world phenomenon: Toward an analytical model for data exchange in product lifecycle management. International Journal of Internet Manufacturing and Services. 2008;213–30.
6. Elliot AJ, Maier MA. Color psychology: Effects of perceiving color on psychological functioning in humans. Annual Review of Psychology. 2014;95–120.
7. Hugine AL, Guerlain SA, Turrentine FE. Visualizing surgical quality data with treemaps. Journal of Surgical Research. 2014;74–83.
8. Cui W, Liu S, Tan L, Shi C, Song Y, Gao Z, et al. Text flow: Towards better understanding of evolving topics in text. IEEE Transactions on Visualization and Computer Graphics. 2011;2412–21.
9. Schreck T, Keim D. Visual analysis of social media data. Computer. 2013;68–75.
10. Thom D, Bosch H, Koch S, Wörner M, Ertl T. Spatiotemporal anomaly detection through visual analysis of geolocated Twitter messages. Visualization Symposium (PacificVis), 2012 IEEE Pacific. 2012;41–8.

Peer-review history:
The peer review history for this paper can be accessed here:
http://sciencedomain.org/review-history/15653