Multivariate Visualization Techniques for exploring Traffic Data

1. Introduction

Visualizing data has proven to be an important part of the data mining process. Data visualization techniques have been promoted with the increasing data stream. In the field of traffic data visualization, many researches are focusing on presenting spatial and temporal attributes of datasets. There are multiple variables contained in traffic data. Wei Chen et al. classified attributes expecting spatial and time attribute of traffic data into three types (Numerical, categorical and textual) [1]. How to integrate other attributes into the visualization workspace and make good use of multiple properties also deserves to be discussed.

As one type of information visualization, multivariate data visualization facilitates users in understanding multidimensional data [2]. Scientists attempt to classify multivariate visualization techniques differently. In 1996, Keim and Kriegel divided multivariate visualization techniques into six categories, namely, geometric projection, icon-based, pixel-oriented, hierarchical, graph-based and hybrid [3]. M.C Ferreira adopted namely and hybrid parts and grouped them into four classes [6].

This literature review summarizes five most frequently used multivariate visualization techniques (parallel coordinates, pixel-oriented technique, data glyphs, tree map and hybrid technique) for traffic data and demonstrates their capabilities using one or two examples.

2 Parallel coordinates

Parallel coordinates is a widely used technique which is famous for revealing correlations among multiple attributes. In 2009, Michael L et al. introduced a comprehensive visual analytic tool (ICE) for transportation incident data [4]. In ICE, they implemented a parallel coordinates plot (Figure 1) to visualize the multivariate of transportation incidents. In this plot, vertical axes present different variables and each line stands for one incident. The density of lines between two axes indicates where incidents frequently occur. Users could find certain value range where incidents always happen. Besides, ICE add filter function on the PCP to which users can select interesting values along an axis as well.

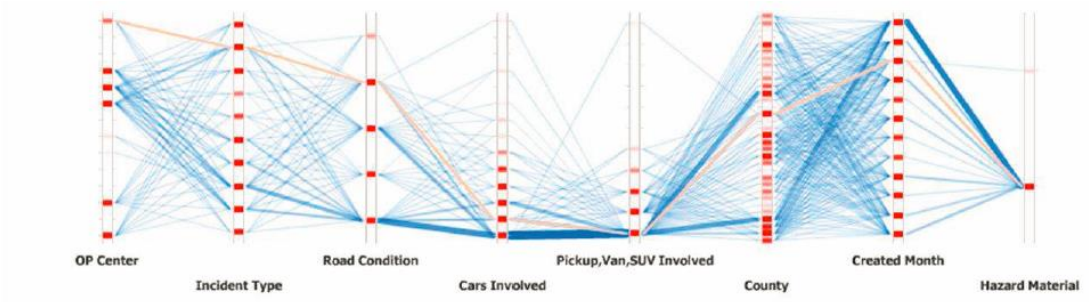

Figure 1

In 2014, Mike Barry and Brian Card used PCP to visualize approximately 1150 train trips

on one day in Boston[5]. As Figure 2 shows, vertical axes with time values present train stations. Colored curves connecting each station at different time depict detailed train schedule in Boston. On the plot, users can track the train route, point out rash hour for trains and the degree of connectivity between different stations.

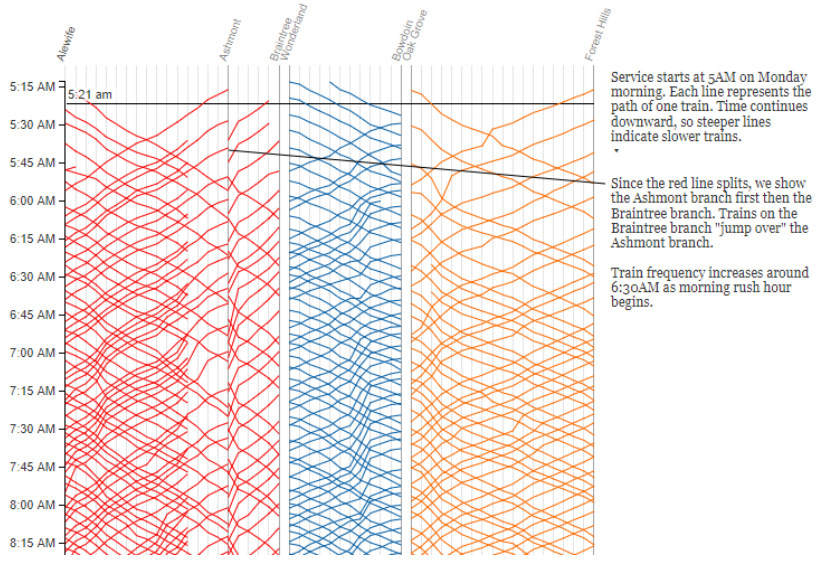

Figure 2

3 Pixel-oriented technique

The idea of pixel-oriented techniques is using colored pixels to present attribute values. As each data stored in one pixel uniquely, it helps to avoid visual cluttering and overlapping[6]. Zuchao Wand et al. applied pixel table (Figure 3) to overview the traffic speed data for whole year in 2012[7]. In the table, pixels stand for different geographical cells, speed value is encoded by different color and each row records daily data. This table could be extended by time and conveniences users to grasp general concept of traffic speed distribution in spatial-temporal pattern.

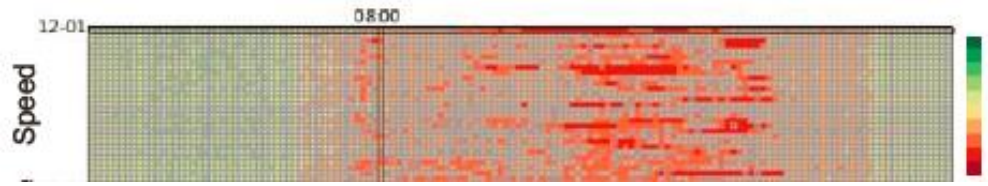

Figure 3

4 Data Glyphs

Since the composition of multi visual variables, glyph-based visualization technique is useful at encoding high-dimension data [8]. Furthermore, it provides a wide design space for different application scenarios. Like in T-Watcher (Jiangsu Pu et al.), scientists designed “Cell Glyph” to visualize the historical and the instant vehicle speed data [12]. As Figure 5 shows, in the cell glyph, the tail illustrates changes of vehicle speed in a certain period. Meanwhile, the body color and shape indicate the instant speed quantity. The core in the body helps to compare the instant speed with the mean speed.

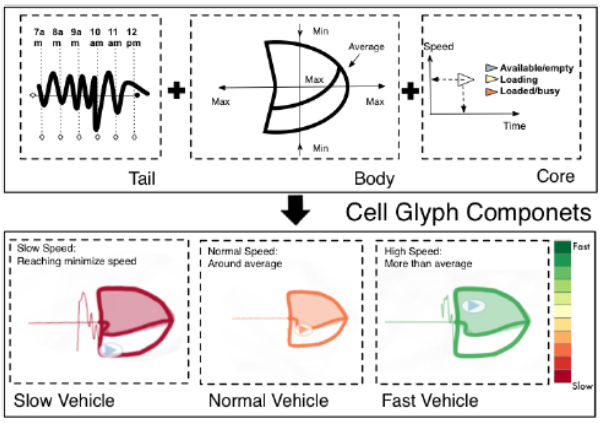

Figure 4

5 Tree Map

Tree map aims to show hierarchical data by subdividing the display space according to attribute values. This technique makes full use of visualizing space and is suitable for obtaining ordinal attributes [9]. In 2010, Aidan Slingsby et al. used tree map to show different traffic variables (vehicle type, average speed and time) in one limited rectangle [10]. As Figure 5 shows, the largest rectangle presents the research area. From A to C, there are 3 levels for the tree map. In the first level, the largest rectangle is proportionated into five parts according to the vehicle quantity. Then, they are segmented into seven parts by the day of a week. Further, rectangles created in the second level are partitioned into twenty-four parts by the hour of a day. Meanwhile, color indicates the average vehicle speed.

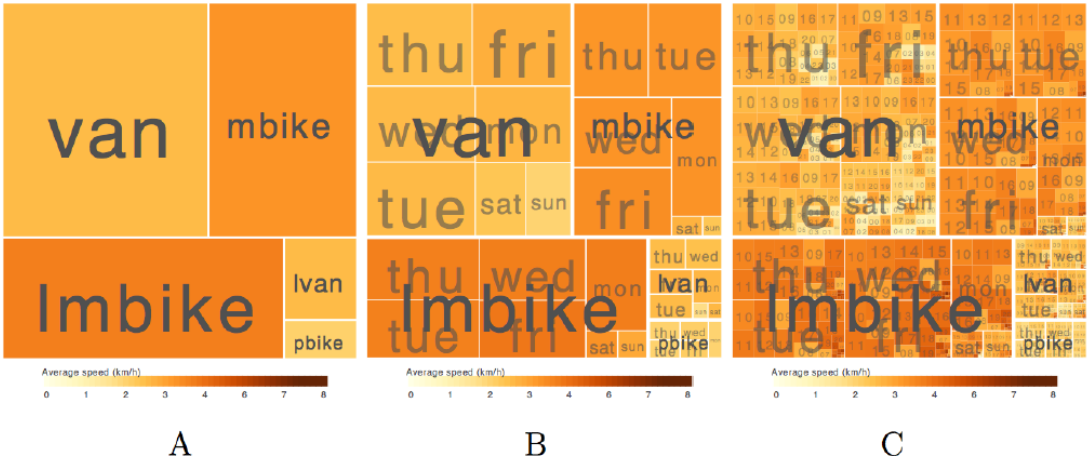

Figure 5

6 Hybrid technique

In some situations, single visualization technique is limited to present complex dataset comprehensively. Hybrid technique aims to integrate multi methods together to enhance the expressiveness of visualization. Based on this concept, Christian Tominski et al. combined 2D map and 3D pixel band together to visualize trajectory data in 2012(Figure 6)[13]. The 2D map provides spatial reference. pixel bands containing speed information

for each trajectory are sequenced by time series. This innovated visualization technique shows more data components in the limited workspace and helps users to understand spatial-temporal trajectory better.

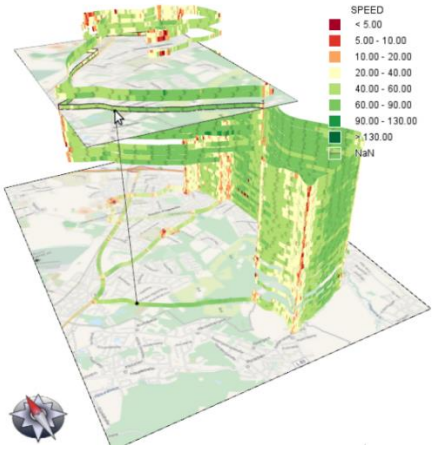

Figure 6

7 Conclusion

In this survey, we reviewed several works on utilizing multivariate visualization techniques to explore traffic data. Benefits of using them on interpreting valuable data patterns and uncovering hidden information are obvious. Like, users can detect outlies of datasets on the parallel coordinates plot and point out different clusters on the pixel-oriented chart easily. Multivariate visualization techniques help to visualize more data dimensions in the screen and have great flexibility to be designed.

However, there are also some limitations for the single visualization technique. For example, the parallel coordinates plot is incapable of presenting categorical properties and textual properties. Data Glyph is only suitable for depicting textual data. Meanwhile, it is hard for users to detect data outliers and cluster center using tree map. Due to the multiple properties of traffic data, composing various visualization techniques is necessary as well.

References

[1] W. Chen, F. Guo and F. Wang, "A Survey of Traffic Data Visualization," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 2970-2984, Dec. 2015.

[2] W Y Chan, W. (2006). A Survey on Multivariate Data Visualization.

[3] J. Pu, S. Liu, Y. Ding, H. Qu and L. Ni, "T-Watcher: A New Visual Analytic System for Effective Traffic Surveillance," *2013 IEEE 14th International Conference on Mobile Data Management*, Milan, 2013, pp. 127-136.

[4] M. L. Pack, K. Wongsuphasawat, M. VanDaniker and D. Filippova, "ICE--visual analytics for transportation incident datasets," *2009 IEEE International Conference on Information Reuse & Integration*, Las Vegas, NV, 2009, pp. 200-205.

[5] M. Barry and B. Card, “Boston’s massachusetts bay transit authority(mbta)”

[6] M. C. Ferreira de Oliveira and H. Levkowitz, "From visual data exploration to visual data mining: a survey," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 9, no. 3, pp. 378-394, July-Sept. 2003.

[7] Z. Wang *et al*., "Visual Exploration of Sparse Traffic Trajectory Data," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 1813-1822, 31 Dec. 2014.
[8] Fuchs, von Johannes Hermann. “Glyph Design for Temporal and Multi-Dimensional Data : Design Considerations and Evaluation.” (2016).
[9] KEIM, Daniel A., 1997. *Visual Techniques for Exploring Databases*. *Knowledge Discovery in Databases (KDD'97)*. Newport Beach, CA, USA, Nov, 1997. In: *Intern. Conference on Knowledge Discovery in Databases (KDD'97), Newport Beach, CA, USA, November, 1997*
[10] Slingsby, A., Wood, J. and Dykes, J. (2010). Treemap Cartography for showing Spatial and Temporal Traffic Patterns. Journal of Maps, pp. 135-146.
[11] D. A. Keim and H. -. Kriegel, "Visualization techniques for mining large databases: a comparison," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 8, no. 6, pp. 923-938, Dec. 1996.
[12] A. Rusu, C. Santiago, A. Crowell and E. Thomas, "Enhanced Star Glyphs for Multiple-Source Data Analysis," *2009 13th International Conference Information Visualisation*, Barcelona, 2009, pp. 183-190.
[13] C. Tominski, H. Schumann, G. Andrienko and N. Andrienko, "Stacking-Based Visualization of Trajectory Attribute Data," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2565-2574, Dec. 2012.