A Survey of Traffic Data Visualization

Wei Chen, Fangzhou Guo, and Fei-Yue Wang, *Fellow, IEEE*

***Abstract*—Data-driven intelligent transportation systems utilize data resources generated within intelligent systems to improve the performance of transportation systems and provide convenient and reliable services. Traffic data refer to datasets generated and collected on moving vehicles and objects. Data visualization is an efficient means to represent distributions and structures of datasets and reveal hidden patterns in the data. This paper introduces the basic concept and pipeline of traffic data visualization, provides an overview of related data processing techniques, and summarizes existing methods for depicting the temporal, spatial, numerical, and categorical properties of traffic data.**

***Index Terms*—Traffic, traffic data visualization, visual analysis, data-driven intelligent transportation system.**

I. Motivation

TRAFFIC is the flux or passage of motorized vehicles, non-motorized vehicles, and pedestrians on the road, or the movement of passengers (e.g., metro interchanges) [1]. Traffic can take place in urban regions, lands, seas, air, or even underground. With the rapid development of transportation systems, traffic has become a vital part of human life and significantly influenced the quality of life. For instance, an estimated average of 40% of the population spends at least 1 hour on the road every day [2].

In modern cities, massive population and large number of vehicles cause problems, like congestion, accidents and air pollution. A number of efforts to address these problems have been proven effective, including intelligent transportation systems (ITSs), public transportation systems, safety seat belts, and air-bags. However, the ever-increasing number of private cars greatly neutralizes the achievement of traffic regulation and control. Among these solutions, ITSs are deemed attractive because they enhance the efficiency and functionalities of transportation systems with advanced information technology [3]. In particular, the role of data in ITS has become essential because of the size of the collected data. The data contain information and can also be used to generate new functions and services in ITS [2]. A data-driven ITS allows users to interactively utilize data resources that pertain to transportation systems, as well as access and employ data through more convenient and reliable services to improve the performance of transportation systems [2].

Data visualization employs visual channels to represent datasets [30], transforming various types of data into appropriate visual representations, so that data understanding and analysis can be completed efficiently. The advantage of data visualization is its incorporation of human capabilities into an intuitive visual interface, thereby combining machine intelligence with human intelligence. Scientific visualization, information visualization and visual analytics are three major fields in data visualization. Scientific visualization illustrates structures and evolutions of physical or chemical properties in the spatial domain. Information visualization focuses on the representation of abstract, unstructured, and high-dimensional data, including business data, social networks data, and textual data, among others. [31]. Iterative, interactive and dynamic integration of human intelligence and data analysis establishes a novel analysis strategy, namely, visual analytics [32]. Traffic datasets are generally high-dimensional or spatial-temporal; thus visualizing traffic data mostly employs information visualization and visual analytics.

Visualization and visual analysis are important for a highly efficient data-driven ITS. Specifically, traffic data visualization can facilitate understanding of the behavior of moving objects (vehicles) and discovery of traffic, social, geo-spatial, and even economic patterns. In general, an analytic system consists of four main components: data collection, data preprocessing, data query and data analysis. Each component requires specialized visualization techniques. For instance, visual data cleaning can help the user transform data to become usable for downstream analysis tasks [33]. Other processes such as aggregation and clustering can also be enhanced with a visual interface [25], [34]. A user-friendly query interface is required to retrieve the needed data [14]. Furthermore, traffic situation monitoring and traffic pattern recognition are widely studied for the purposes of intelligent control and analysis [11], [35]. Considering the tasks of existing traffic data analysis applications, the tasks of traffic data visualization could be classified as follows:

- **Visual monitoring of traffic situations** Interesting events may be hidden in traffic data such as traffic jams. With real-time monitored data (e.g., video surveillance in tunnels or cross-intersections), live traffic situations

Manuscript received September 11, 2014; revised February 25, 2015 and May 14, 2015; accepted May 18, 2015. Date of publication June 10, 2015; date of current version November 23, 2015. This work was supported in part by the Major Program of the National Natural Science Foundation of China under Grant 61232012, by the National Natural Science Foundation of China under Grant 61422211, by the National 973 Program of China under Grant 2015CB352503, by the Zhejiang Provincial Natural Science Foundation of China under Grant LR13F020001, by the Fundamental Research Funds for the Central Universities, and by the NUS–ZJU SeSama Center. The Associate Editor for this paper was J. Zhang.

W. Chen is with the State Key Laboratory of Computer Aided Design and Computer Graphics, Zhejiang University, Hangzhou 310058, China, and also with Cyber Innovation Center, Zhejiang University, Hangzhou 310027, China (e-mail: chenwei@cad.zju.edu.cn).

F. Guo is with the State Key Laboratory of Computer Aided Design and Computer Graphics, Zhejiang University, Hangzhou 310058, China (e-mail: 11321016@zju.edu.cn).

F.-Y. Wang is with the Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: feiyue@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2015.2436897

1524-9050 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

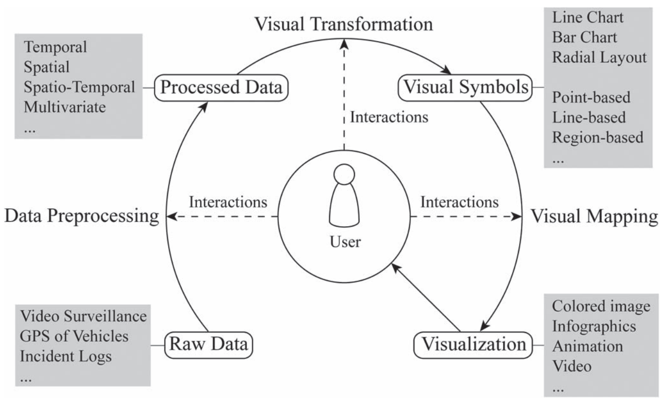


Fig. 1. Conceptual pipeline of traffic data visualization.

and comprehensive events can be observed and tracked to understand the cause and mechanism of a traffic jam along a long road. An example is the visual analysis system [11] based on taxi trajectories.

- **Pattern discovery and clustering** An important goal of traffic data visualization is to discover the mobility patterns of objects and cluster these patterns. These patterns reflect the characteristics of individual movements, their evolutions as well as their associations with other parameters. For instance, Gennady Andrienko *et al.* [36] presented an interactive visual clustering method to classify trajectories of massive objects.
- **Situation-aware exploration and prediction** Data analysis tasks can be categorized into two classes: description and prediction. Many analytic systems are capable of exploring and explaining traffic situations, for example, visually querying taxi trips in a city [14] and predicting the trajectories of massive cars in a city [37].
- **Route planning and recommendation** Traffic regulations and route recommendations are essential components of ITS. Data-driven control and planning have been proven efficient in achieving satisfactory results. Incorporating human capabilities into the analysis process can further improve efficiency, as demonstrated in the visualization-assisted route recommendation system [10].

A traffic data visualization system typically contains four data states and three process stages, as shown in the representation of a general visualization pipeline [38] in Fig. 1. The data flow includes four states, namely, raw data, processed data, visual symbols and visualization. Raw data may be collected from different data sources including video surveillance, GPS of vehicles, and incident logs. Preprocessed data contain temporal, spatial, spatio-temporal, and multivariate properties. Thereafter, visual transformation is performed to convert data into appropriately designed and placed visual symbols, such as line chart, bar chart, dot chart, and star plots, among others. Finally, the visual symbols and metaphors are mapped with various visual channels (color, transparency, texture, etc.) and composed into various visualization forms, such as colored image, infographics, animation, and video. The user is allowed to modulate the parameters in each stage using the user interface. In the data preprocessing stage, the user can optimize the preprocessing functions by iteratively adjusting the parameters. In the visual transformation stage, the user can transform and filter the data to determine the shown data. In the visual mapping stage, the user can manipulate visual mapping types and interact with visual symbols. In this way, the patterns and knowledge hidden in the data can be easily understood and discovered.

The remainder of the text is organized as follows: Details of traffic data and preprocessing techniques are presented in Section II. Section III elaborates on the various visualization techniques in terms of time, locations, and other aggregated or deducted variables. Section IV shows how visualization can be combined with analysis techniques to enhance understanding and mining of traffic data. Finally, this survey is concluded and future works are highlight in Section V.

II. Traffic Data Preprocess

Different types of traffic data demand different visualization and analysis methods. Real-captured data are typically raw, erroneous, and contain uncertainties, outliers, missing values, or mismatched items. Raw data must be processed for visualization and analysis.

A. *Traffic Data*

Traffic data refer to the datasets generated and collected by sensors in traffic vehicles or monitors installed along the roads. Examples of traffic data include GPS data of vehicles, GSM locations or cell station records of human mobility, and video/image/counting records of surveillance devices. The working modes of sensors can be roughly categorized into the following three classes [39], [40]:

- **Location-Based**. The location of an object is recorded upon entering the sensor range. For instance, in a cross-intersection, a video monitoring device captures the location and movement orientation of a pedestrian if and only if he/she passes by the monitor.
- **Activity-Based**. When an object carries out a certain activity, related or derived information is recorded. For instance, the location of a GSM user is automatically recorded when he/she makes a call.
- **Device-based**. A device carried by an object actively records and sends back positional and other information. For instance, a taxi with a GPS device delivers a signal to the data center every 20 seconds.

Trajectory is the most common form of traffic data. A trajectory contains temporal information, which records the timeline of movement, and spatial information, which records positions at each time point. Previous studies have focuses on the visualization and analysis of trajectories, including [4], [7], [13], [18], [23], [41]–[43]. Other types of information accompanying trajectories can also be utilized [39], including movement directions, change of direction [23], movement speed [11], and change of speed [13].

Incident logs are based on events and contain such attributes as event type, event location and other information on related entities.

TABLE I
EXAMPLES OF TRAFFIC DATA. HERE, N, C, AND T STAND FOR NUMERICAL, CATEGORICAL, AND TEXTUAL, RESPECTIVELY

	Data	Properties	Data Types			Representative Datasets
			N	C	T	
Trajectory	Shipping trajectories	Time	✓			Vessel traffic data [4]
		Location	✓			
		Ship type		✓		
		Destination			✓	
		Velocity	✓			
	Aircraft trajectories	Location	✓			Flight in France [5], Europe 24 [6]
		Flight level	✓			
		Time	✓			
		Velocity	✓			
		Aircraft ID			✓	
	Automobile trajectories	Time	✓			Taxi GPS data of Beijing [7], [8], Shenzhen [9], Shanghai [10], [11], [12], San Francisco [13], New York City [14], Wuhan [15], [16], and Sweden [17]; Traffic monitoring cells data in Nanjing [18]; GPS data in Louisiana [19]
		Location	✓			
		Direction		✓		
		Change of direction		✓		
		Velocity	✓			
		Acceleration	✓			
		Pick-up/drop-off		✓		
	Train/Metro trajectories	Location	✓			Train data in France [20], Boston's metro data [21]
		Time	✓			
		Station			✓	
	Pedestrian trajectories	Location	✓			Human mobility traces [22]
		Time	✓			
		Velocity	✓			
	Mixed trajectories	Object type		✓		Intersection count [23]
		Position	✓			
		Velocity	✓			
		Direction		✓		
		Time	✓			
Incident	Tunnel incident	Stateful events		✓		Incident detection system (IDS) data [24]
		Stateless events	✓			
		Video				
	Highway incident	Location	✓			Maryland highway & traffic information [25], Traffic Management Centers Data [26], traffic incident in Singapore [27]
		Time of date	✓			
		Weather conditions		✓		
		Vehicle involved			✓	
		Incident type		✓		
	Metro incident	Time	✓			Metro smartcard records in Shenzhen [28], urban rail transit system data [29]
		Station			✓	
		Check in/out		✓		

In addition to the aforementioned two data types, other multivariate data can be derived from trajectories or incident logs, or recorded by special sensors, including, velocity, direction, and acceleration. Table I summarizes the existing datasets and their respective attributes.

B. Data Preprocessing

A sequence of data preprocessing operations is required before data analysis; such operations include data cleaning, data matching, data organization and data aggregation.

- **Data cleaning** Data errors, outliers, and conflicting values of the raw data must be cleaned [33]. In [7], a step called GPS data cleaning was applied to remove GPS errors and filter out useless records.

  A typical data cleaning process has three phases: auditing data to find discrepancies, choosing transformations to fix such discrepancies, and applying the transformations to the dataset [44].

  The first phase detects errors in the raw data. Rahm *et al.* enumerated the major problems in raw data [45], including uniqueness, referential integrity, misspelling, redundancy, and contradictory values. Conventional approaches to detect errors in raw data include data profiling and data mining.

  In the second phase, data transformation is carefully designed and chosen depending on the number of data sources and dirtiness of the data. This phase can be completed manually or automatically. For example, the user can write custom scripts to control the whole procedure of cleaning or use extraction/transformation/loading (ETL) tools to transform the data.

  The third phase executes transformations of datasets and replaces dirty data with cleaned data. In traffic visualization systems, cleaned data need to be processed further to fit in the analysis tasks.
- **Data matching** Raw traffic data records are discrete sample points and may not match road networks in cities. Map matching, that is, aligning a sequence of observed user positions with the road network on a digital map, is an indispensable step in data preprocessing [46]. Existing map matching algorithms can be categorized into four classes, namely, geometric, topological, probabilistic,

TABLE II
EXISTING MAP MATCHING ALGORITHMS

Class	Characteristics	Ref.
Geometric	Two-stage matching	[48]
	Segment-based	[49]
Topological	Weight-based	[50]
Probabilistic	Particle filtering	[51]
	Able to reconcile inaccurate location	[52]
	Able to determine potential true path	[53]
	Multi-hypothesis road tracking-based	[54]
Advanced	Genetic MMA	[55]
	Interactive voting-based	[56]
	Hidden Markov model-based	[57]

and other advanced techniques [47], as summarized in Table II.

- **Data organization** The preprocessed data need to be organized in a database or data warehouse. A well-studied database should support interactive query and visualization of queried results, and should be compatible with data of moving objects such as trajectories. Indexing methods fall into two classes. The first class includes multidimensional index methods, such as 3D R-Tree [58], STR-Tree [59], and HR-Tree [60]. The second class includes indexing methods that divide space into grids and build a time index for each grid, such as SETI [61] and MTSB-Tree [62]. Data cube is another standard data structure that provides fast responses to data queries [63]. Recently, Nanocubes [64] have been developed to support quick indexing over time and aggregation queries over spatial regions. Several relational databases, such as PostGis (extension of PostgreSQL) [65] and MySQL Spatial (extension of MySQL) [66] provide spatial extensions for spatial data.
- **Data aggregation** Traffic datasets commonly contain spatial and temporal properties and span a large scale of space and time. Data aggregation [25] is effective in reducing the data size and provides convenience in subsequent analysis. The basic aggregation operations for traffic data are spatial (S), temporal (T), directional (D), and attribute (A)-related aggregations. Their combinations generate different types of aggregation: $S \times T$ aggregation, $S \times T \times A$ aggregation [25], [67], $S \times S \times T \times T$ aggregation [68], $S \times T \times D$ aggregation [69], and $S \times S$ aggregation [69]. S aggregation is mainly done by calculating the density of data points inside each grid of an area. T aggregation is employed to show changes along the time axis and is accomplished by merging data points in each time interval. The most common visualization corresponding to T aggregation is time histogram. $S \times T$ aggregation simply computes the density at consecutive time intervals [69]. The time-varying density can be visualized through an animated density map. $S \times T \times A$ aggregation [67] firstly groups spatial records on the basis of regularly sampled grids and then aggregating temporal attributes in each grid. $S \times S \times T \times T$ aggregation combines aggregations based on start location, end location, start time, and end time [68]. It counts the number of entities that move from one place to another in a period of time. $S \times T \times D$ aggregation not only aggregates data by space and time, but also aggregates by movement direction [69]. $S \times S$ aggregation groups trajectories or movements that have the same start locations and the same end locations [69]. Different aggregation strategies meet different requirements of analysis tasks.

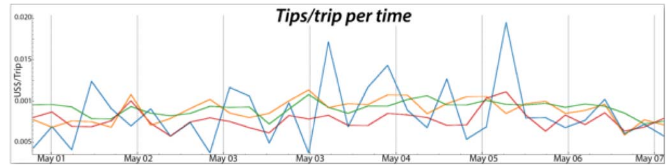


Fig. 2. Line charts representing linear time [14]. It shows tips per trip for taxi trips originating in different regions in the period of May 1, 2011 to May 7, 2011. Each line represents tips per trip in one region.

III. VISUALIZATION OF TRAFFIC DATA

Traffic data contain multiple variables, of which the most important ones are time and space. This section describes the visualization techniques specifically designed for the time, the locations, spatial-temporal information, and other properties in traffic data.

A. *Visualization of Time*

Generally, time can be classified into linear time, periodic time and branching time. Time-oriented visualization [70] emphasizes on the display of trend, periodicity, and abnormality of data along the time axis.

1) Linear Time: Linear time regards time as a linear field from a starting time point to an ending time point. It is the most widely used time representation and yields a sequence of timeline visualization techniques. For instance, in a line chart, the time is represented along the X-axis, and another variate is represented along the Y-axis. Fig. 2 shows the tips per taxi trip on a given date in New York City, USA [14]. Line charts are easy to read but they are not the right choice to show multiple variables due to the clutter problem.

Stacked graph [71] is another popular visual form. It can be used to show multiple quantities that are orderly accumulated along the Y-axis. The quantity of each variate at a certain time point is depicted along the length of a streaming chart. In this way, not only the individual quantity of each variate, but also its ratio to the sum of all variates can be disclosed. Among variants of stacked graph layout algorithms, ThemeRiver [72] is one of the best-known. It can create smooth, symmetrical and artistic stacked graph. In [23], ThemeRiver is employed to show the traffic volume at a road intersection, as shown in Fig. 17(b). Stacked graphs have no clutter problem but need more space than line charts.

Linear time is capable of expressing how traffic data vary with time and indicating peaks or valleys of variable evolutions over time.

2) Periodic Time: Many recursive processes occur in our natural world. Many of them are relevant to time, e.g., iterations of seasons, weeks, and days. A common way of visualizing periodicity is to use a radial layout, such as the visualization

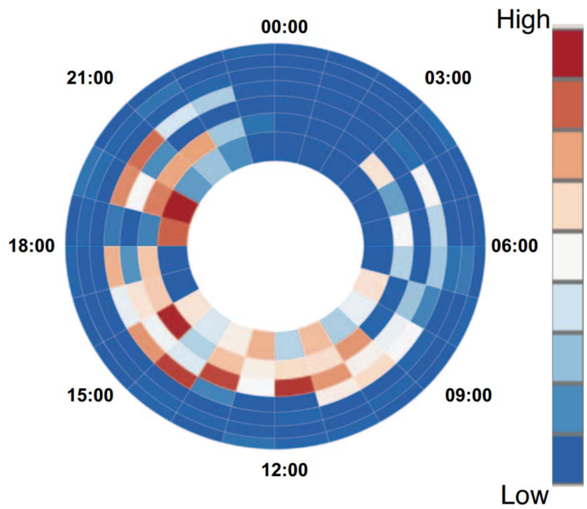


Fig. 3. Visualizing periodic time in radial layout [11]. Time in a day is shown on the circular axis and each ring represents a day. The sector color represents a selected traffic quantity with the color map shown on the right.

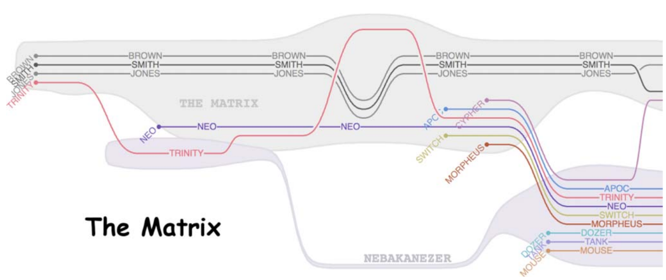


Fig. 4. A storyline of *The Matrix* shows how characters interact with each other in the movie [74].

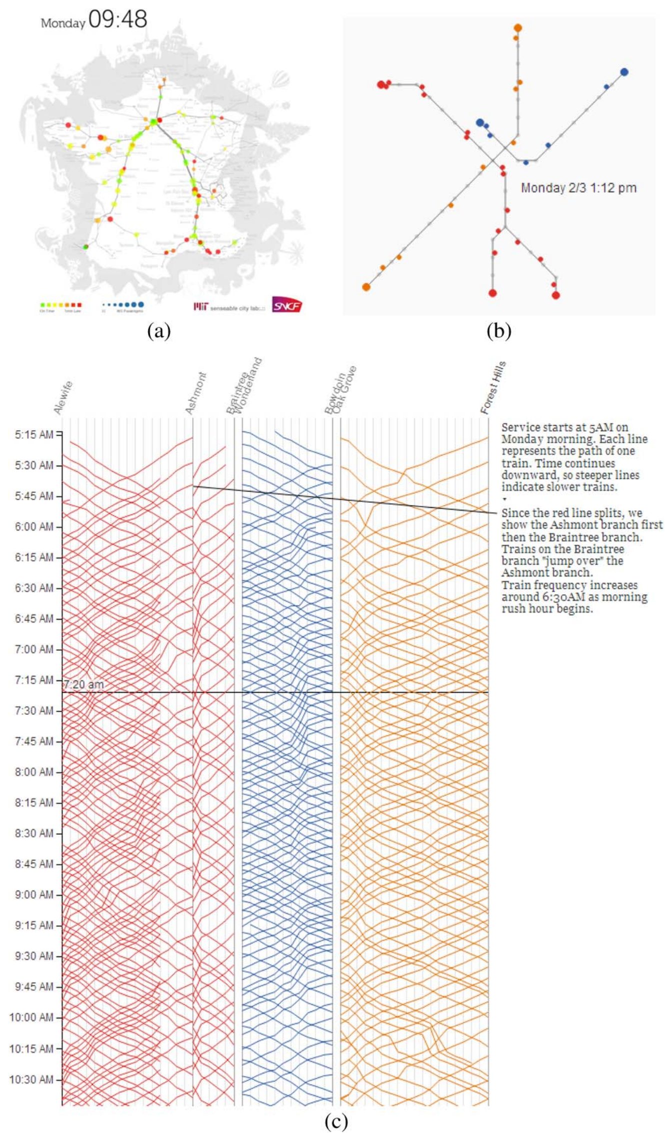


Fig. 5. (a) Train status at 9:48 a.m. in France [20], in which the visualization is based on a railway map and the running trains are labeled by colored points at their locations. (b) and (c) Boston subway status [21] is created based on Boston metro map: (b) the positions of running trains at 1:12 p.m. Monday; (c) an overview of the running status of the metro for one day.

of traffic information in Fig. 3 [11]. In the figure, each circle represents a day in a week, and each sector of a circle represents an hour. The advantage of a radial layout is it easily shows periodic patterns and the disadvantage is that it has low space efficiency.

3) Branching Time: An evolving event or story has many branching structures. An event or story can be described in many aspects. Visualizing branching time is commonly accomplished with a visual metaphor, called Storylines [73] (see Fig. 4), which can depict the progress, joining, branching, and disappearance of a specific event. To the best of our knowledge, traffic data has not yet been visualized through branching time.

B. Visualization of Spatial Properties

Location is the main spatial property of traffic data. It refers to where actions, incidents, or events happen. A series of locations distributed along the time axis forms a trajectory. Based on the aggregation level of location information, visualization of spatial properties can be categorized into three classes: point-based visualization (no aggregation), line-based visualization (first-order aggregation), and region-based visualization (second-order aggregation).

1) Point-Based Visualization: Point-based visualization considers samples of traffic information as individual discrete dots and presents these samples by leveraging point-relevant visual channels. Most traffic data are movement records of cars, planes, and pedestrians, thus this technique can intuitively show the position of objects at a certain time point. With the use of animation techniques, trajectories of objects can be observed straightforward.

In Trains of Data [20] project, the visualization represents each train as a moving dot, which runs in a 2D map. One example is the train transportation status at 9:48 a.m. in France, as shown in Fig. 5(a). Specifically, the dot size indicates the passenger number, and the dot color indicates whether the train is delayed, that is, green means on time and red means delayed.

Mike Barry and Brian Card [21] visualized the Boston subway system with the use of points to represent running trains in the map, and line charts to show the timing of trains [see Fig. 5(b) and (c)]. An overview of the subway system can be clearly obtained by examining these two figures.

Fig. 6. Visualization of hot spots in a city through the heatmap technique [75]. The red regions represent high volume of traffic, whereas the blue regions indicate low volume of traffic.

Dot-based representation typically places dots individually. The advantage of this method is that it enables the user to observe the states of every objects in the data. But when the data contains a large number of objects, the visualization becomes unclear and hard to understand. One can use a heatmap to show the integrated quantity of a large scale of objects in a map. For instance, the hot regions or roads in a traffic network can be depicted with a color-coded heatmap [75], as shown in Fig. 6. Kernel density estimation (KDE) is a commonly used algorithm for generating a heatmap. Network KDE (NKDE) is a modified KDE algorithm that is capable of characterizing certain point events along road networks [76], [77].

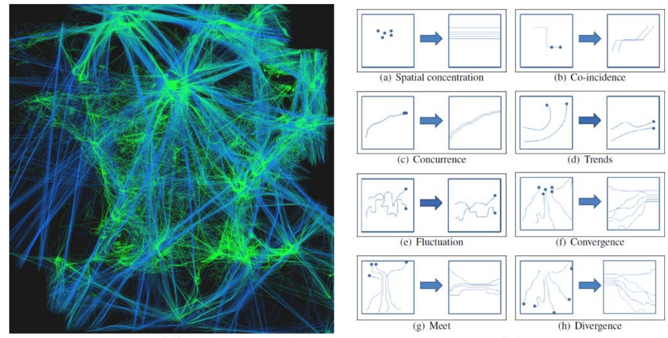
The advantage of point-based visualization is that it can show the distribution of vehicles. It can help the user explore where the busiest region of a city is but is inefficient in showing continuous information, such as how many vehicles travel from one location to another.

2) Line-Based Visualization: Line-based visualization techniques are designed to display traffic trajectories, roadmaps in a large-scale region, or traffic flow in a distributed network. Extended analysis on the basis of trajectories has proven to be useful in many applications, like semantic mining of trajectories [9], trajectory clustering [36], and route recommendation [10].

Conventionally, a trajectory is represented by a line or a curve and is scaled or colored with respect to its properties. The user can interactively navigate, select, and even analyze the set of trajectories. In [5], Hurter *et al.* presented an interactive system that analyzes aircraft trajectories over France. Each trajectory is represented by a line that connects the initial point and the last point, as shown in Fig. 7(a).

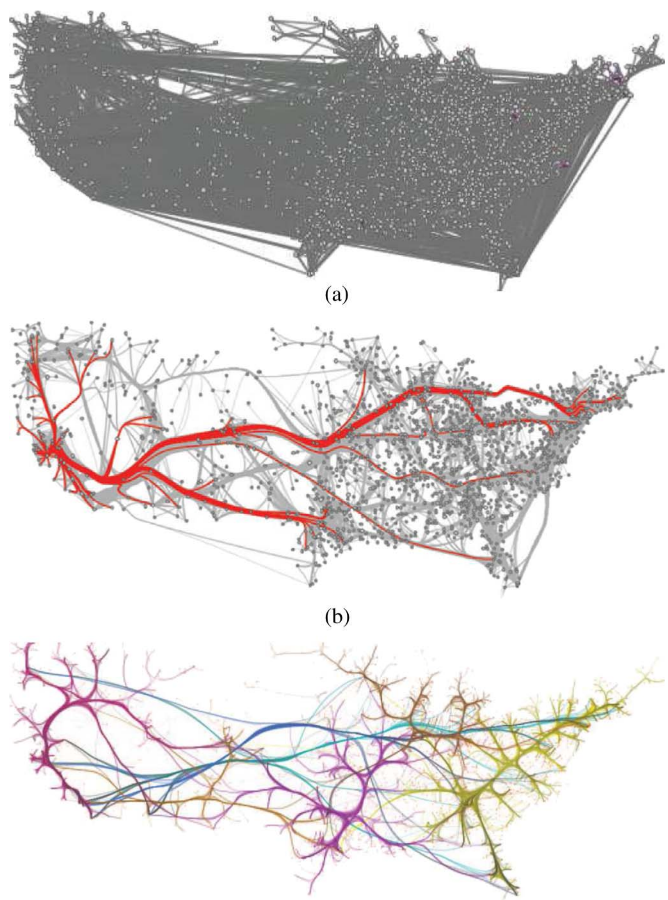
To overcome the complexity of trajectories, they can be transformed into other forms or simplified using topological and geometric algorithms. For instance, Tarik Crnovrsanin *et al.* [78] proposed to transform trajectories from the given spatial layout into an abstract space. This approach can effectively reveal patterns such as hazard prevention, migration patterns, and other behaviours (see Fig. 7). In the given spatial map, temporal information of entities is difficult to show especially for a considerable number of entities. In the abstract space, however, spatial information is shown on the Y-axis and temporal information is shown on the X-axis; therefore, patterns become clearer and more comprehensive. In Fig. 7(b), eight relationships among trajectories and associated transformed representations, namely, spatial concentration, co-incidence, concurrence, trends, fluctuation, convergence, meet, and divergence are defined.

When the number of trajectories becomes large, heavy visual clutter appears, making the visualization result unclear and disorderly. Many approaches have been developed to solve this



(a) (b)

Fig. 7. (a) Aircraft trajectories over France [75]. (b) By mapping trajectories from absolute coordinates to relative coordinates, eight patterns are discovered [78].



(a)

(b)

(c)

Fig. 8. (a) Line-based visualization of the immigration information in the USA [79]; (b) edge bundling result generated by [80]; (c) edge bundling result generated by [81].

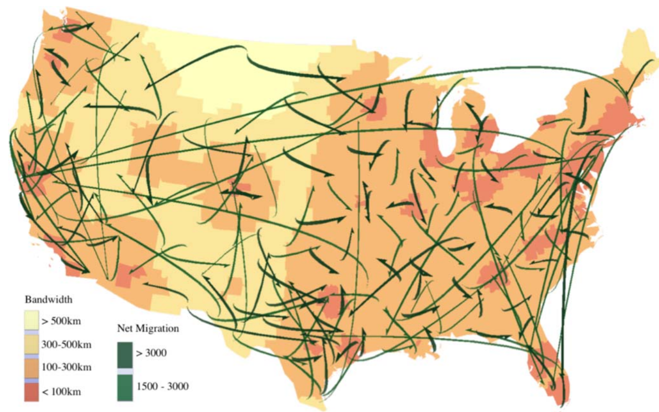


Fig. 9. The major immigration flows are visualized with arrows [82].

problem; the most effective of which is edge bundling [79]. Edge bundling transforms and groups similar edges into bundles. After twisting and clustering the original edges, the clutter is reduced. Fig. 8(a) shows the visualization of the immigration information among different states in the USA. Fig. 8(b) shows the result generated by [80] through a geometry-based edge bundling algorithm. Fig. 8(c) shows the result generated using a skeleton-based edge bundling algorithm [81].

Although edge bundling can reduce the clutter problem, recognizing the actual direction of connection between two locations becomes difficult. Guo *et al.* [82] presented an alternative solution that can extract major flow patterns in massive flow data without bundling or altering paths. It estimates the flow density for each pair of locations by a vector-based density model. Then a subset of smoothed paths is selected to represent the major flow in the flow map. Fig. 9 shows the immigration flows generated by this approach.

KDE can be applied to trajectories as well. In [4], [41], [83], density maps of trajectories are visualized (Fig. 10). Fig. 10(a) is the edge KDE result of the USA air traffic. Instead of twisting and clustering edges as edge bundling does, this technique uses color to indicate the density of trajectories: darker regions mean larger edge density. Fig. 10(b) shows the density map of vessel traffic around Rotterdam.

Line-based visualization can handle the task of analyzing trajectories. However, when the number of trajectories grows, the clutter problem becomes severe. Region-based visualization can be used to reduce the complexity of the visualization result.

3) Region-Based Visualization: Region-based visualization shows the traffic situation based on individual regions. Typically, traffic data are aggregated into regions based on predetermined rules. For instance, the traffic flow of cars is summarized along the streets, or the demographics are collected based on the administrative divisions. Zeng *et al.* applied region-based visualization to visualize the interchange patterns among different regions of a city. A radial metaphor is designed to represent the interchanges of one region to other regions, and other visual channels are employed to represent additional properties (see Fig. 11).

Region-based techniques have advantages in revealing macro patterns in traffic data. For example, when analyzing the pattern of vehicular movements from one region to another, region-

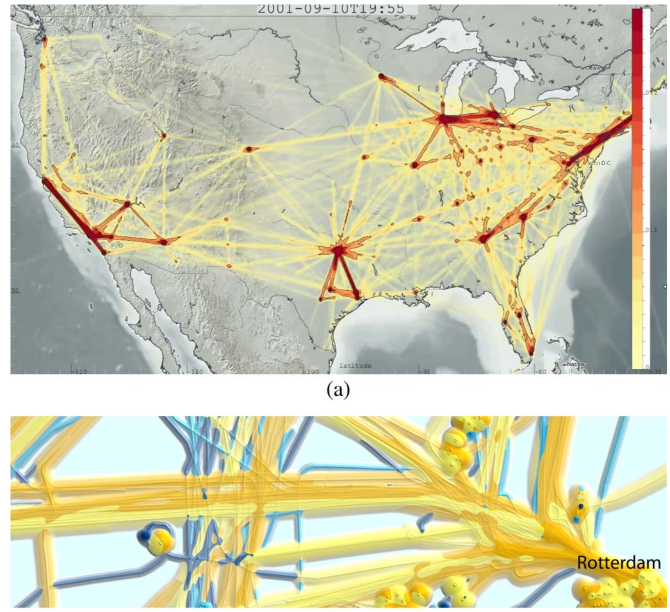


Fig. 10. Density maps of trajectories: (a) air traffic in the USA [83]; (b) vessel traffic around Rotterdam [4].

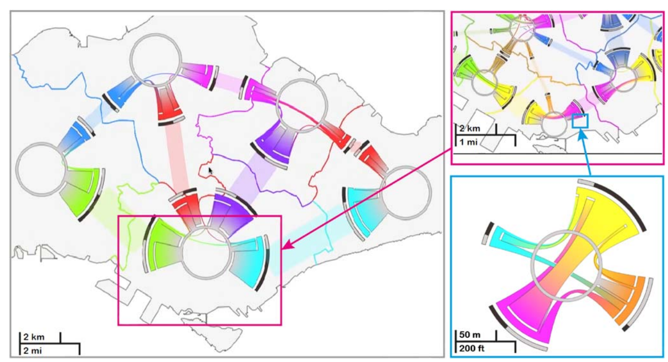


Fig. 11. A region-based visualization example: interchange patterns of metro systems in city scale, regional scale and road network scale [35].

based techniques work perfectly. However, they are inadequate for analyzing micro patterns, such as the patterns of a single vehicle. Level of detail techniques can be used to combine information at different scales and study both macro and micro patterns.

C. Spatio-Temporal Visualization

Space-Time-Cube (STC) [84] is a widely studied method for data with spatio-temporal attributes. In an STC (Fig. 12), a 3D trajectory is visualized in a 3D coordinate system, in which the plane consisting of the X-axis and the Y-axis is used for mapping spatial geographic information, and the Z-axis represents the time axis. In this way, spatio-temporal changes in an arbitrary object are depicted in a canonical space. The STC method has many variants, which are introduced in the following subsection.

D. Visualization of Multiple Properties

In many situations, traffic data contain various attributes in addition to spatial and temporal information. These attributes

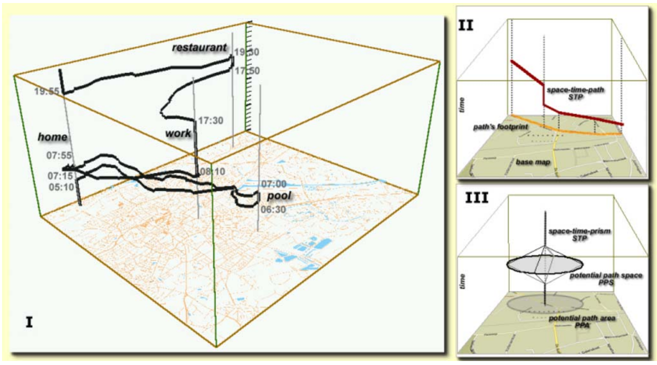
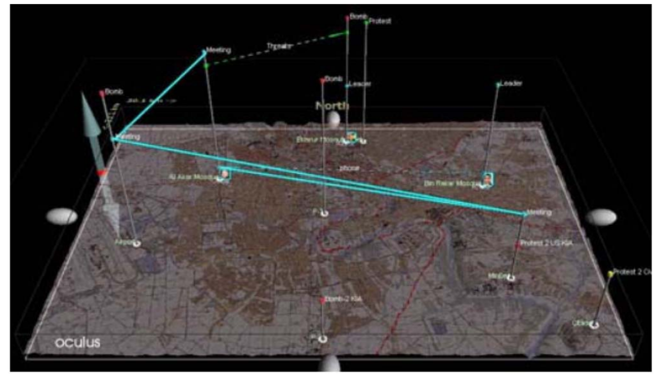


Fig. 12. Space-time-cube: both spatial information and temporal information are visualized in a cube [84]. The X-axis and Y-axis represent spatial information, whereas the Z-axis represents temporal information.

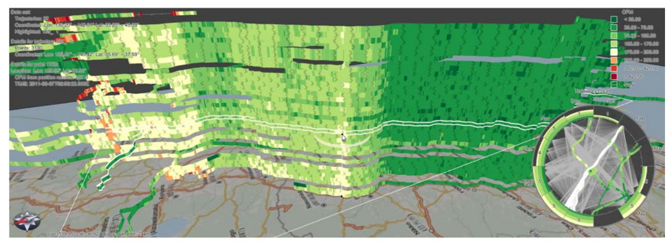
or properties can be roughly divided into three classes:

- Numerical Properties
  Numerical properties are continuous variables that represent quantitative values of data objects. Each numerical property describes one particular aspect of the data object, such as velocity, acceleration, weight, etc. Most of these properties are time variant, and thus the aforementioned time-oriented visualization techniques should be employed. However, in many applications, the user may focus on statistics of these properties. In this case, the histogram is a good choice for visualization.
- Categorical Properties
  Categorical properties are discrete variables that describe the state of data objects. Directions, vehicle types and incident types are representative categorical properties. The simplest visualization for categorical properties is color mapping, which assigns a particular color to represent a value. A popular color map scheme in terms of information visualization is the ColorBrewer system [85].
- Textual Properties
  Textual properties refer to words, lexical information, or logs that describe extra information about the traffic, such as the vehicle names involved in an incident, point of interests, and so on. These properties often contain semantic information and are essential for analyzing and explaining traffic situations. Text-based visualization techniques like TagCloud [86] and Wordle [87] can be employed to show a set of words. Some studies provide instructions on how to layout multiple labels on a 2D map effectively [88], [89].

To depict both spatio-temporal information and related properties, the standard STC can be enhanced. Representative ones include the GeoTime [43] and stacking-based STC [13]. The former adds objects and events at the corresponding points in the STC, as shown in Fig. 13(a). Specifically, each event is added to the track and placed around a corresponding time node to identify the event sender. Dotted lines are used to connect related objects and events. The latter method stacks multiple trajectories along the Z-axis, and visualizes them as stacked bands to depict velocity [see Fig. 13(b)].

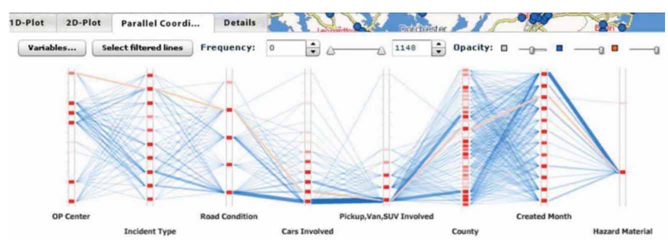


(a)

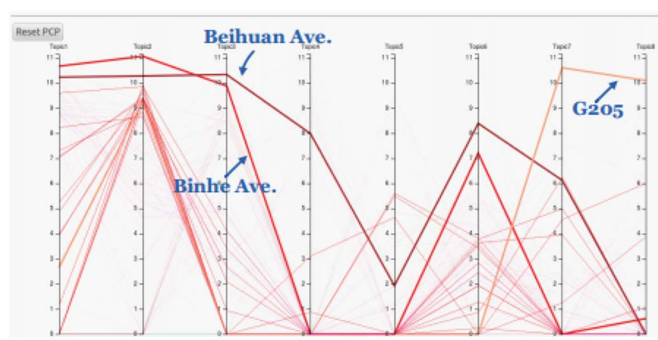


(b)

Fig. 13. Two variants of STC that illustrate related properties of traffic data: (a) GeoTime [43] in which events, objects, and activities are represented as the 3-D trajectories; (b) stack-based visualization [13], in which trajectories are stacked together.



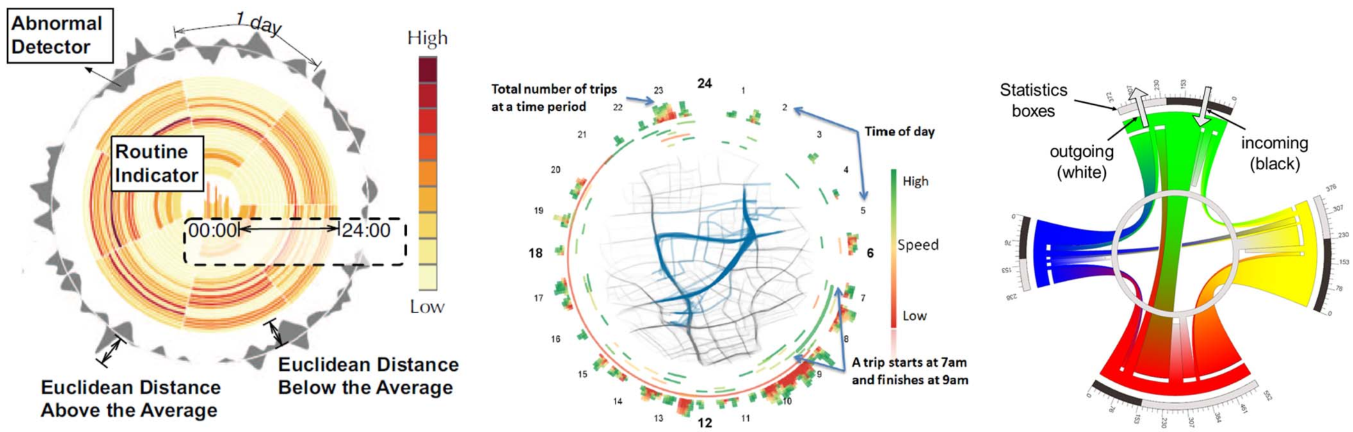
(a)



(b)

Fig. 14. Two examples of PCP in traffic visualization: (a) visualizing traffic incident data [26]; (b) showing the possibility of each street belonging to a particular topic represented by an axis [9].

Many traffic data visualization applications [26] employ parallel coordinate plot (PCP) [90] to show multiple attributes [Fig. 14(a)]. A parallel coordinate plot uses multiple parallel coordinate axes, each of which represents an attribute. Each data object is mapped into a set of connected lines that pass through all axes. In [9], PCP is used to help the user interactively discover knowledge from probability distributions over topics after clustering trajectories into different topics [Fig. 14(b)].



(a) (b) (c)

Fig. 15. Visualization of multivariate traffic information: (a) the distribution of taxis [11]; (b) taxi trips and statistical information [10]; (c) interchange patterns in movement data [35].

TABLE III
EXISTING VISUAL ANALYTICS SYSTEMS FOR TRAFFIC DATA

Tasks	Tools	Characteristics	Datasets	Platform
Situation-aware exploration and prediction	Ferreira et al. [14]	Visual query of taxi data	Taxi GPS data	-
	Liu et al. [75]	Visual analytics for ITS	Taxi GPS data	Java and Oracle
	Wang et al. [18]	Exploration of sparse traffic trajectory data	Traffic cell monitoring data	-
	Zeng et al. [94]	Exploration of mobility in transportation systems	Metro and bus passenger data	-
	Aurisano et al. [95]	User-driven predictive visual analytics	Lord's Resistance Army incident data	JS and HTML5
	Andrienko et al. [96]	Prediction by visual modeling	GPS data	-
Pattern discovery and clustering	Schreck et al. [34]	Visual clustering of trajectories	Trajectories extracted from financial data	-
	Andrienko et al. [36]	Visual clustering and classification trajectories	GPS data	-
	Rinzivillo et al. [92]	Progressive clustering of trajectories	GPS data	-
	Guo et al. [23]	Analysis of complex trajectory data	Traffic data at a road intersection	C++, Qt, boost and OpenGL
	Zeng et al. [35]	Visualization of interchange patterns	Metro and bus passenger data	Java
	Chu et al. [9]	Visualization of taxi topics	Taxi trajectories	-
Traffic situations monitoring	Pack et al. [26]	Traffic incident visualization	Highway incident data in Maryland	ColdFusion, Flex and PostgreSQL
	Anwar et al. [27]	Visualization of the impact of road incidents	Loop detector data and incident data	Java and Processing
	Piringer et al. [24]	Tunnel incidents visualization	Incident detection system data	C#
	Pu et al. [11]	Multilevel surveillance and analysis	Taxi GPS data	-
	Vandaniker et al. [93]	Situation-awareness and decision-making	Traffic incident data	ColdFusion, Flex and PostgreSQL
	Wang et al. [7]	Analysis of urban traffic congestion	Taxi GPS data	-

We can enhance conventional multivariate data visualization techniques with specially designed visual encoding and interaction schemes. For instance, a so-called "similarity lens" [Fig. 15(a)] was invented to show the distribution of taxis, average speed of taxis and pick-up/drop-off activities on roads [11]. Statistics of attributes of taxis (e.g., speed and vehicle density), value of attributes at each time, and the Euclidean distance against the average are simultaneously illustrated. Likewise, a trip view [Fig. 15(b)] is designed to show temporal, spatial, statistical and other attributes of taxi trajectory data [10]. Temporal statistical information is contained in the circle outside the area. Spatial information is shown by the map at the center of the illustration. The start and end time points of trips are shown by circular traces around the map. Fig. 15(c) shows an interchange circos diagram (ICD) [35] showing junction node, flow volume, direction of flow and statistics of flow volume.

IV. VISUAL ANALYSIS OF TRAFFIC DATA

A large number of visual analytics tools and applications have been developed for traffic data. They cover situation-aware exploration and prediction [5], [14], [75], [91], [94], pattern discovery and clustering [9], [23], [34]–[36], [68], [92], and traffic situation monitoring [7], [11], [24], [26], [27], [93]. In this section, we present representative works, and summarize these works in Table III.

A. *Situation-Aware Exploration and Prediction*

Querying unstructured data, especially moving objects is challenging for traditional data cube-based query model. Many studies focuses on new query models for fast response to query on traffic data and interfaces for exploring traffic information.

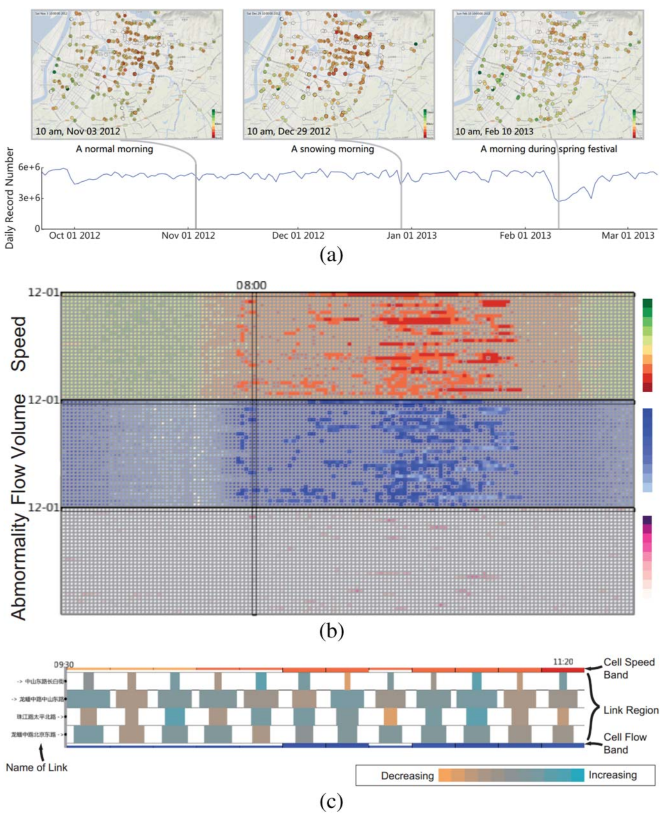


Fig. 16. A visual analytics system that supports the exploration of sparse traffic trajectory data [18]. Three views, namely, the map view, the pixel map view, and the linked view are provided.

Ferreira *et al.* proposed a new model that allows the user to visually query taxi trips [14]. In this model, three types of query constraints exist: spatial, temporal and attribute constraints. These constraints are visually specified and modulated through five views. The map view displays the query results and specifies the spatial constraints. The time selection widget specifies the temporal constraints. The data summary view shows the information associated with the results of the queries. The tool bar provides several operations (e.g., pickups, drop-offs, etc.). Multiple coordinated views specify both temporal and attribute constraints.

Wang *et al.* presented a system that evaluates the real traffic situations based on taxi trajectory data [91]. A road-based query model and a hash-based data structure are proposed to support dynamic query for trajectories. The presented system effectively fulfills the tasks of data-driven road evaluation.

Visual analytics for intelligent transportation (VAIT) is a system that visualizes traffic data and supports analytics queries through an interactive visual interface [75]. Twelve queries are defined for traffic data. The analysis procedure includes three steps: overview, distribution exploration and evolution exploration. Thereafter, query and recheck are performed to investigate the findings.

Wang *et al.* recently presented a visual analytics system to explore sparse traffic data [18] (see Fig. 16). Local animation and aggregation techniques are employed to address the uncertainty problem in sparse data. The analysis procedure is composed of three steps, namely, global exploration, cell exploration and correlation exploration, which are facilitated by three views. With the help of an integrated visual interface, the user can easily observe the hidden macro-patterns.

Aurisano *et al.* [95] presented a user-driven visual analytics system on multivariate and spatio-temporal incident reporting data on the Lord's Resistance Army activity in Central Africa. The data used in this system are similar to traffic incident data. The proposed prediction method is also feasible for traffic incident prediction scenarios. Andrienko *et al.* [96] developed interactive visual interfaces that represent the interdependencies between traffic intensities and speeds in an abstracted road network, which can be utilized for forecasting the expectable normal traffic situation at a given moment, and its development over time.

B. Pattern Discovery and Clustering

Detecting patterns in object movement and clustering the trajectories can be greatly enhanced by visualization and interactions.

Schreck *et al.* proposed a visual-interactive monitoring and control framework [34] that extends the Kohonen feature map (or self-organizing map, SOM). One distinctive feature of this work is the combination of automatic data analysis and human expert supervision. The user can monitor and control the SOM clustering process and obtain appropriate cluster results. Although the data used in this system are not traffic data, the clustering method can be used for clustering traffic trajectories.

In [36], [92], the OPTICS algorithm, a member of the DBSCAN family, is applied to cluster the trajectory data. The user can refine the clustering result by interactive visualization and manipulation, such as excluding one or several subclusters from a cluster, making a new cluster, or dividing a subcluster into two or more smaller subclusters. This approach is more efficient than traditional approaches because it incorporates human intelligence in the analysis loop.

Zeng *et al.* presented a suite of visual analysis techniques to study interchange patterns in movement data [35]. ICD is designed to examine interchange patterns [see Fig. 15(c)]. It supports depiction in three scales, namely, city, region and road network scales (see Fig. 11). At the road network scale, each ICD represents a road junction. At the city/region scale, each ICD represents a partitioned area.

Triple Perspective Visual Trajectory Analytics (TripVista) is an interactive visual analytics system for exploring and analyzing complex traffic trajectory data [23]. The system mainly consists of three views: traffic, ThemeRiver, and PCP, as shown in Fig. 17. The traffic view displays spatial information. The ThemeRiver view displays directional information, and the PCP view displays multidimensional data. Utilizing these views, the user can effectively detect regular and irregular traffic flow patterns.

Visual Analytics of Taxi Topic is a visual analytics system that discovers the movement patterns in taxi trajectories [9]. It integrates four views: taxi topic maps, street cloud, PCP and topic routes. The main contribution of this system is the transformation of taxi trajectories into documents and the use of latent Dirichlet allocation (LDA) to find hidden semantics from

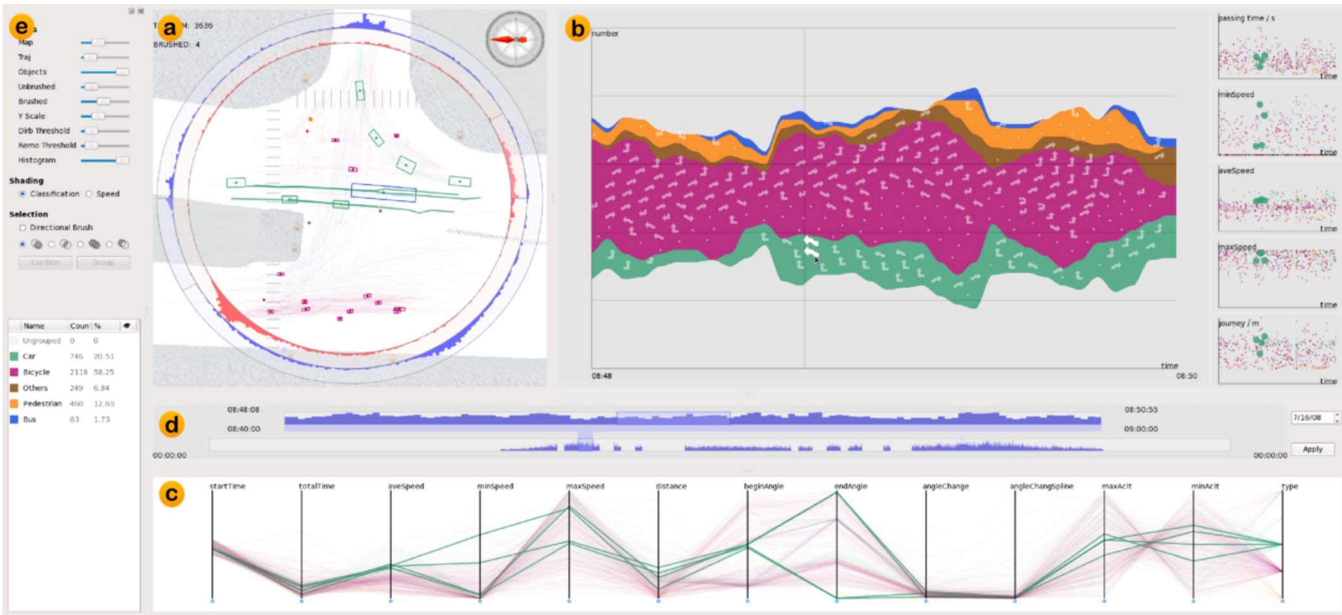


Fig. 17. TripVista is a visual analytics system for finding traffic flow patterns at a road intersection [23].

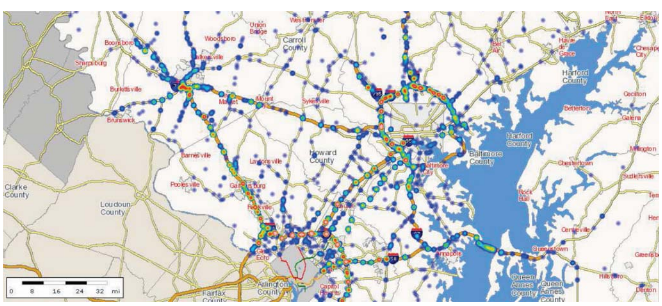


Fig. 18. ICE is a visually assisted application for studying transportation incident data [26].

the documents, namely, taxi topics that disclose interesting movement patterns of taxis.

C. Visual Monitoring of Traffic Situations

Traffic monitoring focuses on investigating traffic incidents. Datasets used by traffic monitoring systems either have incident records or not. For the latter, incidents can be extracted from raw data.

Incident Cluster Explorer (ICE) is an application that studies transportation incident datasets [26] (Fig. 18). Geospatial visualization (map), histogram, 2D plots, and PCP are integrated in the application. Incidents are visualized on the map in two modes: icon mode, which employs colored dots, and heatmap mode, which depicts density distribution.

The Traffic Origins system is designed to perform traffic incident analysis [27]. Animation transition techniques are employed to emphasize the start and the end of an incident. When an incident occurs, a circle appears and surrounds the site of the incident. When an incident ends, the corresponding circle fades out progressively. The color of each road indicates the average speed of vehicles on it.

AIVis [24] is a system that monitors traffic situations in road tunnels. The incidents are detected automatically from video sequences in real time. As shown in Fig. 19, the system comprise spatialCtemporal views, including future view, present view, history view, temporal overview, and additional windows. In particular, the present view is depicted in a tunnel shape and shows special locations, such as positions of cameras and emergency exists in tunnels. The history view shows incidents that occurred in the last three minutes, and the future view predicts incidents that may happen in a minute.

T-Watcher is an interactive visual analytics system for monitoring and analyzing complex traffic situations in big cities [11]. The monitoring task is accomplished in three views: region view, road view, and vehicle view (Fig. 20). Each view corresponds to a specially designed fingerprint that allows the user to complete a specialized task.

The transportation incident management explorer (TIME) is a system that combines temporal and spatial data with incident logs [93]. TIME integrates six visualizations, namely, communications, variable message signs, responders, lane status, traffic speed and traffic volumes.

Wang *et al.* presented a visual analytics system for traffic jams [7]. Traffic jams are automatically detected by setting a threshold for road speed. The system integrates five views: the spatial view presents an overview of traffic jam; the road speed view shows the speed patterns of each road; the graph list view shows the list of the propagation graphs; the graph projection view shows the topological relationships of propagation graphs; and the multifaceted filter view provides a dynamic query tool for querying propagation graphs, as shown in Fig. 21.

V. Conclusion

Large data brings numerous opportunities and challenges to the field of traffic data analysis. Traffic data visualization

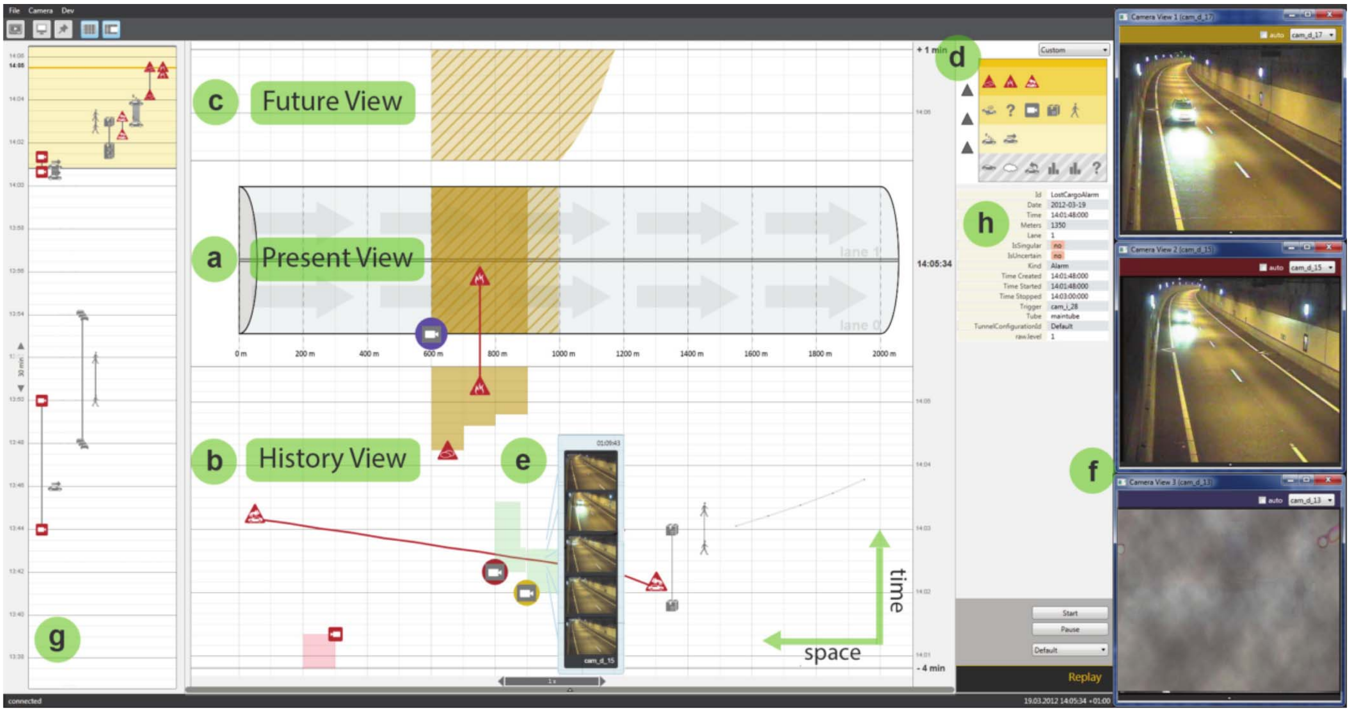


Fig. 19. AIVis is a system that monitors traffic situation in road tunnels [24].

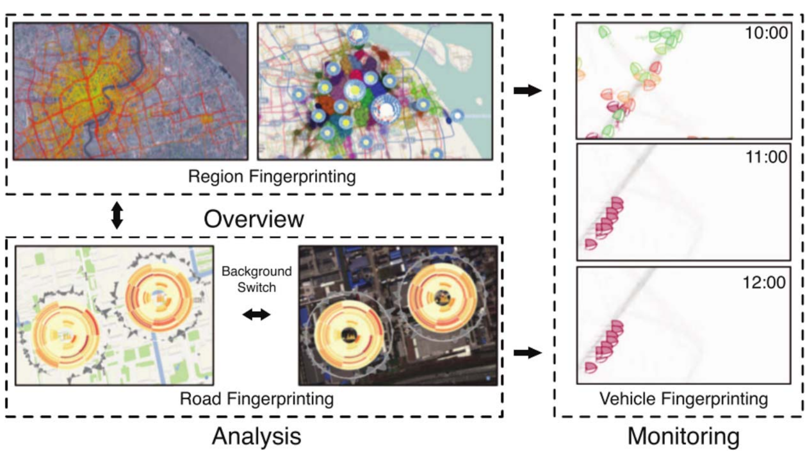


Fig. 20. T-Watcher system monitors and analyzes traffic by means of three levels of fingerprints [11].

performs a key function in addressing the problems arising from large-scale, multi-modal, and unstructured data. This paper provides an overview of relevant visualization techniques and visual analysis systems in the context of traffic analysis, and presents the common data flow in traffic data visualization. According to the characteristics of traffic data, visualization techniques for traffic analysis are presented in four aspects: temporal, spatial, spatio-temporal, and multi-variable. Existing traffic visualization applications and visual analytics systems are presented based on the analysis tasks.

Performing analysis tasks in real-time is difficult when the data size if large. Few works supporting visual analysis of big traffic data are available. Ferreira *et al.* [14] presented a system that supports visual exploration of huge spatio-temporal data. However, developments do not stop in the design of such a system. Thus, the analysis in situation-aware and immersive environments are promising directions.

Visual analytics provides a comprehensible way to analyze data and consequently significantly improves the efficiency and accuracy of the analysis. In the context of ITS, visual analytics can accomplish various tasks, such as route planning, traffic jam detection, accident monitoring and flow patterns recognition. However, most existing traffic visualization and visual analytics systems employ offline data. Designing and implementing systems using on-line and streaming data may be a potential research direction.

Benefiting from the development and popularization of sensor technology, data sources related to traffic data are currently growing in number. For instance, video surveillance [24] has been combined with road incident data for better monitoring.

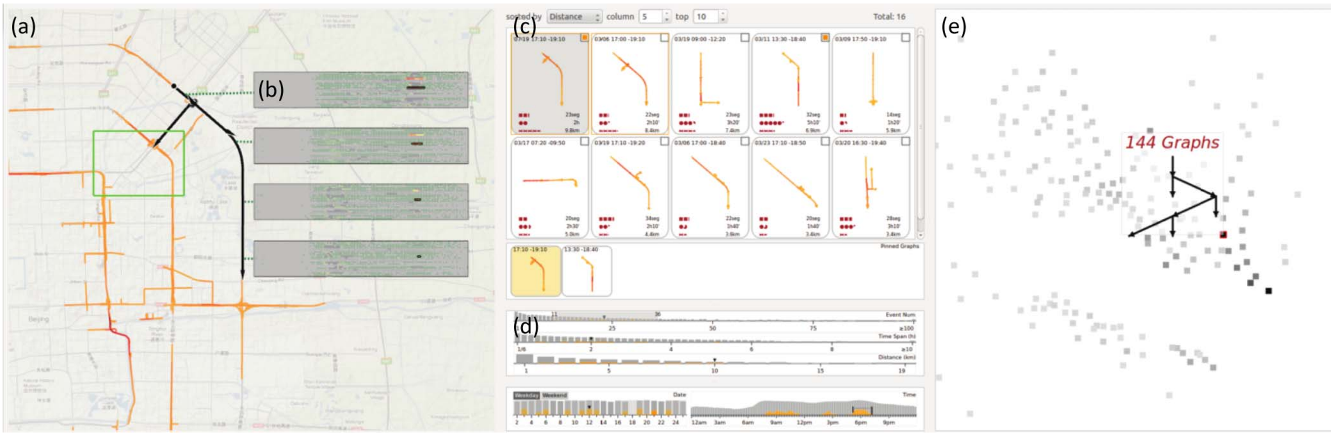


Fig. 21. A visual analytics system for analyzing traffic jams comprising five views: (a) spatial view; (b) road speed view; (c) graph list view; (d) multifaceted filter view; (e) graph projection view [7].

We believe that visual analysis of heterogeneous data from different sources (e.g., social media) will be the next research topic on data-driven ITS.

Another interesting direction is the visualization and visual analytics of social transportation, with an aim of collecting, analyzing, and utilizing data from cyber, physical, and social spaces for ITS.

REFERENCES

[1] "Traffic." [Online]. Available: http://en.wikipedia.org/wiki/Traffic_(disambiguation)

[2] J. Zhang *et al.*, "Data-driven intelligent transportation systems: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1624–1639, Dec. 2011.

[3] L. Figueiredo, I. Jesus, J. T. Machado, J. Ferreira, and J. M. de Carvalho, "Towards the development of intelligent transportation systems," in *Proc. IEEE Intell. Transp. Syst.*, 2001, pp. 1206–1211.

[4] R. Scheepens, N. Willems, H. van de Wetering, and J. J. van Wijk, "Interactive visualization of multivariate trajectory data with density maps," in *Proc. IEEE Pac. Vis. Symp.*, 2011, pp. 147–154.

[5] C. Hurter, B. Tissoires, and S. Conversy, "FromDaDy: Spreading aircraft trajectories across views to support iterative queries," *IEEE Trans. Vis. Comput. Graphics*, vol. 15, no. 6, pp. 1017–1024, Nov./Dec. 2009.

[6] "Europe 24—An Air Traffic Data Visualisation." [Online]. Available: http://nats.aero/blog/2014/03/europe-24-air-traffic-data-visualisation/

[7] Z. Wang, M. Lu, X. Yuan, J. Zhang, and H. V. D. Wetering, "Visual traffic jam analysis based on trajectory data," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2159–2168, Dec. 2013.

[8] X. Liang, X. Zheng, W. Lv, T. Zhu, and K. Xu, "The scaling of human mobility by taxis is exponential," *Phys. A, Stat. Mech. Appl.*, vol. 391, no. 5, pp. 2135–2144, Mar. 2012.

[9] D. Chu *et al.*, "Visualizing hidden themes of taxi movement with semantic transformation," in *Proc. IEEE Pac. Vis. Symp.*, 2014, pp. 137–144.

[10] H. Liu *et al.*, "Visual analysis of route diversity," in *Proc. IEEE Conf. Visual Anal. Sci. Technol.*, 2011, pp. 171–180.

[11] J. Pu, S. Liu, Y. Ding, H. Qu, and L. Ni, "T-Watcher: A new visual analytic system for effective traffic surveillance," in *Proc. IEEE 14th Int. Conf. MDM*, 2013, vol. 1, pp. 127–136.

[12] Y. Liu, C. Kang, S. Gao, Y. Xiao, and Y. Tian, "Understanding intra-urban trip patterns from taxi trajectory data," *J. Geogr. Syst.*, vol. 14, no. 4, pp. 463–483, Oct. 2012.

[13] C. Tominski, H. Schumann, G. Andrienko, and N. Andrienko, "Stacking-based visualization of trajectory attribute data," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 12, pp. 2565–2574, Dec. 2012.

[14] N. Ferreira, J. Poco, H. T. Vo, J. Freire, and C. T. Silva, "Visual exploration of big spatio-temporal urban data: A study of New York City taxi trips," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2149–2158, Dec. 2013.

[15] Y. Yue, Y. Zhuang, Q. Li, and Q. Mao, "Mining time-dependent attractive areas and movement patterns from taxi trajectory data," in *Proc. 17th Int. Conf. Geoinf.*, 2009, pp. 1–6.

[16] Y. Yue, B. Hu, H. D. Wang, and Q. Q. Li, "Identifying shopping center attractiveness using taxi trajectory data," in *Proc. Int. Workshop Trajectory Data Mining Anal.*, 2011, pp. 31–36.

[17] B. Jiang, J. Yin, and S. Zhao, "Characterizing the human mobility pattern in a large street network," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 80, no. 2, Aug. 2009, Art. ID. 021136.

[18] Z. Wang *et al.*, "Visual exploration of sparse traffic trajectory data," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 1813–1822, Dec. 2014.

[19] C. A. Quiroga and D. Bullock, "Travel time studies with global positioning and geographic information systems: An integrated methodology," *Transp. Res. C, Emerging Technol.*, vol. 6, no. 1/2, pp. 101–127, Feb. 1998.

[20] K. Kloeckl, X. Chen, C. Sommer, C. Ratti, and A. Biderman, "Trains of data." [Online]. Available: http://senseable.mit.edu/trainsofdata/

[21] M. Barry and B. Card, "Boston's Massachusetts Bay Transit Authority (MBTA)." [Online]. Available: http://mbtaviz.github.io/

[22] I. Rhee, M. Shin, S. Hong, K. Lee, and S. Chong, "On the levy-walk nature of human mobility," in *Proc. IEEE INFOCOM*, 2008, pp. 924–932.

[23] H. Guo, Z. Wang, B. Yu, H. Zhao, and X. Yuan, "TripVista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection," in *Proc. IEEE Pac. Vis. Symp.*, 2011, pp. 163–170.

[24] H. Piringer, M. Buchetics, and R. Benedik, "AlVis: Situation awareness in the surveillance of road tunnels," in *Proc. IEEE Conf. Visual Anal. Sci. Technol.*, 2012, pp. 153–162.

[25] A. Fredrikson, C. North, C. Plaisant, and B. Shneiderman, "Temporal, geographical and categorical aggregations viewed through coordinated displays: A case study with highway incident data," in *Proc. Workshop New Paradigms Inf. Vis. Manipulation*, 1999, pp. 26–34.

[26] M. L. Pack, K. Wongsuphasawat, M. VanDaniker, and D. Filippova, "Ice-visual analytics for transportation incident datasets," in *Proc. IEEE Int. Conf. Inf. Reuse Integr.*, 2009, pp. 200–205.

[27] A. Anwar, T. Nagel, and C. Ratti, "Traffic origins: A simple visualization technique to support traffic incident analysis," in *IEEE Pac. Vis. Symp.*, 2014, pp. 316–319.

[28] Y. Gong *et al.*, "Exploring spatiotemporal characteristics of intra-urban trips using metro smartcard records," in *Proc. 20th Int. Conf. Geoinf.*, 2012, pp. 1–7.

[29] J. Wang, Y. Li, J. Liu, K. He, and P. Wang, "Vulnerability analysis and passenger source prediction in urban rail transit networks," *PloS One*, vol. 8, no. 11, 2013, Art. ID. e80178.

[30] C. D. Hansen and C. R. Johnson, *The Visualization Handbook*. San Diego, CA, USA: Academic, 2004.

[31] C. Ware, *Information Visualization: Perception for Design (Interactive Technologies)*. San Mateo, CA, USA: Morgan Kaufman, 2012.

[32] J. J. Thomas and K. A. Cook, *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. Richland, WA, USA: National Visualization and Analytics Center, 2005.

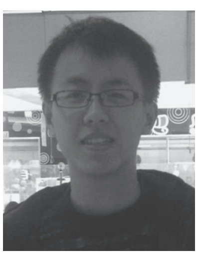
[33] S. Kandel *et al.*, "Research directions in data wrangling: Visualizations and transformations for usable and credible data," *Inf. Vis.*, vol. 10, no. 4, pp. 271–288, Oct. 2011.

[34] T. Schreck, J. Bernard, T. Von Landesberger, and J. Kohlhammer, "Visual cluster analysis of trajectory data with interactive Kohonen maps," *Inf. Vis.*, vol. 8, no. 1, pp. 14–29, Jan. 2009.

[35] W. Zeng, C.-W. Fu, S. M. Arisona, and H. Qu, "Visualizing interchange patterns in massive movement data," *Comput. Graph. Forum*, vol. 32, no. 3. pp. 271–280, Jun. 2013.
[36] G. Andrienko *et al.*, "Interactive visual clustering of large collections of trajectories," in *Proc. IEEE Symp. Visual Anal. Sci. Technol.*, 2009, pp. 3–10.
[37] J. Zhou, A. K. Tung, W. Wu, and W. S. Ng, "A semi-lazy approach to probabilistic path prediction," in *Proc. ACM SIGKDD*, 2013, pp. 748–756.
[38] S. K. Card, J. D. Mackinlay, and B. Shneiderman, *Readings in Information Visualization: Using Vision to Think*. San Mateo, CA, USA: Morgan Kaufmann, 1999.
[39] N. Andrienko, G. Andrienko, N. Pelekis, and S. Spaccapietra, "Basic concepts of movement data," in *Mobility, Data Mining and Privacy*. Berlin, Germany: Springer-Verlag, 2008, pp. 15–38.
[40] N. Andrienko, G. Andrienko, H. Stange, T. Liebig, and D. Hecker, "Visual analytics for understanding spatial situations from episodic movement data," *KI–Künstl. Intell.*, vol. 26, no. 3, pp. 241–251, Aug. 2012.
[41] R. Scheepens *et al.*, "Composite density maps for multivariate trajectories," *IEEE Trans. Vis. Comput. Graphics*, vol. 17, no. 12, pp. 2518–2527, Dec. 2011.
[42] G. Andrienko, N. Andrienko, and S. Wrobel, "Visual analytics tools for analysis of movement data," *ACM SIGKDD Explor. Newslett.*, vol. 9, no. 2, pp. 38–46, Dec. 2007.
[43] T. Kapler and W. Wright, "Geotime information visualization," *Inf. Vis.*, vol. 4, no. 2, pp. 136–146, Summer 2005.
[44] V. Raman and J. M. Hellerstein, "Potter's wheel: An interactive data cleaning system," in *Proc. VLDB*, 2001, vol. 1, pp. 381–390.
[45] E. Rahm and H. H. Do, "Data cleaning: Problems and current approaches," *IEEE Data Eng. Bull.*, vol. 23, no. 4, pp. 3–13, Dec. 2000.
[46] Y. Lou *et al.*, "Map-matching for low-sampling-rate GPS trajectories," in *Proc. ACM SIGSPATIAL*, 2009, pp. 352–361.
[47] M. A. Quddus, W. Y. Ochieng, and R. B. Noland, "Current map-matching algorithms for transport applications: State-of-the art and future research directions," *Transp. Res. C, Emerging Technol.*, vol. 15, no. 5, pp. 312–328, Oct. 2007.
[48] X. Li, M. Li, W. Shu, and M. Wu, "A practical map-matching algorithm for GPS-based vehicular networks in Shanghai urban area," in *Proc. IET Conf. Wireless, Mobile Sensor Netw.*, 2007, pp. 454–457.
[49] S. S. Chawathe, "Segment-based map matching," in *Proc. IEEE Intell. Veh. Symp.*, 2007, pp. 1190–1197.
[50] N. R. Velaga, M. A. Quddus, and A. L. Bristow, "Developing an enhanced weight-based topological map-matching algorithm for intelligent transport systems," *Transp. Res. C, Emerging Technol.*, vol. 17, no. 6, pp. 672–683, Dec. 2009.
[51] P. Bonnifait *et al.*, "Multi-hypothesis map-matching using particle filtering," in *Proc. 16th World Congr. ITS Syst. Serv.*, 2009, pp. 1–8.
[52] W. Y. Ochieng, M. Quddus, and R. B. Noland, "Map-matching in complex urban road networks," *Rev. Brasileira Cartogr.*, vol. 55, no. 2, pp. 1–14, 2009.
[53] M. Bierlaire, J. Chen, and J. Newman, "A probabilistic map matching method for smartphone GPS data," *Transp. Res. C, Emerging Technol.*, vol. 26, pp. 78–98, Jan. 2013.
[54] M. Jabbour, P. Bonnifait, and V. Cherfaoui, "Map-matching integrity using multihypothesis road-tracking," *J. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 189–201, Nov. 2008.
[55] F. C. Pereira, H. Costa, and N. M. Pereira, "An off-line map-matching algorithm for incomplete map databases," *Eur. Transp. Res. Rev.*, vol. 1, no. 3, pp. 107–124, Oct. 2009.
[56] J. Yuan, Y. Zheng, C. Zhang, X. Xie, and G.-Z. Sun, "An interactive-voting based map matching algorithm," in *Proc. 11th Int. Conf. Mobile Data Manage.*, 2010, pp. 43–52.
[57] C. Y. Goh *et al.*, "Online map-matching based on hidden Markov model for real-time traffic sensing applications," in *Proc. IEEE Intell. Transp. Syst.*, 2012, pp. 776–781.
[58] Y. Theoderidis, M. Vazirgiannis, and T. Sellis, "Spatio-temporal indexing for large multimedia applications," in *Proc. 3rd IEEE Int. Conf. Multimedia Comput. Syst.*, 1996, pp. 441–448.
[59] K. Deng, K. Xie, K. Zheng, and X. Zhou, "Trajectory indexing and retrieval," in *Computing With Spatial Trajectories*. Berlin, Germany: Springer-Verlag, 2011, pp. 35–60.
[60] M. A. Nascimento and J. R. Silva, "Towards historical r-trees," in *Proc. ACM Symp. Appl. Comput.*, 1998, pp. 235–240.
[61] V. P. Chakka, A. C. Everspaugh, and J. M. Patel, "Indexing large trajectory data sets with SETI," in *Proc. CIDR*, 2003, pp. 48109–2122.
[62] P. Zhou, D. Zhang, B. Salzberg, G. Cooperman, and G. Kollios, "Close pair queries in moving object databases," in *Proc. 13th Annu. ACM Int. Workshop Geogr. Inf. Syst.*, 2005, pp. 2–11.
[63] J. Gray *et al.*, "Data cube: A relational aggregation operator generalizing group-by, cross-tab, and sub-totals," *Data Mining Knowl. Discov.*, vol. 1, no. 1, pp. 29–53, 1997.
[64] L. Lins, J. T. Klosowski, and C. Scheidegger, "Nanocubes for real-time exploration of spatiotemporal datasets," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2456–2465, Dec. 2013.
[65] "Postgis." [Online]. Available: http://postgis.net/
[66] "Mysql Spatial." [Online]. Available: http://dev.mysql.com/doc/refman/5.0/en/spatial-extensions.html
[67] A. Slingsby, J. Dykes, and J. Wood, "Using treemaps for variable selection in spatio-temporal visualisation," *Inf. Vis.*, vol. 7, no. 3/4, pp. 210–224, Sep. 2008.
[68] D. Guo, J. Chen, A. M. MacEachren, and K. Liao, "A visualization system for space-time and multivariate patterns (VIS-STAMP)," *IEEE Trans. Vis. Comput. Graphics*, vol. 12, no. 6, pp. 1461–1474, Nov./Dec. 2006.
[69] G. Andrienko and N. Andrienko, "Spatio-temporal aggregation for visual analysis of movements," in *Proc. IEEE Symp. Visual Anal. Sci. Technol.*, 2008, pp. 51–58.
[70] W. Aigner, S. Miksch, W. Muller, H. Schumann, and C. Tominski, "Visual methods for analyzing time-oriented data," *IEEE Trans. Vis. Comput. Graphics*, vol. 14, no. 1, pp. 47–60, Jan./Feb. 2008.
[71] L. Byron and M. Wattenberg, "Stacked graphs—Geometry & aesthetics," *IEEE Trans. Vis. Comput. Graphics*, vol. 14, no. 6, pp. 1245–1252, Nov./Dec. 2008.
[72] S. Havre, B. Hetzler, and L. Nowell, "ThemeRiver: Visualizing theme changes over time," in *Proc. IEEE Symp. Inf. Vis.*, 2000, pp. 115–123.
[73] M. Ogawa and K.-L. Ma, "Software evolution storylines," in *Proc. 5th Int. Symp. Softw. Vis.*, 2010, pp. 35–42.
[74] Y. Tanahashi and K.-L. Ma, "Design considerations for optimizing storyline visualizations," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 12, pp. 2679–2688, Dec. 2012.
[75] S. Liu *et al.*, "VAIT: A visual analytics system for metropolitan transportation," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1586–1596, Dec. 2013.
[76] Z. Xie and J. Yan, "Kernel density estimation of traffic accidents in a network space," *Comput., Environ. Urban Syst.*, vol. 32, no. 5, pp. 396–406, Sep. 2008.
[77] G. Borruso, "Network density estimation: A GIS approach for analysing point patterns in a network space," *Trans. GIS*, vol. 12, no. 3, pp. 377–402, Jun. 2008.
[78] T. Crnovrsanin, C. Muelder, C. Correa, and K.-L. Ma, "Proximity-based visualization of movement trace data," in *Proc. IEEE Symp. Visual Anal. Sci. Technol.*, 2009, pp. 11–18.
[79] H. Zhou, P. Xu, X. Yuan, and H. Qu, "Edge bundling in information visualization," *Tsinghua Sci. Technol.*, vol. 18, no. 2, pp. 145–156, Apr. 2013.
[80] W. Cui, H. Zhou, H. Qu, P. C. Wong, and X. Li, "Geometry-based edge clustering for graph visualization," *IEEE Trans. Vis. Comput. Graphics*, vol. 14, no. 6, pp. 1277–1284, Nov./Dec. 2008.
[81] O. Ersoy, C. Hurter, F. V. Paulovich, G. Cantareiro, and A. Telea, "Skeleton-based edge bundling for graph visualization," *IEEE Trans. Vis. Comput. Graphics*, vol. 17, no. 12, pp. 2364–2373, Dec. 2011.
[82] D. Guo and X. Zhu, "Origin–destination flow data smoothing and mapping," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 2043–2052, Dec. 2014.
[83] O. D. Lampe and H. Hauser, "Interactive visualization of streaming data with kernel density estimation," in *Proc. IEEE Pac. Vis. Symp.*, 2011, pp. 171–178.
[84] M.-J. Kraak, "The space-time cube revisited from a geovisualization perspective," in *Proc. 21st Int. Cartogr. Conf.*, 2003, pp. 1988–1996.
[85] "Colorbrewer 2.0." [Online]. Available: http://colorbrewer2.org/
[86] A. W. Rivadeneira, D. M. Gruen, M. J. Muller, and D. R. Millen, "Getting our head in the clouds: Toward evaluation studies of tagclouds," in *Proc. ACM SIGCHI*, 2007, pp. 995–998.
[87] "Wordle." [Online]. Available: http://www.wordle.net/
[88] J.-D. Fekete and C. Plaisant, "Excentric labeling: Dynamic neighborhood labeling for data visualization," in *Proc. ACM SIGCHI*, 1999, pp. 512–519.
[89] P. Goffin, W. Willett, J.-D. Fekete, and P. Isenberg, "Exploring the placement and design of word-scale visualizations," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 2291–2300, Dec. 2014.
[90] A. Inselberg and B. Dimsdale, "Parallel coordinates," in *Human–Machine Interactive Systems*. New York, NY, USA: Springer-Verlag, 1991, pp. 199–233.
[91] F. Wang *et al.*, "A visual reasoning approach for data-driven transport assessment on urban road," in *Proc. IEEE Conf. Visual Anal. Sci. Technol.*, 2014, pp. 103–112.

[92] S. Rinzivillo *et al.*, "Visually driven analysis of movement data by progressive clustering," *Inf. Vis.*, vol. 7, no. 3/4, pp. 225–239, Jun. 2008.
[93] M. VanDaniker, "Visualizing real-time and archived traffic incident data," in *Proc. IEEE Int. Conf. Inf. Reuse Integr.*, 2009, pp. 206–211.
[94] W. Zeng, C.-W. Fu, S. Arisona, A. Erath, and H. Qu, "Visualizing mobility of public transportation system," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 1833–1842, Dec. 2014.
[95] J. Aurisano, M. Cha, and M. Snyder, "User-driven predictive visual analytics on multivariate, spatio-temporal incident reports," in *Proc. IEEE VIS Workshop Vis. Predictive Anal.*, Paris, France, 2014.
[96] N. Andrienko, G. Andrienko, and S. Rinzivillo, "Experiences from supporting predictive analytics of vehicle traffic," in *Proc. IEEE VIS Workshop Vis. Predictive Anal.*, Paris, France, 2014.

Wei Chen received the Ph.D. degree from Fraunhofer Institute for Graphics, Darmstadt, Germany, in 2002. His Ph.D. advisors were Prof. Q. Peng and Prof. G. Sakas. He is currently a Professor at the State Key Laboratory of Computer Aided Design and Computer Graphics, Zhejiang University, Hangzhou, China. He has published more than 60 papers in international journals and conferences. Prof. Chen served as a Steering Committee Member of the IEEE Pacific Visualization, the Conference Chair of the IEEE Pacific Visualization 2015, and a Paper Co-Chair of the IEEE Pacific Visualization 2013. For more information, please refer to http://www.cad.zju.edu.cn/home/chenwei/.



Fangzhou Guo is currently working toward the Ph.D. degree with the State Key Laboratory of Computer Aided Design and Computer Graphics, Zhejiang University, Hangzhou, China. His research interests include visualization and visual analytics.

Fei-Yue Wang (S'87–M'89–SM'94–F'03) received the Ph.D. degree in computer and systems engineering from Rensselaer Polytechnic Institute, Troy, NY, in 1990.

In 1990, he joined the University of Arizona, Tucson, AZ, USA, where he became a Professor and the Director of the Robotics and Automation Laboratory and the Program for Advanced Research in Complex Systems. In 1999, he founded the Intelligent Control and Systems Engineering Center, Chinese Academy of Sciences (CAS), Beijing, China, under the support of the Outstanding Oversea Chinese Talents Program. Since 2002, he has been the Director of the Key Laboratory on Complex Systems and Intelligence Science, Institute of Automation, CAS. From 1995 to 2000, he was the Editor-in-Chief of the *International Journal of Intelligent Control and Systems* and the *World Scientific Series in Intelligent Control and Intelligent Automation*. His research interests include social computing, web science, complex systems, computational intelligence and intelligent control.

Dr. Wang is a member of Sigma Xi; an elected Fellow of the International Council on Systems Engineering, the International Federation of Automatic Control, the American Society of Mechanical Engineers, and the American Association for the Advancement of Science; an Association for Computing Machinery (ACM) Council Member at Large; and Vice President and Secretary-General of the Chinese Association of Automation. He has served as Chair of more than 20 IEEE, ACM, and Institute for Operations Research and the Management Sciences conferences. He was the President of the IEEE Intelligent Transportation Systems Society from 2005 to 2007; the Chinese Association for Science and Technology, USA, in 2005; and the American Zhu Kezhen Education Foundation from 2007 to 2008. He is currently the Editor-in-Chief of the IEEE INTELLIGENT SYSTEMS MAGAZINE and the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS. He was the recipient of the National Prize in Natural Sciences of China in 2007 and the Outstanding Scientist Award by the ACM for his work in intelligent control and social computing.