

TripVista: Triple Perspective Visual Trajectory Analytics and Its Application on Microscopic Traffic Data at a Road Intersection

Hanqi Guo^{1,2} Zuchao Wang¹ Bowen Yu¹ Huijing Zhao¹ Xiaoru Yuan^{1,2*}

¹) Key Laboratory of Machine Perception (Ministry of Education), and School of EECS, Peking University, Beijing, P.R. China
²) Center for Computational Science and Engineering, Peking University, Beijing, P.R. China

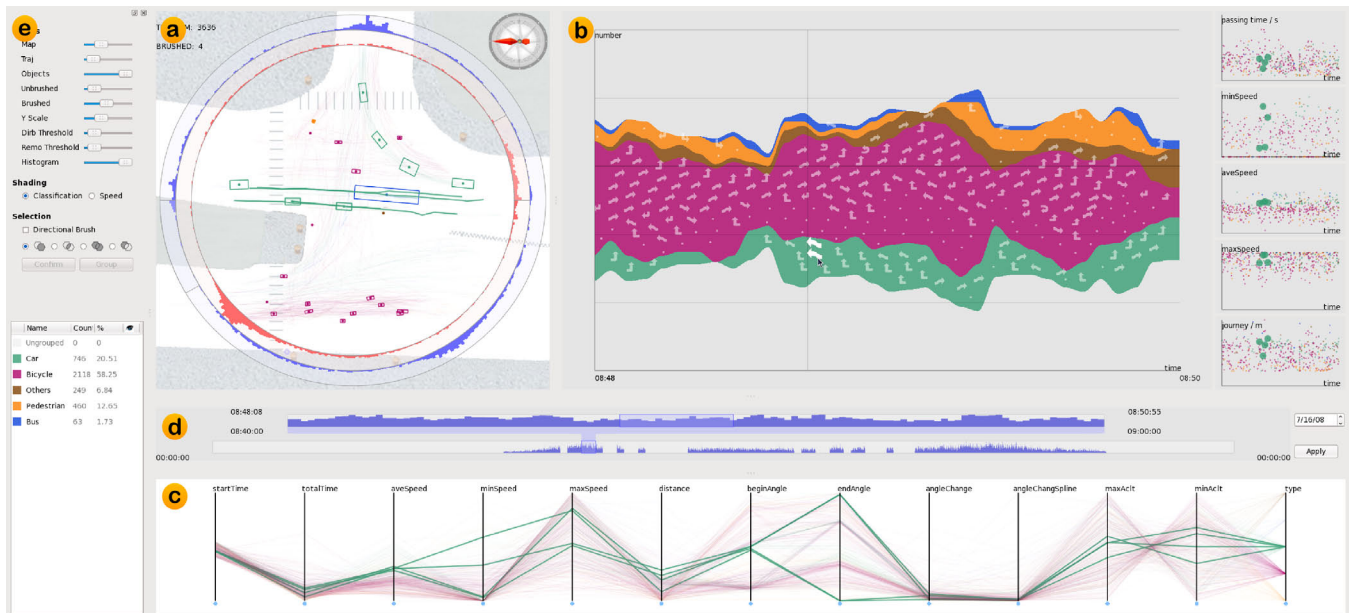


Figure 1: Interface of Triple Perspective Visual Trajectory Analytics (TripVista) visualizing traffic trajectory data at a road intersection. (a) Spatial traffic view showing geometrical trajectory information; (b) Temporal views of ThemeRiver and scatterplots; (c) Parallel coordinates plot showing multiple properties of the multi-dimensional data; (d) Time sliders for two-level time range selection; (e) Control panel for system parameter settings and data classification.

ABSTRACT

In this paper, we present an interactive visual analytics system, Triple Perspective Visual Trajectory Analytics (TripVista), for exploring and analyzing complex traffic trajectory data. Users are equipped with a carefully designed interface to inspect data interactively from three perspectives (spatial, temporal and multi-dimensional views). While most previous works, in both visualization and transportation research, focused on the macro aspects of traffic flows, we develop visualization methods to investigate and analyze microscopic traffic patterns and abnormal behaviors. In the spatial view of our system, traffic trajectories with various presentation styles are directly interactive with user brushing, together with convenient pattern exploration and selection through ring-style sliders. Improved ThemeRiver, embedded with glyphs indicating directional information, and multiple scatterplots with time as horizontal axes illustrate temporal information of the traffic flows. Our system also harnesses the power of parallel coordinates to visualize the multi-dimensional aspects of the traffic trajectory data. The above three view components are linked closely and interactively

*e-mail: {hanqi.guo, zuchao.wang, bowen.yu, zhaohj, xiaoru.yuan}@pku.edu.cn

to provide access to multiple perspectives for users. Experiments show that our system is capable of effectively finding both regular and abnormal traffic flow patterns.

Keywords: Visual analytics, Linked view, Spatiotemporal, Multi-dimensional data visualization, Traffic visualization

Index Terms: I.3.6 [Computer Graphics]: Methodology and techniques—Interaction Techniques; H.5.2 [Information Interfaces and Presentation]: User Interfaces—Graphical User Interfaces (GUI)

1 INTRODUCTION

With rapid economic growth, large increases in both motorization and urbanization have been witnessed in many countries since the beginning of the last century. As a consequence, the considerably increased number of automobiles in many cities has quickly congested transportation pathways. Heavy traffic may give rise to air pollution, aggravate time consumption of people who need to commute to work or school, pose a significant safety risk, and even exacerbate feelings of inequities in the society. In late August, 2010, the world witnessed the longest traffic jam in China [34]. Road construction, traffic accidents and breakdowns created a 60-mile, 11-day bumper-to-bumper gridlock on National Expressway 110 between the capital Beijing and inner Mongolia. It is indisputable that transportation has become a key issue in city development and management in today's world, which leads to the urgent require-

ment of advanced technologies for us to monitor, model, and optimize traffic flows in our transportation systems.

Recently, much work has been devoted to improving transportation systems with the state-of-the-art information technologies. Sensing devices, such as RFID tags, video recorders, laser scanners, GPS tracking units and cellphones have been utilized to collect various data and construct intelligent transportation systems. With more and more data available, one major challenge coming into our sight is how to effectively analyze the traffic data we have gathered and extract insights. Many researchers have focused on macro simulation data or observations in order to find network bottlenecks. As a matter of fact, it is also essential to analyze and monitor the micro behaviors of traffic, find out the origin of accidents and jams, and evaluate traffic light and crossing configurations for potential adjustments. In addition to the traffic scenario that can be generated by micro-simulation software like Paramics [12] and VisSim[13], data collected from real traffic flow is vital to research as there are many features and exceptions of real-life situations that cannot be well modeled. Micro traffic data is a collection of object movements, including the position, speed, size and other properties of the vehicles and pedestrians. The movement of an object is usually described as a specific trajectory.

In this work, we analyze the traffic data at a road intersection collected through several laser scanners and some other auxiliary devices, which is a typical microscopic traffic dataset. The dataset is generated from the raw point cloud by post processing techniques [37]. Laser-scanned data provides a possibility of studying micro behaviors of individual vehicles and pedestrians. Much more information can be uncovered from such detailed data in comparison with data obtained through video or other traditional technologies, if appropriate analytic tools are provided. However, such datasets are challenging for various reasons. First, at a busy road intersection, thousands of moving objects can be captured in a short period of time. It is not a trivial task visualizing and analyzing such datasets efficiently and conveniently. Second, the collection of the point clouds inherently includes noises, containing short, incomplete, even nonsense trajectories that can hardly be filtered out by automatic algorithms. Therefore new approaches are demanded to deal with the magnitude of data and to reduce the influence of the inherent noises.

In this paper, we design a visual analytics system, Triple Perspective Visual Trajectory Analytics (TripVista), for exploring microscopic traffic data, as illustrated in Figure 1. The developed visualization system enables the user to investigate trajectories from different angles, including spatial, temporal and multi-dimensional perspectives. In our system, by taking advantages of the linked views, the user can perceive the underlying features of the dataset and filter out noises and irrelevant trajectories for further investigation of interesting cases. Experiments show that our system is capable of effectively finding regular patterns and anomalies of traffic flows.

The remainder of this paper is organized as follows. First the related works are reviewed in Section 2. Our design philosophy, together with the data to be visualized, is given in Section 3. Detailed description of the proposed visual analytics system is in Section 4, followed by several analysis results from the traffic data presented in Section 5. Important issues of the work and possible improvements are discussed in Section 6 before the paper is concluded in Section 7.

2 RELATED WORKS

Traffic data acquisition, collection and processing have been widely studied in intelligent traffic system research. Micro traffic data can be obtained by simulation from existing software like VisSim [13] and Paramics [12]. In recent years, laser scanners and camera-based methods have been developed to capture vehicle data, thus

moving objects can be detected and tracked to estimate their status parameters, including location, speed and direction at each time instance [37]. The advent of GPS and cellphone tracking methods also makes those properties more accessible [24, 10].

Trajectory and movement data have been studied with various approaches, including visual analysis [3], machine vision [31], clustering [5], feature extraction [4] and movement pattern taxonomy [14]. Visual analysis tools enable interactive and intuitive data exploration. Andrienko and Andrienko [2] investigate ways of using aggregation for visual analysis of movement data. A variety of visualizations and interaction techniques are designed to represent results of aggregations and enable comprehensive exploration of the data. Mosaic diagrams are suggested for the exploration of cyclical traffic patterns and directional bar diagrams are proposed for the study of movements in different directions. Visualizations play vital roles in finding significant locations, extracting tracks and exploring movement dynamics [1]. Anomalies and other activities can also be detected and analyzed by exploiting the intuition and experience of security and surveillance experts through an easy-to-use visual feedback loop [21]. Viewing datasets from different perspectives is of high importance to trajectory data investigation. Movement data has been studied in geospatial and spatiotemporal visualization systems [22, 36]. Those systems provide connections between different perspectives to enhance the capability of visualizations. Slingsby [30] proposes a treemap cartography method to show spatiotemporal traffic patterns. To select a few interesting trajectories from a large number, Bouvier and Oates [9] suggest staining and Hurter [19] proposes a brush-pick-drop interaction scheme. Their methods are general for 2D trajectory data, but with limited perspectives provided. In our work, we provide comprehensive perspectives for data exploration. Our system mainly focuses on specific types of data, such as traffic data collected at a road intersection.

Several visualization metaphors have been included in our system. Parallel coordinates [20] have been developed for multi-dimensional visualization. The ThemeRiver metaphor introduced by Harve et al. [18] provides an intuitive way for time-varying data visualization with clustering information. Byron et al. [11] have a detailed discussion on the geometry and aesthetics of ThemeRiver as one kind of stacked graph. Wei et al. [33] develop the method of putting labels into the space midst the ThemeRiver in order to show text information on email contents. In our work we embed glyphs into ThemeRiver to illustrate time-varying characteristics as well as directional patterns. Glyph representations have been commonly applied in information visualization [32]. Visualization tools can provide interfaces for visually-driven data clustering [28]. Schreck et al. [29] propose a visual-interactive monitoring and controlling framework extending the basic Kohonen Feature Map algorithm for trajectory clustering. Coordinated views have been widely applied to provide effective visualizations and user interactions [8, 16]. In our system, visualizations of spatial, temporal and multi-dimensional perspectives are linked together to provide visual analytics from multiple aspects simultaneously.

3 OVERVIEW

In this section, we first describe the traffic dataset to be explored by our proposed visual analytics system. Then the design philosophy of the proposed system is introduced.

3.1 Data Description

In this work, we focus on a microscopic trajectory dataset collected at a road intersection. The data is captured with roadside laser scanners employed to profile the road conditions horizontally from different viewpoints [37]. The traffic direction and the traffic light configuration of the cross are shown in Figure 3(a). It is a T-shape crossroad with a joint one-way road and another multi-lane two-

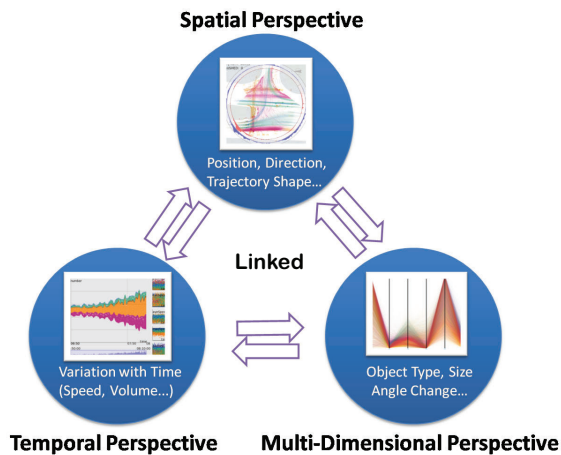


Figure 2: Design Philosophy of Triple Perspective Visual Trajectory Analytics (TripVista). Spatial, temporal and multi-dimensional perspectives are closely linked and respectively represented by different metaphors.

way road. In part of the two-way road, lanes with different directions are separated with a safety island. The compass indicates the north direction on the data map. The contour points of moving objects are captured on a horizontal plane at a scanning rate of 26ms per frame. When the objects entered the intersection, they were detected, tracked and further classified as cars, buses, bicycles, pedestrians and others. The dataset contains 209,426 trajectories represented by consecutive sampled points recorded over two days. There are 33,362,651 sampled points in total and each point has the attributes *position*, *speed*, *direction* and *timeInstance*. This dataset collected from a real situation is very challenging to be analyzed due to its noisy nature and uncertainties caused by occlusions. The noises resulting from tracking confusions and errors are mainly in the form of those with an impossibly small fraction of movement. We filter out the trajectories that have very short passing distance (<10 meters) or time (<1 second). Through a carefully designed visual analytics system, we enable direct exploration and manipulation on such a dataset.

3.2 Design Philosophy

Based on the nature of the traffic trajectory data, we design a visual analytics system, Triple Perspective Visual Trajectory Analytics (TripVista) with three perspectives embedded: spatial, temporal and multi-dimensional respectively to study microscopic patterns and discover abnormal behaviors in this data. As illustrated in Figure 2, each perspective has its own visual representation. They are closely linked to provide powerful visual exploration capabilities.

Spatial Perspective Spatial information is essential for understanding object movements in physical space. To identify certain types of object movements in traffic flows, flexible and thorough exploration of the geometrical information is indispensable. In our design, spatial perspective (the traffic view) provides the intuitive geometrical information.

Temporal Perspective Traffic movements are typically time-varying data. Observation of the traffic variation along the temporal axis is vital for identifying corresponding patterns and features, and especially for making operational decisions. The temporal perspective of the visualization is able to provide both an overview spanning a long time period and details at a specific time point. We have employed ThemeRiver with glyphs together with scatterplots to illustrate temporal features of the dataset.

Multi-dimensional Perspective In addition to the spatiotemporal

properties, traffic trajectory data has multi-dimensional attributes such as *object type*, *path length* and *direction*. Extra information can also be derived from the raw data. For example, *minimum speed* and *maximum speed* of each individual track can be extracted from original speed information. We include original and derived dimensions in the parallel coordinates plot for interactive user manipulation. The multi-dimensional perspective also helps to filter out data noises, and to find interesting patterns.

Linking all the above three perspectives for analysis is a design of consequence. Any separation would lead to visualization gap and result in information loss or an incomplete understanding of the data. Duplex interactions among all views have been established.

4 INTERFACE

The interface of the system consists of the multiple coordinated views shown in Figure 1: the traffic view for displaying spatial information, the ThemeRiver with embedded glyphs and scatterplots for showing temporal variations, and the parallel coordinates for multi-dimensional visualization. Two-level time sliders supporting quick time range selection are integrated to augment data exploration. In the following, we will introduce design details of each view in our visual analytics system.

4.1 Traffic View

The major function of traffic view is to directly display spatial information by rendering each trajectory as a polyline. Each trajectory is drawn according to its scanned position. The traffic view shows the accumulation of all trajectories by blending the tracks as semi-transparent polylines (Figure 3(b)). This gives the user an intuitive overview of the data over the selected time span. The color of each line indicates object type: *pedestrian*, *bus*, *car*, *bicycle* or *others*. Alternatively, considering the fact that speed information is also essential to pattern study, we provide another shading mode called *speed*, in which one trajectory is drawn with gradually changing colors. Red color denotes low speed while green denotes higher speed as shown in (Figure 3(c)). A map of the intersection is displayed as background. The user may recognize general clusters of trajectories through an intuitive perception of the plot. Box-like representations in the view denote moving objects at a specified time instance (determined interactively in the temporal views or by keyboard play/rewind hotkeys).

Convenient selections are integrated into the traffic view. In addition to normal brushing, the user may apply directional brushing to pick up trajectories with a specific shape by sketching. Furthermore, ring-style sliders are overlaid on top of the trajectory paths. The inner ring slider is used to select the entrance direction, while the outer ring is for exiting angle filtering. By adjusting the slider bars, the user can specify patterns with a specific entrance and exit range. Histograms showing the density of the trajectories at different angles are mapped along the ring circumference, which provides additional information concerning the traffic under investigation. When the mouse hovers over the histograms, the corresponding region will be shown in an expanded mode to display more details. The histogram information of selected trajectories will be highlighted with a darker color (Figure 3(b)). Rotation of the view is also supported for the user's convenience.

4.2 ThemeRiver Embedded with Glyphs

Although the standard ThemeRiver can bring forth flow volume information, it alone cannot show details of direction-related traffic patterns. We design a ThemeRiver view with embedded glyphs in TripVista to display directional information. Arrow-shape glyphs representing specific directional movement of objects are integrated into the ThemeRiver. The ThemeRiver view supports convenient user interactions, e.g., mouse-hovering highlighting, glyph brushing and zooming. These interactions can help the user get intuitive

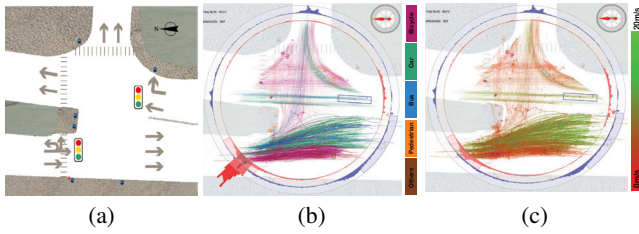


Figure 3: Traffic View: (a) The map of the road intersection where the data is collected. The arrows on the map indicate the permitted traffic directions and the traffic light configurations; (b) Traffic view - colored according to object types; (c) Traffic view - colored according to speed variation.

information both on the flow volume and the flow direction of each traffic object type. By browsing along the time axis, the user may easily compare different groups of trajectories and simultaneously identify patterns through the embedded glyphs. In terms of our design philosophy, the ThemeRiver with embedded glyphs presents a way of connecting different perspectives such as direction (spatial), volume statistics (multi-dimensional) and its time-varying features (temporal) are coordinately visualized.

In our implementation of the ThemeRiver, two layers are computed: standard ThemeRiver as background, and added glyphs. Harve’s algorithm [18] is adopted to draw the background ThemeRiver (Figure 4(a)). The glyphs are overlaid on the background ThemeRiver according to the following criteria: 1) the glyphs should clearly and faithfully represent local characteristics of each traffic type; 2) the glyphs should be distributed evenly; 3) each glyph should be placed inside one specific river for aesthetics and unambiguity; 4) local patterns of glyphs should remain consistent when zoomed in and out. To determine the positions of glyphs, Fast Hierarchical Importance Sampling [26] with uniform importance is applied. As this sampling method has the blue noise property, the sampled points will be distributed randomly and evenly. The points outside the rivers or intersecting the boundaries are discarded. The red dots in Figure 4(b) show the computed candidate glyph positions. To decide the type of the glyph at each point, every river is subdivided into several subrivers according to the directional clustering of trajectories. Each sampled point is placed exactly inside one subriver (Figure 4(c)). Thus one type of glyph represents one directional pattern. Due to the fact that the directional information at a road intersection is naturally described by entrance and exit, for simplification we use discrete encoding for the directional patterns and glyph generations rather than a fully-developed clustering algorithm based on trajectory comparisons. The details of this process will be discussed in Section 6.2. After the types of glyphs are determined, the enhanced ThemeRiver is drawn with corresponding glyphs rendered at each sampled position, as shown in Figure 4(d). When zoomed in or out, the sampling frequency of the time points for flux calculation will adapt to the current observation level. The changes of river contours in this procedure are usually minor so that the drawing coherence is maintained. We use the local densities of glyphs to show the importance of the direction pattern feature instead of size [33]. Each traffic trajectory has a certain group of corresponding glyphs, which guarantees the uniqueness of representation and also enables lasso selections in the ThemeRiver.

Five scatterplots are also included in the temporal view on the right of the ThemeRiver to enhance TripVista, where each point represents one trajectory path. The horizontal coordinate of each point in the plot is determined by the occurring time of the trajectory. Vertical axes are the total passing time, the minimum/maximum/average speed, and the moving distance of the objects in each plot respectively. To balance point density and maximize space utilization, we performed logarithmically scaled distortion transformation on the vertical axes [27] in addition to the linear

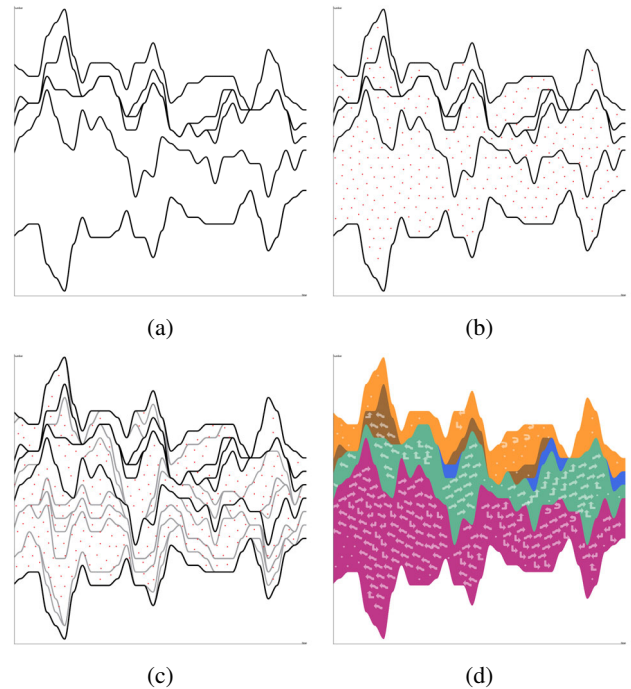


Figure 4: Algorithm Illustration for glyph embedding in ThemeRiver: (a) Original ThemeRiver; (b) Possible glyph positions determined by Fast Hierarchical Importance Sampling; (c) Each river is subdivided and the same type of glyphs are located in the same subriver; (d) Resulting ThemeRiver with embedded glyphs.

mapping. Scatterplots serve as a medium between the temporal and the multi-dimensional perspective. The user can brush out a group of trajectories quickly based on their occurring time or clustering in the scatterplots.

4.3 Parallel Coordinates

Parallel coordinates are integrated into TripVista to support multi-dimensional data exploration. As one of the most widely applied methods of multi-dimensional visualization, parallel coordinates plot amplifies the system’s ability of viewing data in a multi-dimensional sense. We set the dimensions of parallel coordinates to the entrance time, passing time/distance, minimum/maximum/average speed, start/end bearing, angle change (for both raw and preprocessed trajectories), maximum/minimum acceleration and object type. Through parallel coordinates, the user can perceive more features of trajectories rather than be limited to basic spatiotemporal properties. The user can clearly see singularities among the common patterns if they have any distinctive properties. Generally, brushing on the parallel coordinates offers a convenient way to select those trajectories with interesting features. Noises can also be filtered out with parallel coordinates.

4.4 User Interactions

Each view in TripVista supports convenient interactions. Besides click and brush, pertinent interactions such as directional brush and ring sliders in the traffic view are provided. A free-form lasso is automatically recognized as a rectangle when the moving distance of the user’s mouse is close to the straight-line distance between the beginning and ending points of the lasso [25]. Selection results of any of the views support intersection, union and subtraction operations. The user may apply *group* operations to the current selection, e.g. they may create or modify groups through the control panel, which provides a way for user-defined classification. User-defined

color schemes for groups are supported for multi-purpose suitability. The two time sliders can be used to select a target time range at lower and higher levels, to which flow volume histograms are mapped for a quick overview of distributions. In detail, the lower-positioned time slider performs on the higher level — enabling the selection of time spans in tens of minutes or hours from the traffic on a single day (July 16th, 2008 in Figure 1), while the upper-positioned time slider functions at a lower level — enabling the selection of time ranges in seconds and minutes. The ThemeRiver view may be switched to scatterplots by moving the thumbnails of scatterplots in a drag-and-drop manner in order to assign the scatterplots more space for display when necessary. We leave the decisions on the parameters including rendering transparency of trajectories and objects, size of brushed and unbrushed points in the scatterplots, scale of histograms, etc. to the user.

5 VISUAL ANALYSIS RESULTS

By using TripVista, not only can the user obtain overview information of traffic flow over a long time period, but they may also drill down to traffic trajectory details and discover interesting micro behaviors or patterns through interactions. Three cases are demonstrated in this section to show how TripVista helps analyze traffic flow data from integrated spatial, temporal and multi-dimensional perspectives. Our results illustrate the importance of linked perspectives, which is the key part of our design philosophy.

5.1 Case 1: Investigate Specific Behaviors

The first case is an example of recognizing special spatial patterns in the traffic view with the advantage of ring sliders and the directional brush.

TripVista is capable of identifying different types of traffic trajectory patterns according to differences in their geometry. Dense rendering line bundles in the traffic view provide an intuitive visual cue of the directions of the trajectory groups.

In the dataset we are working on, drivers from the low-left main street are allowed to make a U-turn at the intersection, as illustrated in Figure 3(a). We are interested in investigating how frequently such U-turn actions occurred. With TripVista, we first filter out pedestrians and bicycles by brushing the dimension of *type* in the parallel coordinates accordingly. The U-turn traffic then can be immediately isolated by adjusting the ring slider bars to the corresponding positions and ranges as shown in Figure 5(a). Alternatively, directional brush sketched by the user can also disclose a similar result as illustrated in Figure 5(b).

The selected trajectories are highlighted accordingly in other views, which enables the user to investigate other properties at the same time in addition to the geometrical information. For example, the user can explore the scatterplots or the parallel coordinates for information of vehicle speed. The visualization shows that the average speeds of these trajectory patterns are generally low and the minimum speed values are nearly zero. There are only four cases in which the minimum speed of the vehicle is above 2 m/s. This indicates that most cars drive cautiously and make a stop during the execution of a U-turn. Our system also provides play-back function of the selected scenes. Figure 5(c) shows individual moving cars with box-like representations after the time span is narrowed to a very short period. The system also provides information on the selected traffic volume. Among the total number of 32,777 moving objects during the selected time span, cars and buses together account for 9,421, as illustrated in the top-left corners of the figures. The number of trajectories selected by ring sliders is only 85. This indicates that there were only 85 cars that made a U-turn in that time period. The directional brush gives the smaller number of 42 due to more restricted shape matching.

This example demonstrates how to use TripVista to discover and investigate specific trajectory patterns. The process usually starts

with spatial understanding through brushing. Then the user can go into more detailed information, such as speed, object type and others through the linked views. We emphasize the internal connections between space and other perspectives, as well as the natural perceptual progressing from one to another.

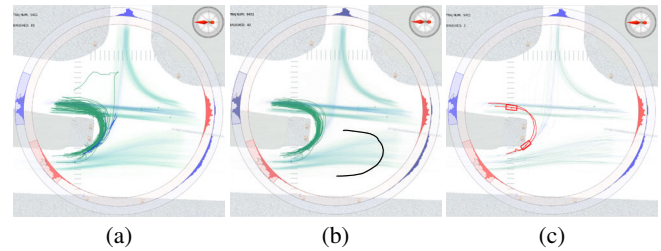


Figure 5: Investigation of U-turn Patterns (a) Select the U-turn trajectories with ring sliders; (b) Select the U-turn trajectories using directional brush; (c) Replay the scene (U-turn patterns in red).

5.2 Case 2: Find Patterns and Violations

The second case focuses on regular pattern recognition and violation detection.

Through TripVista, the user can perceive traffic flow patterns in multiple angles. The regular traffic light patterns can be discovered through the temporal view, which can be illustrated by showing time variation of the vehicle traffic volume. Such an intermittent pattern can hardly be observed without the help of the temporal view.

We first select the trajectories entering the intersection from the lower-left road. The system then splits and regroups the paths into two groups according to their exit directions, which can be done with the ring sliders or the directional brush as we have described previously. The result is shown in Figure 6.

It is clear in the glyph-embedded ThemeRiver that the left-turn trajectories (brown river with left-to-top arrows) and the straight trajectories (blue river with left-to-right arrows) have different temporal patterns. The difference is derived from the underlying traffic light regulation. The left-turn traffic light turns green following the go-forward traffic light turning into green with a short lag period for the pedestrians to cross the road. Here embedded glyphs help the user quickly identify the flow direction within a certain river (traffic type). It can also be perceived in the ThemeRiver view that the left-turn traffic flow has smaller volume. Through the rivers the proportion variation of the two flows in a single day can be further investigated.

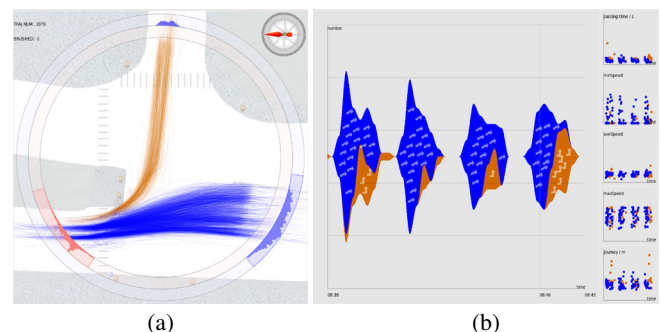


Figure 6: Regular Traffic Patterns (a) Group the concerned trajectories according to their exits; (b) Selected trajectory groups have different volumes as shown in the ThemeRiver.

In addition to understanding general patterns, TripVista also enables us to gain insights into the micro behaviors of pedestrians and

vehicles. Since the vertical road is one-way only as shown in Figure 3(a), any automobile retrograding violates the traffic rule.

After brushing these trajectories with the reversed direction of the one-way road and excluding some obvious noises, the trajectories of the offenders can be shown in Figure 7(a). The cyan trajectories representing cars indicate the violation events. Bicycle traffic flows are represented as trajectories in purple. The cyclists' behaviors are legal since the applicable regulation in this case only covers cars. TripVista can also expose another type of offence at this road intersection. When brushing the traffic from the right to the left, we can observe that a few vehicles turned right first at the intersection, and then made an immediate U-turn. At this road intersection, right turns are always allowed regardless of the status of the traffic lights, right turns are always allowed regardless of the status of the traffic lights. Those identified trajectories indicate drivers who wanted in effect to go left at the intersection during the red light time, which is illegal. Furthermore, with the replay function in the temporal view, we can watch the box-like representation executing the entire maneuver. The identified violation is shown in Figure 7(b).

This example illustrates the process of identifying both regular patterns and violations, which is one of the essential objectives of the TripVista design. Visualization methods in our design present information from different perspectives, thereby offering more opportunities for the user to perform in-depth investigation.

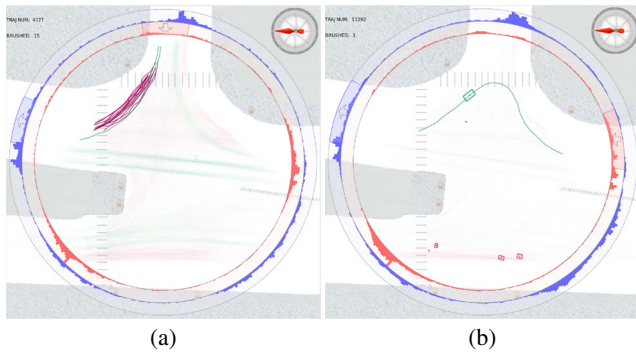


Figure 7: Traffic Violations (a) Wrong-way offenders; (b) Illegal turning pattern.

5.3 Case 3: Discover Hidden Information

In the third case, all views are used in conjunction to figure out interesting but usually hidden information like a chain of incidents among the densely mapped trajectories. This exemplifies that TripVista can be used to detect potentially dangerous cases, and even to help make conclusions concerning social behaviors.

In traffic flows, there often exist chain reactions in which the behavior of one object (vehicle or pedestrian) is the result of the abnormal behavior of other moving objects. For example, an abrupt stop of a moving vehicle would result in several following vehicles slamming on the brakes. However, such chain reactions are usually hidden among the data collected and cannot be directly identified. To discover such events, one needs to fully utilize all the perspectives including space, time and multi-dimension to view the data from multiple angles. TripVista offers the possibility of such complex investigation. The below example shows an actual finding of a dangerous moment that was initiated by a cyclist's violation behavior.

We calculate the *angle change* for each trajectory, which is the sum of the included angles (always positive) of consecutive speed vectors. Trajectories with smooth behaviors would have angle changes no more than 90 degrees except those performing a U-turn. An abnormal trajectory that exhibits sudden turns or evasion actions will have an angle change larger than 90 degrees. In this

case, we select the trajectories with larger angle changes with the parallel coordinates interface (Figure 8(a), (b)) and identify one trajectory of a bicycle passing the center of the intersection as shown in Figure 8(c). By replaying the scene in the traffic view, it can be observed that the large angle change actually resulted from the cyclist's evasive actions to avoid collision with a car. The car made a panic brake in front of the cyclist in that incident. It can be concluded that it is a dangerous event (Figure 8(d)). Besides the single violation, it is interesting to reveal through this example that violations usually appear in groups. Since the violating cyclist passed directly along the diagonal of the intersection, the travel distance is rather long. We brush similar behaviors by selecting long passing distances in one of the scatterplots in a short time span covering the event identified above. Then we find that there were several other bicycles going in a similar fashion at the time of the incident (Figure 8(e)). Such observation helps us identify social behaviors as violations tend to appear in group due to people's false sense of safety.

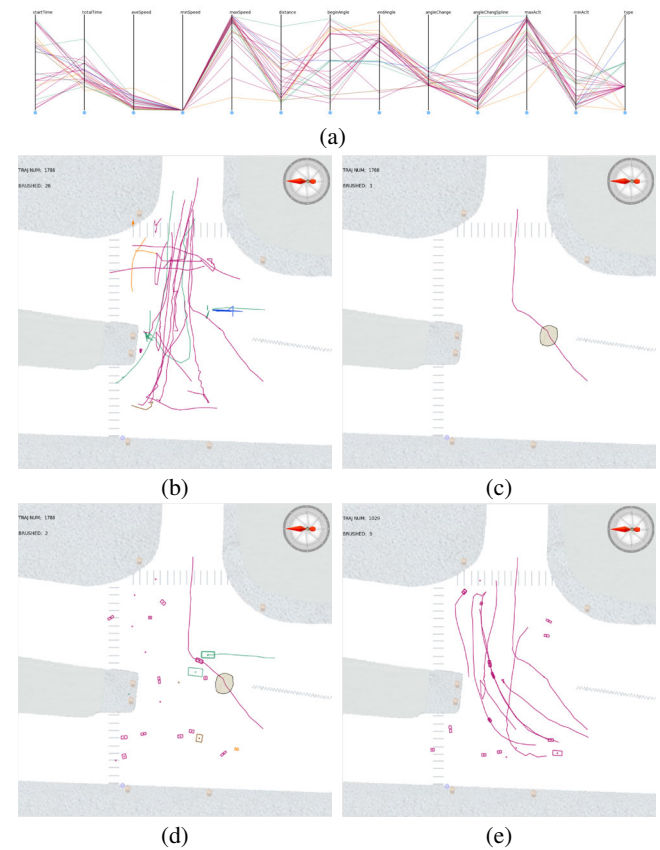


Figure 8: Dangerous Events Discovered in Dense Trajectories: (a) Selection of the trajectories with larger angle changes in the parallel coordinates; (b) Trajectories with larger angle changes in the traffic view; (c) One interesting trajectory identified in the traffic view; (d) A dangerous event: the cyclist narrowly escaped being hit by a car; (e) Other violation behaviors in groups.

6 DISCUSSION

In this section, we summarize a few aspects of the implementation details of TripVista. Important issues including automatic algorithm features and possible improvements to the system are discussed. User feedback is provided to further confirm the application value of the system to its potential users.

6.1 Implementation Details and Scalability

The system was developed in C++ with Qt 4.7, boost 1.43.0 and OpenGL 3.1. The program has been tested on a Dell T3500 workstation, with an Intel Xeon W3503 2.4GHz CPU, 2GB RAM and a NVIDIA GeForce GTX 275 graphics card with 896MB memory. User interactions can be performed smoothly at a frame rate of more than 20 frames/second if middle size data is load.

To mitigate the high demand on computing resources when visualizing large datasets, intermediate results of all views are saved in frame buffers and only necessary parts are updated when small changes are made by interactions. For example, the highlighting effects can be drawn on top of the existing buffer contents. Our system can run on regular workstations or PCs with general hardware settings.

To take advantage of a better user experience during interactions, we have also tested our system with a Wacom Cintiq 12WX display tablet as the input device. Interactions such as brushing and hovering can be performed more efficiently on the tablet with a pen through sketching than through a mouse and keyboard. The advantages of interactive features of TripVista can be further utilized on such devices to enhance data exploration.

6.2 Extension of Applicability

TripVista helps the user discover interesting patterns by including not only visualization representation techniques, but also semi-automatic algorithms such as the directional brush and angle change calculation.

More automatic algorithms can be integrated into our system to enhance its capability. For example, relative motion detection is an interesting topic in microscopic traffic analysis. We have discovered a case in which a vehicle narrowly avoided colliding with a bicycle in Section 5.3. It would be beneficial to automatically extract similar cases in the whole dataset. One possible solution is to search for matches of a predefined patterns [23], such as encounter or approximation [6]. Another is to search for behaviors similar to a given case. We have integrated into our system a prototype of a relative motion detection algorithm by searching the similar behaviors, taking both spatial and temporal information together with neighboring movement information into account. As shown in Figure 9, our system can detect a few similar cases in which moving bicycles and cars appeared very close. The user can further inspect the detected events and evaluate the degree of danger accordingly. The detection is also helpful for data correction. As shown in the last case of Figure 9, one bicycle is inside the space of a bus at a specific moment. This impossible scene results from inaccuracies of the data. Such findings are extremely valuable for data cleaning and validation.

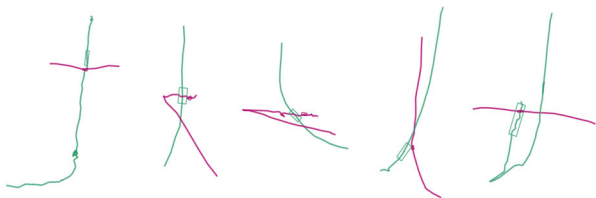


Figure 9: Several cases discovered by our prototype of relative motion detection algorithm, in which moving bicycles (purple) and cars (cyan) appeared very close. However, when speed is considered, none of them is as dangerous as the case in Section 5.3. It is noticeable that in the last case, the bicycle is inside the car. This impossible scene results from the inaccuracies in the data.

In terms of clustering techniques, our present system clusters the trajectories based on the entrance and exit information. It uses predefined glyphs to represent different patterns. This simple cluster-

ing strategy is appropriate and sufficient in the current case due to the fact that patterns, as well as abnormal behaviors at a road intersection have limited variation on possible movement directions. In the future, we would like to extend our approach to more complex traffic movements within larger road networks. More complicated phenomena should be considered, for example the question of how a congestion propagates through a network can potentially be visually analyzed. Some research has been done to cluster general movement, such as probabilistic trajectory clustering [17] and density-based trajectory clustering [15, 7]. Such automatic algorithms can be enhanced by visualization and interaction [28, 29, 5]. It is possible to include comprehensive clustering algorithms in our system. Since it is very difficult to predict all possible trajectory patterns, glyphs in our ThemeRiver can also be automatically generated as a visual summary of the trajectory shapes they correspondingly represent.

Views and interactions can be made adaptable to more complex conditions. For example, ring sliders can be modified to deal with road networks instead of one single road intersection. We are considering enhancing the current system by introducing ring slices, which are partial ring sliders with more freedom in shapes. Our system can also be extended to handle 3D scenarios, such as airline traffic data. The directional brush and ring sliders can be improved to support 3D selections.

6.3 User Feedback

We have consulted people who have expert knowledge in intelligent transportation system research for comments after internally testing our TripVista system.

The feedback is very encouraging. The highly supported interactions are greatly appreciated by the experts. Especially the powerful selection function as well as the animation replaying the scenes receive very positive comments. These features largely enhance the user experiences in traffic data exploration. In comparison with their traditional point cloud analysis tools, in which only limited functions such as point cloud plotting are supported, our system makes the whole exploration process much smoother. According to the feedback, our system is also good for demonstration. The visualization provides intuition for understanding the data.

Our visual analytics cases are also convincing to the domain experts. The analysis of the traffic flow patterns in Section 5.2 is considered very valuable to guide the optimization of traffic lights control. The discovery of the cyclist's case in Section 5.3 greatly arouses the experts' attention. The detected event is later further confirmed by their manual inspection on the initial laser-scanned point cloud data. The experts also indicate that the tablet demonstration of our system shows a data exploration process much more efficient and convenient than traditional methods, which they would like to apply in their work.

7 CONCLUSION AND FUTURE WORK

In this paper, we present a visual analytics system named TripVista for visualizing microscopic traffic trajectory data at a road intersection. Based on the triple perspective design philosophy, coordinated visualizations with convenient interactions are created in our proposed system. By simultaneously inspecting the time-varying trajectory data from spatial, temporal and multi-dimensional views, our system provides powerful tools for data comprehension and exploration. The establishment of close linking among each view is essential to the visualization effectiveness in our system.

As illustrated in the examples of traffic data analysis with TripVista, our system helps the user not only comprehend regular traffic patterns, but also discover abnormal behaviors. Some possible improvements are named in our user feedback. More automatic algorithms are expected, such as advanced clustering and relative motion detections. In the future, we would like to explore

the potential of TripVista by equipping it with improved automatic algorithms and additionally extending it to more complex data, such as traffic in a larger road network, road-free movement and 3D trajectory data. We also plan to integrate more powerful visualization techniques. For example, by introducing point representation in the parallel coordinates plots [35] in our system, we can enhance the ability to analyze high-dimensional aspects of the trajectory data.

ACKNOWLEDGEMENTS

The authors thank the anonymous reviewers for their invaluable comments and suggestions. This work is supported by National Natural Science Foundation of China Project No.60903062, Beijing Natural Science Foundation Project No.4092021, 973 Program Project No.2009CB320903, 863 Program Project 2010AA012400, Chinese Ministry of Education Key Project No.109001 and IJPL-09-016. The work of data collection was supported by NSFC Project No.60975061. The authors would also like to thank Jie Sha for precious comments and feedback.

REFERENCES

- [1] G. Andrienko and N. Andrienko. Exploration of massive movement data: a visual analytics approach. In *AGILE '08*, 2008.
- [2] G. Andrienko and N. Andrienko. Spatio-temporal aggregation for visual analysis of movements. In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology*, pages 51–58, 2008.
- [3] G. Andrienko and N. Andrienko. A visual analytics approach to exploration of large amounts of movement data. In *VISUAL '08: Proceedings of the 10th international conference on Visual Information Systems*, pages 1–4, 2008.
- [4] G. Andrienko, N. Andrienko, J. Dykes, S. I. Fabrikant, and M. Wachowicz. Geovisualization of dynamics, movement and change: key issues and developing approaches in visualization research. *Information Visualization*, 7(3):173–180, 2008.
- [5] G. Andrienko, N. Andrienko, S. Rinzivillo, M. Nanni, D. Pedreschi, and F. Giannotti. Interactive visual clustering of large collections of trajectories. In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology*, pages 3–10, 2009.
- [6] N. Andrienko, G. Andrienko, M. Wachowicz, and D. Orellana. Uncovering interactions between moving objects. In *Extended Abstracts of GIScience 2008*, pages 16–26, 2008.
- [7] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander. Optics: Ordering points to identify the clustering structure. In *SIGMOD'99*, pages 49–60, 1999.
- [8] A. Barsky, T. Munzner, J. Gardy, and R. Kincaid. Cerebral: Visualizing multiple experimental conditions on a graph with biological context. *IEEE Trans. Vis. Comput. Graph.*, 14(6):1253–1260, 2008.
- [9] D. J. Bouvier and B. Oates. Evacuation traces mini challenge award: Innovative trace visualization staining for information discovery. In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology*, pages 219–220, 2008.
- [10] Y.-J. Byon, B. Abdulhai, and A. Shalaby. Real-time transportation mode detection via tracking global positioning system mobile devices. *Journal of Intelligent Transportation Systems*, 13:161–170, 2009.
- [11] L. Byron and M. Wattenberg. Stacked graphs - geometry & aesthetics. *IEEE Trans. Vis. Comput. Graph.*, 14(6):1245–1252, 2008.
- [12] G. Cameron, B. J. N. Wylie, and D. McArthur. Paramics—moving vehicles on the connection machine. In *Supercomputing'94*, pages 291–300, 1994.
- [13] H. Chen, X. Zhang, and G. Liu. Simulation and visualization of empirical traffic models using vissim. In *Proceedings of IEEE International Conference on Networking, Sensing and Control 2007*, pages 879–882, 2007.
- [14] S. Dodge, R. Weibel, and A.-K. Lautenschütz. Towards a taxonomy of movement patterns. *Information Visualization*, 7(3):240–252, 2008.
- [15] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *KDD'96*, pages 226–231, 1996.
- [16] W. Freiler, K. Matkovic, and H. Hauser. Interactive visual analysis of set-typed data. *IEEE Trans. Vis. Comput. Graph.*, 14(6):1340–1347, 2008.
- [17] S. Gaffney and P. Smyth. Trajectory clustering with mixtures of regression models. In *KDD'99*, pages 63–72, 1999.
- [18] S. Havre, E. G. Hetzler, P. Whitney, and L. T. Nowell. ThemeRiver: Visualizing thematic changes in large document collections. *IEEE Trans. Vis. Comput. Graph.*, 8(1):9–20, 2002.
- [19] C. Hurter, B. Tissoires, and S. Conversy. Fromdady: Spreading aircraft trajectories across views to support iterative queries. *IEEE Trans. Vis. Comput. Graph.*, 15(6):1017–1024, 2009.
- [20] A. Inselberg. The plane with parallel coordinates. *The Visual Computer*, 1(2):69–91, 1985.
- [21] F. Janoos, S. Singh, O. Irfanoglu, R. Machiraju, and R. Parent. Activity analysis using spatio-temporal trajectory volumes in surveillance applications. In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology*, pages 3–10, 2007.
- [22] T. Kapler and W. Wright. GeoTime information visualization. In *Proceedings of the IEEE Symposium on Information Visualization*, pages 25–32, 2004.
- [23] P. Laube, S. Imfeld, and R. Weibel. Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science*, 19(6):639–668, 2005.
- [24] G. Leduc. Road traffic data: Collection methods and applications. Technical report, Institute for Prospective Technological Studies, EU, 2008.
- [25] M. J. McGuffin and I. Jurisica. Interaction techniques for selecting and manipulating subgraphs in network visualizations. *IEEE Trans. Vis. Comput. Graph.*, 15(6):937–944, 2009.
- [26] V. Ostromoukhov, C. Donohue, and P.-M. Jodoin. Fast hierarchical importance sampling with blue noise properties. *ACM Transactions on Graphics*, 23(3):488–495, 2004.
- [27] S. Potts and T. Möller. Transfer functions on a logarithmic scale for volume rendering. In *Proceedings of Graphics Interface 2004*, pages 57–64, 2004.
- [28] S. Rinzivillo, D. Pedreschi, M. Nanni, F. Giannotti, N. Andrienko, and G. Andrienko. Visually driven analysis of movement data by progressive clustering. *Information Visualization*, 7(3):225–239, 2008.
- [29] T. Schreck, J. Bernard, T. Tekusova, and J. Kohlhammer. Visual cluster analysis of trajectory data with interactive kohonen maps. In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology*, pages 3–10, 2008.
- [30] A. Slingsby, J. Wood, and J. Dykes. Treemap cartography for showing spatial and temporal traffic patterns. *Journal of Maps*, v2010:135–146, 2010.
- [31] J. Vijverberg, N. de Koning, J. Han, P. de With, and D. Cornelissen. High-level traffic-violation detection for embedded traffic analysis. In *ICASSP'07*, volume 2, pages 793–796, 2007.
- [32] M. O. Ward. A taxonomy of glyph placement strategies for multidimensional data visualization. *Information Visualization*, 1(3/4):194–210, 2002.
- [33] F. Wei, S. Liu, Y. Song, S. Pan, M. X. Zhou, W. Qian, L. Shi, L. Tan, and Q. Zhang. Tiara: a visual exploratory text analytic system. In *KDD'10*, pages 153–162, 2010.
- [34] M. Wines. Chinas 60-mile traffic jam is breaking up, 2010. <http://wheels.blogs.nytimes.com/2010/08/24/chinas-60-mile-traffic-jam-is-breaking-up/>.
- [35] X. Yuan, P. Guo, H. Xiao, H. Zhou, and H. Qu. Scattering points in parallel coordinates. *IEEE Trans. Vis. Comput. Graph.*, 15(6):1001–1008, 2009.
- [36] H.-S. Zhang, Y. Zhang, Z.-H. Li, and D.-C. Hu. Spatial-temporal traffic data analysis based on global data management using mas. *IEEE Transactions on Intelligent Transportation Systems*, 5(4):267–275, 2004.
- [37] H. Zhao, J. Cui, H. Zha, K. Katabira, X. Shao, and R. Shibasaki. Sensing an intersection using a network of laser scanners and video cameras. *IEEE Intelligent Transportation Systems Magazine*, 1(2):31–37, 2009.