Computational Visual Media
DOI 10.1007/s41095-xxx-xxxx-x Vol. x, No. x, month year, xx–xx

Review Article

A survey of urban visual analytics: domain problem, visualization, computation, and system

Zikun Deng[1], Di Weng[2,†] (✉), Shuhan Liu[1], Yuan Tian[1], Mingliang Xu[3,4], and Yingcai Wu[1,†] (✉)

Abstract Developing effective visual analytics systems demands careful considerations in characterizing domain problems and integrating visualization techniques and computational methods. Due to the explosive emergence of urban data and the rapid development of visualization techniques and computational methods, urban visual analytics has achieved remarkable success in tackling urban problems and providing fundamental services for smart cities. To further promote academic research and assist the development of industrial urban analytics systems, we comprehensively review the urban visual analytics studies from four perspectives, namely, domain problem, visualization, computation, and system. In particular, we identify 8 urban domains and 22 types of popular visualizations, analyze 7 types of computational methods, and categorize the existing systems into 4 types based on the integration of visualization techniques and computational methods. Finally, we conclude the summarization of the state-of-the-art progress with potential research directions and opportunities.

Keywords visual analytics; smart city; spatiotemporal data analysis; urban analytics; survey.

1 Introduction

Urban computing has achieved remarkable success in tackling many urban problems [208], such as traffic prediction [130], air quality forecasting [213], bike lane planning [54], transit route planning [171], and location selection [85]. However, since urban analytics is an interdisciplinary field, it is crucial to integrate domain knowledge and expertise in the analysis loop. Hence, urban visual analytics [210] was proposed and extensively studied to empower urban experts with the combination of intuitive data visualizations and fast computational methods, enabling the experts to visually and interactively perceive, explore, manipulate and reason with urban data [102].

When developing an urban visual analytics approach, practitioners like urban analysts and researchers may have the following four questions:

1. Which urban **domain problems** were solved or not solved yet by visual analytics?
2. What **visualization** techniques were applied to visually interpreting urban data?
3. What **computational methods** were employed in urban visual analytics to solve urban problems?
4. How to combine visualizations and computational models in existing **visual analytics systems**?

Without answers to these questions, researchers will find it difficult to obtain a big picture of the state-of-the-art progress of urban visual analytics. The lack of such a big picture prohibits urban analysts from quickly finding feasible approaches from prior studies when designing visual analytics solutions for urban problems.

The recent rapid development of urban visual

1 State Key Lab of CAD & CG, Zhejiang Univerisity, Hangzhou 310058, China. E-mail: {zikun_rain, shliu, ytian, ycwu}@zju.edu.cn.
2 Microsoft Research Asia, Beijing 100080, China. Email: diweng@microsoft.com.
3 School of Information Engineering, Zhengzhou University, Zhengzhou, China and Henan Institute of Advanced Technology, Zhengzhou University, Zhengzhou, China. E-mail: iexumingliang@zzu.edu.cn.
4 Henan Institute of Advanced Technology, Zhengzhou University, Zhengzhou, China.
† Yingcai Wu and Di Weng are the co-corresponding authors.
Manuscript received: 2014-12-31; accepted: 2015-01-30.

analytics urgently demands a comprehensive review to summarize how urban problems can be addressed effectively with visual analytics and reveal the future directions of urban visual analytics [5, 27, 210]. As for question 1, Chen et al.'s [27] and Andrienko et al.'s [5] surveys focused only on the traffic domain. Zheng et al.'s survey [210] discussed the general urban visual analytics approaches but was only from a data point of view and did not summarize the domain problems. As for question 2, Chen et al.'s [27] and Zheng et al.'s [210] surveys categorized visualizations based on whether it is temporal, spatial, or for other properties. However, many recent techniques were not included in these reviews, given that the latest review was published five years ago. As for question 3, none of the existing surveys can answer it. As for question 4, Zheng et al. [210] classified the systems that combine visualizations and computational methods into "data exploration and pattern interpretation" and "visual learning" based on the system output. Yet, how to combine visualizations and computational models, e.g., their respective roles and interrelationships, remains unclear.

In this study, we attempt to develop a systematic and comprehensive survey of urban visual analytics. By answering the four questions, this survey summarizes urban visual analytics studies from multiple perspectives, including domain problems, visualizations, computational methods, and systems. Since urban visual analytics is an interdisciplinary research topic, we investigated the prominent journals and conferences in multiple fields, including visualization, transportation, data mining, and geography. To avoid missing important papers, we generally followed the reference- and search-driven paper collection methods in Guo et al.'s study [52]. The collected papers were mainly published in four journals, namely, IEEE TVCG, CGF, IEEE TITS, and ACM TIST, and four conferences, namely, IEEE VIS, EuroVis, PacificVis, and ACM CHI. These papers range from 2007 to 2022. An interactive tool for exploring these papers is available at `https://urban-va-survey.github.io/`. Our contributions can be summarized from the following aspects.

- **Survey.** We present a systematic survey of urban visual analytics. This survey summarizes the significant progress urban visual analytics has made in the past few years comprehensively from four perspectives. This survey also indicates future research directions and opportunities.
- **Domain problem.** We categorize the domain problems studied in urban visual analytics into 8 domains. Such a taxonomy not only assists urban analysts in finding relevant visual analytics approaches for real-world urban problems, but also reveals the gap between the existing approaches and the urban problems that are yet to be solved.
- **Visualization.** We categorize the visualization techniques for the spatial, temporal and other properties in urban data into 22 types and demonstrate their usages in different analytical tasks in the hope of promoting the better usage and designs of urban data visualization.
- **Computation.** We categorize the computational methods that drive urban visual analytics into 7 types and report the usage of artificial intelligence.
- **System.** We conclude 4 types of urban visual analytics systems based on the integration of visualization techniques and computational methods and discuss the applicable scenarios and trends for each system type.

2 Domain problem

The existing surveys [5, 27, 210] focused on data and attempted to answer what could be done with different types of data. However, real urban analytics scenarios generally start with specific urban problems. In this section, we classify the urban domains that have been studied by visual analytics into 8 categories: traffic, environment, business, economics, public security, architecture, public service, and public opinion. Table 1 and 2 summarize the categories, specific problems, and relevant urban visual analytics studies.

Generally, the classification is made based on the papers' introduction, case studies and usage scenarios. Note that an approach can be applied to multiple domains and address different specific problems. Take VisCas [37] as an example – urban analysts can use VisCas to analyze the cascades of both air pollution and traffic congestion events. Hence, we put VisCas into the environment and traffic domains based on its applications indicated in the usage scenarios.

2.1 Traffic

The traffic domain is the most widely studied due to the explosive growth of mobility data in the past decade. Mobility data contains valuable knowledge because mobility is the most important manifestation of citizens' activities in urban space.

Human mobility. Some approaches are proposed specifically for studying human mobility. They are not

Tab. 1 Domain problems studied by the existing urban visual analytics approaches.

Domain	Problem		Papers
Traffic	Human mobility	Errors and uncertainty of trajectories	[25, 26, 63, 107]
		Mobility patterns	[9, 26, 28, 32, 49, 96, 98, 101, 108, 111, 115, 134, 149, 159, 161, 177, 198]
		Origin-Destination patterns	[6, 44, 69, 96, 109, 110, 146, 200, 202, 214]
		Human co-occurrence	[25, 29, 178, 212]
		Mobility semantics	[3, 71, 191, 199]
	Road network	Road network accessibility	[42, 68, 179]
		Road centrality	[61]
		Road correlation and causality	[38, 168]
		Intersection interchange behavior	[51, 160, 197]
		Tidal lanes analysis	[160, 211]
		Traffic situation understanding	[160, 161]
		Traffic zone division	[32, 67, 152]
	Congestion	Congestion monitoring	[72]
		Congestion discovery	[7, 8, 28, 132, 134]
		Congestion propagation	[167]
		Congestion events' cascades	[37]
		Congestion causal inference	[28, 132, 160]
		Congestion prediction	[72]
	Public transportation	Network accessibility	[11]
		Bus schedule analysis	[129]
		Interchange behavior	[197]
		System usage	[200]
		System efficiency	[105, 172, 196]
		Bus network optimization	[68, 105, 172]
		Shuttle bus planning	[3, 100]
	Traffic safety	Tunnel surveillance	[133]
		Vehicle monitoring	[135]
		Abnormal driver detection	[3, 92]
		Traffic violation identification	[51]
		Abnormal traffic pattern detection	[135]
	Autonomous driving	Traffic light detection evaluation	[47]
		Semantic segmentation evaluation	[55]
		Autonomous driving action evaluation	[66]
		Autonomous driving system evaluation	[57]
	Traffic volume forecasting & simulation		[12, 201]

limited to a specific problem but provide implications for a wide range of scenarios.

Mobility data collected by GPS is usually inaccurate owing to measurement errors and the low sampling rate. Map-matching [106, 194] is the first step toward obtaining clean mobility data. Lu et al. [107] assisted this process with visual analytics. Chen et al. [25] summarized five types of uncertainty in trajectory data and proposed a semi-automatic refinement method. In addition to GPS-based data, Chen et al. [26] worked on the uncertainty of the trajectories inferred from sparse geo-tagged posts.

Given the clean mobility data, various visual analytics approaches were developed for studying mobility patterns [9, 26, 28, 32, 49, 96, 101, 108, 115, 149, 159, 161, 177, 198], origin-destination (OD) patterns [6, 44, 69, 96, 109, 110, 146, 200, 202, 214], co-occurrences [25, 29, 178, 212], and mobility semantics [3, 63, 71, 191, 199]. These analyses supported the in-depth understanding of how citizens move within the urban space. Notably, some interesting studies attempted to infer the semantics of mobility using contextual information like the surrounding points of interest (POIs) [191, 199]. Such semantics forms the basis of convenient text-based mobility query in [3, 63].

Some other approaches involving human mobility focus on addressing domain-specific problems. Thus we categorize these approaches into other domains.

Road network. As the foundation of urban traffic, road networks have been analyzed in many visual analytics systems. These analyses can be divided into macro-, meso-, and micro-levels.

At the macro-level analysis, accessibility globally measures how fast a place can be accessed from another in a road network. Accessibility can be computed based on the vehicles running on the roads. To analyze accessibility across the entire road network, a density-map-based visualization [42] and a visual query-based approach [68] were developed. Wu et al. [179] also designed visualizations to reveal how the reachability boundary changes over time. Huang et al. [61] analyzed road centrality (i.e., importance) in the road network based on the PageRank algorithm.

The meso-level analysis studies the relationships among multiple roads. Roads can be grouped into zones based on the inherent traffic patterns [32, 67, 152]. The implicit correlation and causality of roads are visually investigated in [38, 168].

The micro-level analysis dives into the individual roads. Wang et al. [160] allowed users to assess roads based on the passing trajectories. Zheng et al. [211] and Wang et al. [160] studied the bi-directional traffic flow of individual roads, supporting the policy-making of tidal lanes. Road intersections can be investigated by visualizing traffic interchange behaviors [51, 160, 197].

Congestion. The most notorious problem of road traffic is congestion. Detecting and understanding congestion are critical to reducing traffic congestion. Congestion can be visually discovered based on the velocity of trajectories [7, 8, 28, 132, 134]. Lee et al. [72] further developed a visual analytics system for predicting traffic congestion. Congestion events can propagate or spread over space and time, leading to a large area of congestion. Many approaches were proposed for understanding such processes, shedding light on effective congestion control. Wang et al. [167] are the first to visually analyze the congestion propagation on the road network topology. Deng et al. [37] studied the implicit cascading processes of congestion events based on the spatiotemporal relationships of the events instead of the road network topology. Tailored visualizations were designed to support experts to infer the causes of congestions from multiple aspects [28, 132, 160].

Public transportation. Efficient, convenient, and comfortable public transportation can satisfy more travel demands, alleviating traffic congestion and improving cities. Representative public transportation methods include subways, shared bicycles, and buses. Visual analytics can facilitate the evaluation and optimization of public transportation systems.

Evaluation. The evaluation can be performed based on the timetable (or schedule), travel time, and the satisfaction of travel demands. Andrienko et al. [11] exploited the planned bus timetables to reveal how a transportation system connected urban space (i.e., accessibility). Considering the discrepancies between the actual and planned operations, Palomo et al. [129] compared the timetables against the actual service to diagnose the low efficiency, e.g., a serious delay.

A passenger's trajectory along with its interchange behaviors [197, 200] can be inferred from the OD data in a public transportation system. Thus, various metrics, such as waiting and transfer time, can be derived based on the passengers' ODs and visualized to understand the efficiency [105, 196]. Beyond the numeric metrics, Weng et al. [172] visually encoded the movement of bus passengers to discover the gaps between the current system and the travel demands.

Optimization. Experts can optimize the system after understanding and evaluating the public transportation systems. For example, Lorenzo et al.'s method [105] supported the comparison between the old and the new routes with respect to the derived metrics. Kamw et al.'s method [68] could identify the areas unreachable on foot and propose candidate bus stops in these areas. Liu et al. [100] designed interactive visualizations that allow users to visually evaluate potential bus stations and select ideal ones to create a shuttle bus route. Weng et al.'s method [172] integrated a Monte-Carlo tree search model [171] into a progressive and interactive route replacement procedure.

Traffic safety. Traffic safety can be improved via traffic monitoring, for which visual analytics provides an interactive and situation-aware environment. Piringer et al. [133] supported situation awareness in the surveillance of road tunnels. Pu et al. [135] designed a vehicle fingerprint to monitor their movements on roads. Besides, visual analytics also empowered experts to detect abnormal traffic behaviors [3, 51, 92]. Timely detection can reduce the occurrence of accidents.

Autonomous driving. Autonomous driving is an emerging research area. Deep learning models are incorporated to understand complex road environments and produce driving actions. Thus, the performance of these models is critical. Recently, researchers have started to use visual analytics to evaluate traffic light detection [47], semantic segmentation [55], action prediction models [66], and even the entire system [57].

Prediction & Simulation. Andrienko et al. [12] proposed a visual analytics framework for accessing, forecasting, and developing what-if options. Zeng et al. [201] adopted visual analytics for diagnosing the impact of the flow aggregations on traffic prediction. Zeng et al.'s method essentially uses visualizations to improve the prediction models based on deep learning. Two approaches were used to predict traffic volumes, but they can be extended to other similar scenarios, such as predicting traffic speed and travel time.

2.2 Environment

Environment, on the one hand, is of great significance to environmental science and geographic science. On the other hand, it greatly affects people's lives. The widely deployed environment monitoring stations provide a sheer volume of environment data that can be used to gain insights into environmental issues, such as water quality, air quality, and meteorology changes. Environmental issues have also been widely studied in the visualization community.

Air quality. Air quality has increasingly been a critical problem and has a profound impact on the economy and health. Qu et al. [136] firstly proposed a visual analysis approach for understanding the air pollution problem. Li et al. [73] proposed a real-time visual query method for retrieving the spatiotemporal distribution of air pollution. To control air pollution, understanding its influence and propagation processes is significant. Various visual analytics techniques were developed to study the propagation processes based on co-occurrences [76, 184], simulation [36], event cascades [37], correlation [50], and causality [38]. Besides, Shen et al. [144] proposed a visual analytics approach combined with deep learning to predict air quality.

Meteorology. Meteorological issues also affect people's daily lives. Gautier et al. [45] overlaid temperature data onto a 3D city model to analyze the temperature. Qu et al. [136] designed a representation to reason the correlation between temperature and air quality. Wu et al. [179] analyzed the boundary changes of the ozone hole over Antarctica. Li et al. [77] visualized how climate changes in large-scale geographic space over a long period of time. Among many meteorological issues, weather forecasts are the most related to people. Weather prediction calibration [91] and prediction result comparison [138] have been studied with visual analytics.

Water quality, noise, and radiation. Water quality, noise, and radiation were less studied in the visualization community. For water quality, Maciejewski et al. [119] studied the effect of industrial wastewater on animals. Accorsi et al. [1] designed an interface for visually exploring the water quality of rivers. For noise pollution, Li et al.'s method [73] supported a fast visual query of the spatiotemporal distribution of flight noise. Malik et al. [120] explored the spatiotemporal correlation among noise complaints, traffic accidents, and drunkenness. Radiation pollution, especially nuclear pollution, is an extremely harmful environmental problem. Wei et al. [169] allowed the visual understanding of radioactive contamination based on the static and mobile sensors.

2.3 Business

At present, the application of urban visual analytics in business intelligence is mainly location selection.

Selecting facility locations. Selecting suitable facility locations is an important guarantee for business profitability. An informed selection process requires the integration of an intelligent recommendation model and a human-centered multi-faceted evaluation. Visual analytics is a good solution for that.

Different types of facilities have different selection criteria. Traffic flow needs to be considered in many location selection scenarios, such as commercial sites [146], billboards [95], and stores [170]. Some scenarios have specific criteria. Selecting houses focuses on the reachability over time [173]. Selecting warehouses should consider the delivery distance [173]. However, unlike the criteria mentioned above, there are some aspects that cannot be easily quantified, for example, the spatial context. To this end, Weng et al. [170] developed a context-integrated solution for location selection.

Selecting POIs for visit. In addition to selecting the location of facilities, visual analytics can also help people choose POIs to visit. Li et al. [74] embed the keywords extracted from social media into a metro map, which is called Metro-Wordle. Metro-Wordle allowed users to seek a restaurant for eating beefsteak based on the keywords over a city. Kamw et al.'s method [68] supported choosing a restaurant whose place is conveniently reachable to friends in different places.

2.4 Public security

Public security is to protect individuals, property, and objects from threats such as disasters or accidents. The threats can be categorized into human-made events and natural disasters.

Human-made events. Identifying unexpected human-made events with potential threats can be

Tab. 2 Domain problems studied by the existing urban visual analytics approaches (continued).

Domain	Problem		Papers
Environment	Air quality	Air pollution situation	[73, 136]
		Air pollution co-occurrence	[76, 184]
		Air pollution propagation simulation	[36]
		Air pollution correlation	[50]
		Air pollution event cascades	[37]
		Air pollution causality	[38]
		Air pollution prediction	[144]
	Meteorology	Temperature analysis	[45]
		Correlation of temperature and air pollution	[136]
		Ozone hole boundary changes	[179]
		Climate changes	[77]
		Weather prediction	[91, 138]
	Water quality	Water quality understanding	[1]
		Water pollution effect on animals	[119]
	Noise	Noise spatiotemporal distribution	[73]
		Correlation of noise and crimes	[120]
	Radiation	Nuclear contamination understanding	[169]
Business	Location selection	Commercial site selection	[146]
		Billboard selection	[95]
		Store selection	[170]
		House selection	[170, 173]
		Warehouse selection	[81]
		POI selection	[68, 74]
Public security	Human-made events	Suspect finding	[3]
		Abnormal event detection	[18, 21, 146]
		Marathon monitoring	[80]
		Crime pattern analysis	[113, 118, 120, 195]
		Resource allocation	[69, 121, 122]
		Fire station selection	[68]
		Crisis management	[116]
		Disease analysis	[2, 117–119]
		Surveillance video inspection	[123]
	Natural disaster	Flood impact analysis	[60]
		Evacuation monitoring and understanding	[82]
Architecture	Understanding	Human-scale scene sense-making	[127, 145]
		Non-visual city attribute prediction	[14]
		Shadow distribution analysis	[125]
		Location functionality analysis	[218]
		3D environment exploration	[137, 203]
	Planning	Impact analysis of new buildings	[43, 125]
		Urban space design precedent seeking	[126]
Economics	Real estate market understanding		[154]
	Spatiotemporal understanding of sales		[96]
	Trade network analysis		[163]
	Economic influence between countries		[76]
Public service	Public service events' hotspot, causality, and emergency		[204]
	Park management		[149]
	Locating new hospitals		[42]
	Lost and found		[28]
Public opinion	Spatial and temporal patterns		[4, 75, 112]

based on various data collected in cities, such as human trajectories [3, 146], social posts [21], and taxi trips [18]. Afterward, experts can further investigate these anomalies and take measures accordingly. Organized large-scale activities also demand emergency countermeasures. For example, marathons should be monitored in real time such that medical assistance is timely [80]. The aforementioned analyses should be near real-time, because experts need to quickly find abnormalities and respond to them.

Besides, visual analytics also supports in-depth post-analysis for public security, such as, crime analysis [113, 118, 120, 195], police resource allocation [69, 121, 122], fire station selection [68], crisis management [116], disease analysis [2, 117–119], and surveillance video inspection [123]. For example, given history crime data, crime hotspots with frequent occurrences of crimes can be visually analyzed [113, 118, 195]. Afterward, urban experts can take informed countermeasures, for example, reasonable police arrangements.

Traffic safety is also an important part of public security. Please refer to Sec. 2.1 for details.

Natural disasters. Timely responses to natural disasters can reduce many losses. Such decision making scenarios can be enhanced through visual analytics. Huang et al. [60] demonstrated the effectiveness of visual analytics in assessing the impact of floods. Li et al.'s approach [82] allowed the visual analysis of the emergency evacuation plan simulation.

2.5 Architecture

Architects or urban planners can use visual analytics to understand urban space and conduct space planning.

Understanding. The human-scale environment is the urban space that people are most directly exposed to, e.g., the things people see when walking on the street. Street views are the proper materials for understanding the human-scale environment. Thus, researchers proposed an efficient query method [127] and an exploration system [145]. These street views can also be exploited to predict non-visual city attributes [14]. In addition to street views, Miranda et al. [125] simulated and visualized the shadows in the urban physical environment, which helps understand the environmental quality of public spaces. Besides, Zhu et al. [218] explored and analyzed the functionality of locations in urban space.

On a larger scale, a zooming technique by Qu et al. [137] allowed users to explore a 3D urban environment in an occlusion-free way. Zeng and Ye [203] visually combined 3D physical entities and numeric urban design metrics to study urban vitality.

Planning. Urban space planning requires many considerations. For example, when developing new buildings, the impacts should be accessed. Ferreira et al. [43] and Miranda et al [125] combined information visualizations with a 3D urban environment to support such an analysis. A representative and mature city can be viewed as a precedent for other developing cities. Miranda et al. [126] visualized human behaviors in cities to derive urban space design precedents.

2.6 Economics

Data in economic domains has spatial and temporal characteristics, posing challenges of analysis. Sun et al. [154] analyzed the spatiotemporal development of the real estate market and the correlation among multiple economic attributes. Liu et al. [96] extracted and visualized the spatiotemporal patterns from sales volumes of different regions. In the era of globalization, the economies of different regions influence each other. Wang et al. [163] analyzed the trade network of countries. Li et al. [76] used co-occurrence patterns of the capita income data to infer the country-wise influences.

2.7 Public service

Public services in this survey refer to those serving society and urban residents. Zhang et al. [204] proposed a visual analytics system for investigating heat, water, gas supply, drainage, and road divisions issues. Hotspot and causality analyses and emergency discovery were supported in their system. Steptoe et al. [149] visually analyzed tourists' trajectories and visiting and communication behaviors in parks, and gained insights into improving park services. Chen et al. [28] connected multi-source heterogeneous urban spatiotemporal data through a novel spatiotemporal visual query, and applied it to finding lost objects. Feng et al.'s method [42] can facilitate locating a new hospital that serves more citizens and balances medical resources

2.8 Public opinion

Some methods utilized social media data to analyze the public opinion in the spatial or urban context [4, 75]. Combined with geographic information, users can have a deeper understanding of public opinion, for example, what opinions the people in a region have. Many studies have studied public opinion in the absence of spatial and urban context [19, 23, 53, 59, 70, 156, 183, 187]. We consider they are out of the scope of our survey. Please refer to the prior surveys for more information related to public opinion [24, 180].

3 Visualization

In this section, we summarize the visualizations in urban visual analytics studies into three categories, spatial, temporal, and other property visualizations.

Tab. 3 Visualizations in urban visual analytics.

Visualization		Main usage	Paper
Spatial	Map (dot)	Show distribution	[1, 7, 8, 11, 12, 21, 37, 39, 44, 50, 68, 71, 73, 74, 76, 77, 95, 100, 101, 105, 107, 116, 119, 121, 127, 135, 145, 152, 154, 170, 173, 177, 178, 191, 199, 200, 211, 212]
	Map (line)	Show distribution; visualize movement or influence	[3, 7, 8, 12, 25, 28, 29, 32, 49, 51, 61, 63, 66, 68, 71, 72, 74, 81, 92, 95, 97, 98, 101, 107, 108, 111, 113, 115, 123, 134, 146, 149, 167, 177, 200, 202, 214]
	Map (heat)	Show distribution	[2, 14, 18, 28, 38, 39, 42, 44, 45, 63, 66, 69, 73, 91, 95, 101, 112, 115–118, 120–122, 125–127, 160, 163, 173, 178, 179, 195, 198, 199, 201, 203, 211, 212, 218]
	Map (glyph)	Summarize or compare multi-dimensional spatial data	[6, 11, 18, 29, 36, 37, 61, 81, 91, 96, 109, 110, 115, 132, 135, 168, 172, 184, 191, 197, 199, 218]
	Map (area)	Show regions with the same attributes	[9, 12, 25, 44, 67, 68, 100, 107, 113, 115, 132, 138, 152, 161, 169, 172, 196]
	Map (graph)	Visualize movement, influence, or relations	[6–8, 12, 26, 36, 38, 82, 98, 105, 159, 168]
	Flow map	Visualize crowd movement	[69, 161]
	3D map	Provide physical context and a sense of presence	[7–9, 11, 12, 43, 45, 57, 60, 80, 82, 123, 125, 137, 191, 203]
Temporal	Timeline	Show temporal features	[3, 4, 6–9, 18, 25, 26, 29, 36, 37, 50, 63, 71–73, 77, 80, 95, 98, 109, 113, 115, 120–123, 126, 132, 133, 135, 146, 149, 159, 167–169, 173, 178, 195, 196, 199, 201, 212, 214, 218]
	Line/Area chart	Show statistics given two dimensions; show temporal features	[2, 7–9, 11, 12, 18, 26, 28, 32, 38, 44, 50, 57, 61, 66, 73, 80, 91, 96, 111, 112, 115, 117–121, 126, 132, 144, 161, 163, 169, 172, 173, 178, 179, 195, 199, 204, 211, 212]
	Streamgraph	Show temporal features of multi-objects	[51, 75, 144, 154, 179]
	Sankey	Show temporal features of multi-objects	[67, 111, 200, 212]
	Others	-	[38, 49, 76, 100, 101, 129, 177, 195, 196, 198]
Other property	Bar chart	Show statistics given two dimensions; show temporal features	[4, 7, 8, 21, 26, 28, 36, 39, 44, 45, 47, 49, 51, 55, 57, 63, 66, 72, 74, 76, 80, 91, 92, 95, 96, 98, 100, 101, 105, 108, 111, 115, 120, 121, 123, 129, 145, 146, 149, 152, 167, 170, 172, 173, 177, 184, 191, 195, 203, 212, 214]
	Tree/Graph	Visualize movement, influence, or relations	[1, 28, 36, 61, 82, 96, 136, 163, 167, 178, 196, 198]
	Scatterplot	Visualize projected multi-dimensional data; Show value distribution given two dimensions	[1, 3, 12, 18, 26, 32, 36, 38, 39, 44, 47, 51, 55, 60, 60, 67, 77, 80, 81, 95, 101, 117, 119, 126, 132, 136, 144, 145, 152, 160, 163, 167, 169, 179, 184, 201, 212, 214, 218]
	PCP	Show multi-dimensional data	[3, 32, 43, 47, 49, 51, 57, 101, 132, 136, 145, 152, 177, 178, 184, 203, 211]
	Radar	Show multi-dimensional data	[57, 100, 115, 214]
	Glyph	Summarize or compare multi-dimensional data	[36, 67, 76, 95]
	Matrix	Visualize movement or relations	[49, 55, 77, 152, 163, 170, 172, 178, 214, 218]
	Wordle	Show semantics	[21, 26, 32, 74, 75, 116]
	Video	Provide details and real-world context	[47, 55, 57, 66, 72, 72, 82, 123, 133]

This review further specifies 22 visualization types under these three categories shown in Table 3, which is different from the previous surveys [27, 210].

3.1 Spatial visualization

As the basis of urban analytics, spatial context is introduced in almost all visualization studies. Visual elements are then depicted in the spatial context, which constitutes spatial visualizations.

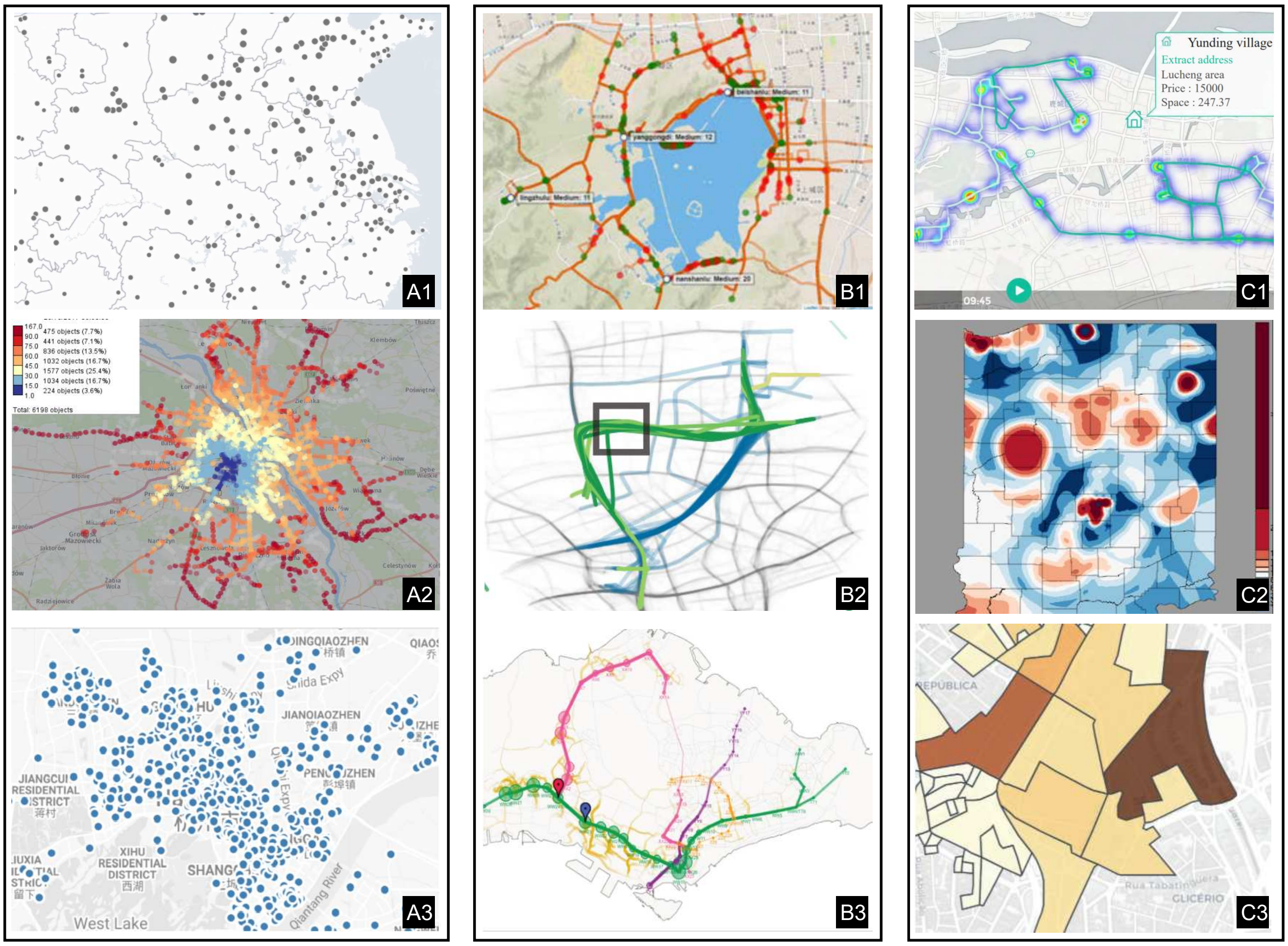

Fig. 1 Examples of dot-on-the-map visualization in [37](A1), [11] (A2), and [170](A3). Examples of line-on-the-map visualization in [3] (B1), [98](B2), and [200](B3). Examples of heatmap visualization In [28](C1), [118](C2), and [195](C3).

Such visualizations enable the urban analysis to be performed in a geographic context, such as obtaining data distributions and anomalies in geographic space. To obtain more precise insights, we further divide the spatial visualization into eight types.

Map (dot). Data points like geographic locations and spatial events can be directly plotted as dots on the map (Fig. 1A). Only simple visual channels, such as size and color, are adopted (e.g., in [11, 37] of Fig. 1A1 and A2). There is even no visual encoding in some cases (e.g., in [170] of Fig. 1A3), such that a large number of data points can be displayed in a scalable manner without overwhelming information.

Map (line). Lines or curves on the map can depict trajectories (like human mobilities) or facilities (like bus and subway routes). For trajectories, the line-based visualization is intuitive to show the mobility patterns. Fig. 1B1 shows the spatial distribution of the trajectories involving a target location [3]. Fig. 1B2 shows people's preference for routes [98]. For facilities, the line-based visualization retains the urban geographic context that is itself linear [172, 200], for example, the Singapore subway routes in Fig. 1B3.

Map (heat). There are continuous and discrete spatial heatmaps. The continuous heatmap is a smooth representation of aggregated geo-referred objects, usually generated by kernel density estimation (KDE). Both the lines and dots on the map can be aggregated to generate heatmaps. For example, Liu et al. [95] summarized the trajectories' pick-up and drop-off locations as the heatmap in Fig. 9B. Chen et al. [28] summarized the trajectories as heatmaps directly in Fig. 1C1. Besides, the spatial distribution of geo-referred events can be modeled using heatmaps. Maciejewski et al. [118] visualized the syndromic population over space based on syndrome events, shown in Fig. 1C2.

Discrete heatmap refers to choropleths [13, 62, 195,

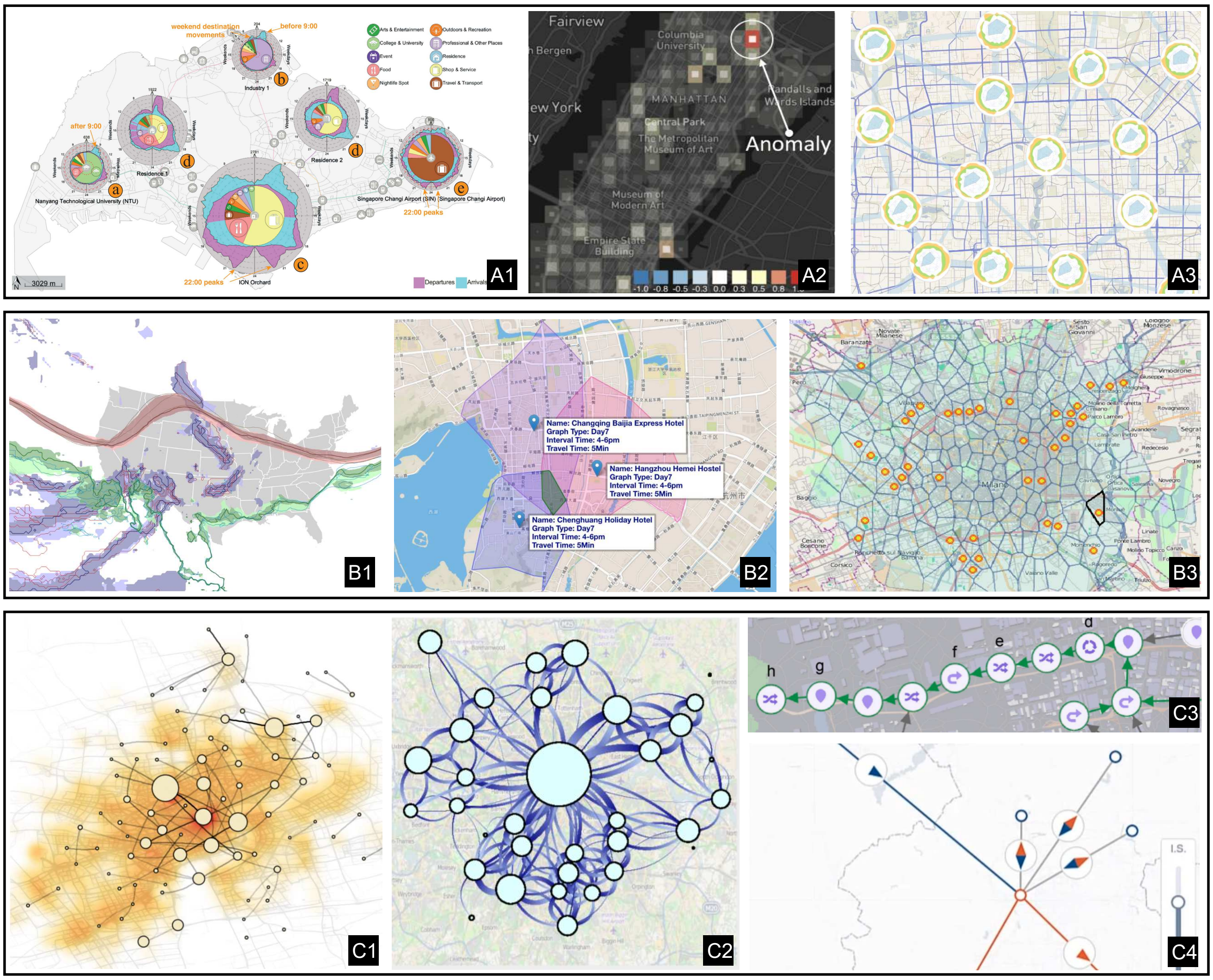

Fig. 2 Examples of glyph-on-the-map visualization in [199](A1), [18](A2), and [172](A3). Examples of area-on-the-map visualization in [138](B1), [68](B2), and [9](B3). Examples of graph-on-the-map visualization in [98](C1), [159](C2), [82](C3), and [38](C4).

205]. Semantic geographic space division avoids the use of smoothing algorithms such as KDE. Heat is used to show the value level of attributes in each region. For instance, Fig. 1C3 visualizes the number of crimes in different regions [195].

Map (glyph). Glyphs are effective visualizations for multi-dimensional data [190]. Glyphs on the map can summarize complex data together within a spatial context and support comparison and in-depth understanding, which can thereby serve as an overview that guides further exploration. Geographical information is indicated by glyphs' geographic positions. The visual channels of glyphs encode temporal information, attributes, or associated information. For example, in Zeng et al.'s glyphs [199], the inner pie chart visualized the portions of the POIs associated with the location, while the outer radial area chart encodes the weekly temporal distribution of mobility (Fig. 2A1). Cao et al. [18] layered rectangle glyphs on the map to reveal the attributes of detected anomalies (Fig. 2A2). Weng et al. [172] derived attributes related to the bus service for each region and encoded them with radar charts of glyphs (Fig. 2A3).

Map (area). A area (e.g., a polygon) on the map indicates that the region under this area have the same attributes. For example, the area constituted by isocontour means that an attribute of the region is higher than a certain value, which is widely seen in meteorology visualizations (e.g., Quinan et al's method [138] in Fig. 2B1). In the transportation domain, areas are used to visualize reachable regions from a location under constraints [179, 196]. Fig. 2B2 shows a concrete example [68]. Voronoi diagram [164] also belongs to areas on the map. In Andrienko et al.'s study [9] (Fig. 2B3), the urban space was divided based on the significant locations using a Voronoi diagram. Each region of a polygon is covered by a significant location.

Map (graph). A graph on the map comprises a set of nodes with geographic positions and edges between nodes. Undirected graphs on the map represent the mutual relationships between nodes. For example, in Liu et al.'s work [98] (Fig. 2C1), each edge represents the variety of routes between the two nodes of locations.

Directed graphs can represent the human mobilities (e.g., in [82, 159]) and spatial influence (e.g., in [36–38]). Landesberger et al. [159] applied spatial and temporal simplification to massive human mobilities and derived a concise mobility graph as Fig. 2C2. Li et al. [82] used the graph of Fig. 2C3 to visualize how people (or agents) move during an emergency evacuation. Some graph visualizations on the map encode the spatial influence [36–38]. Fig. 2C4 is an example from Compass [38], where edges encode the causal relations between the nodes of regions. Following the graphs, urban experts can obtain the urban deterioration patterns.

More information can refer to a recent review that comprehensively explores the visualization of geographic networks [140].

Flow map. The flow map treats human mobilities as a flow. Such a technique is often used to summarize massive crowd movement on a large spatial scale [69, 161]. For example, Kim et al. [69] visualized citywide movement patterns extracted from non-directional discrete events using flow maps.

3D map. 3D maps provide a realistic urban context in a more immersive and engaging way than the 2D geographic map [31]. It is commonly used in public security [60, 80, 82, 123] and architecture domains [43, 125, 137, 203] because it provides a sense of presence. For example, Li et al. [80] and Li et al. [82] adopted 3D views to track the people who are moving in the urban space with optimal scene navigation. Ferreira et al. [43] developed a 3D decision making framework, where users can well perceive the impact of new buildings on urban space (Fig. 3A).

Note that the third dimension may result in the occlusion of information visualizations. Qu et al. [137] (Fig. 3B) attempted to alleviate the issues by deforming urban space under constraints [153, 155].

3.2 Temporal visualization

Temporal visualizations display temporal features along a timeline. Such visualizations support time-oriented exploration and analyses, such as identifying urban data's temporal distribution and trends and drilling down for in-depth reasoning. We classify them into six types based on the graphical shape.

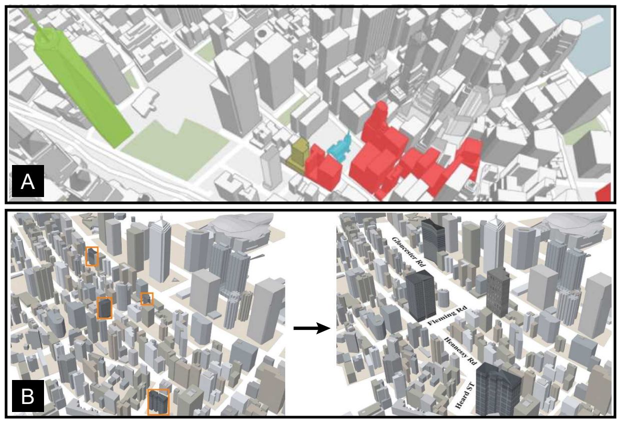

Fig. 3 Users can perceive the impact of new buildings on urban space in a 3D environment [43]. 3D urban space is deformed to highlight the target buildings [137].

Timeline. Timeline-based visualizations refer to those visualizations compactly encoded along a timeline. Numerical, boolean, and categorical data can be visually encoded on it.

The timeline can be in linear or circular shapes. Linear timelines (e.g., in [18, 37, 126, 132, 168, 173]) are easily accepted by general users. For example, Wang et al. [167] created a horizontal timeline (Fig. 4A1). Each row represents 24 hours of a day and each small grid encodes the (numeric) traffic speed with the size and color. Deng et al. [37] designed a folded timeline (Fig. 4A2) to show the (boolean) occurrences of event cascading processes. Each vertical bar indicates that a cascading process occurred during that time period. Andrienko et al. [6] adopted a calendar as the timeline (Fig. 4A3), where the (categorical) clustering results were encoded with the color.

Circular timelines are more artistic (e.g., in [26, 36, 146, 214, 218]). They are also easy to understand, because they borrow the clock metaphor. For example, Chen et al. [26] designed a wheel to show the periodic temporal distribution of different mobilities (Fig. 4A4). Similarly, in Deng et al.'s glyph [36], the circular timeline showed the temporal occurrences of an air pollution propagation pattern.

Line/Area chart. The line (or area) chart uses lines or areas above a timeline to show the temporal evolution. As a basic chart, it is widely and undoubtedly accepted by general or expert users and thereby can be safely used in many scenarios (e.g.,

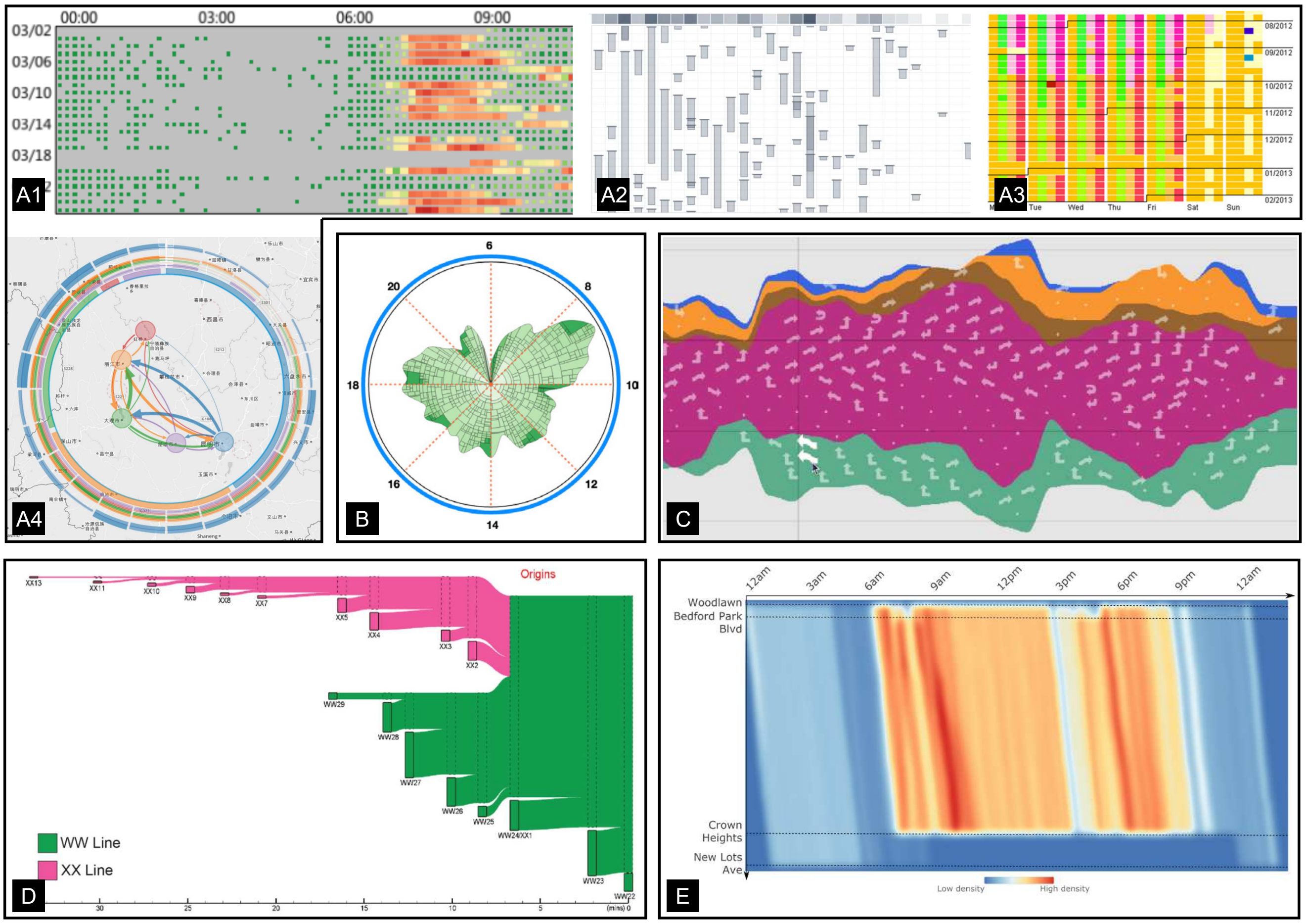

Fig. 4 Examples of timeline visualization in [167](A1), [37](A2), [6](A3), and [26](A4). Circular area chart in [178] (B). Streamgraph in [51](C), Sankey diagram in [200](D). Marey graph equipped with KDE in [129](E).

in [2, 38, 163, 195]).

It is worth mentioning that a line (or area) chart combined with the circular timeline can also produce more expressive and artistic visualizations [178, 199, 212]. For example, Wu et al. [178] designed a glyph with a circular area chart to show the temporal distribution of mobilities (Fig. 4B). Zeng et al. [199] wrapped their area charts of in and out volumes around pie charts to design an effective glyph (Fig. 2A1).

Line charts can also visualize non-temporal data, if the x-axis does not represent the time (e.g., in [7, 12]).

Streamgraph. Stacked area charts constitute a streamgraph. Each stacked area is called flow. Besides the quantity of each variable, a streamgraph also displays each variable's percentage to the sum of all variables. Many researchers use streamgraphs because of this advantage [51, 75, 154, 179]. Fig. 4C shows the streamgraph in Guo et al.'s work [51]. It visualized the evolution of different traffic flows.

Sankey. Compared with streamgraphs, sankey diagram can show not only the evolution of object groups but also the transitions between groups over time [111, 200, 212]. The transitions are encoded by the split and merging of flow. For instance, the sankey diagram in Zeng et al.'s work [200] visualizes the paths of passengers coming to a station and the corresponding volume along the time (Fig. 4D).

Others. There are interesting urban temporal visualizations [38, 49, 76, 100, 101, 129, 177, 195, 196, 198] that cannot be classified into the above categories. For example, in Fig. 6A, the journeys started from a station (the leftmost red node) are organized according to the travel time in a parallel isotime fashion. That is to say, the travel times from the leftmost station to any station with the same vertical position are the same. A Marey graph smoothed by kernel density estimation (Fig. 4E) was designed by Palomo et al. [129]. It visualizes the movement or schedules of buses departing at different times. These designs were generally proposed for specific data and problems.

3.3 Visualization of other properties

In addition to spatial and temporal information, urban data will also have high-dimensional, relational, and semantic information, etc. These data also need to be presented and comprehended visually so that analysts can comprehensively analyze urban problems from more aspects. We identify nine popular visualization components for non-temporal and non-spatial data.

Bar chart. The bar chart is also a basic chart. It can not only show the temporal information [4, 184, 195] but also statistical distributions because its coordinate axis can be discrete [63, 95, 96]. For temporal data, the usage of bar charts is similar to the line (or area) chart. We would not repeat here. For statistics, bar charts can be used to display and compare the attribute values of different items (e.g., performances of different prediction models).

Bars for multiple attributes can be stacked as a stacked bar chart. An interesting form of the stacked bar chart is ValueChart [20] or LineUp [48] (Fig. 5 and Fig. 9F) in decision making and ranking scenarios [95, 145, 152, 157, 170, 172, 173, 178], supporting intuitive comparison of data points with multiple attributes. By stacking the bars of attributes, the stacked bars' heights encode the sum of the corresponding attributes. In LineUp, users are allowed to interactively choose which attributes and bars to be stacked, thereby enabling informed decision making and ranking.

Tree/Graph. The nodes in a tree or graph may have geographic positions or not.

If nodes have geographic positions, tree and graph visualizations usually serve the same functionalities as they are on the map. The positions of nodes are free by separating from the map [1, 163] and so that the visualization layouts can be improved or optimized to pursue aesthetics, legibility, and faithfulness. In the transportation domain, researchers tended to use tree visualizations to visualize the mobilities involving a

Fig. 5 LineUp adopted in [173]. The bars of the first two attributes are stacked by users.

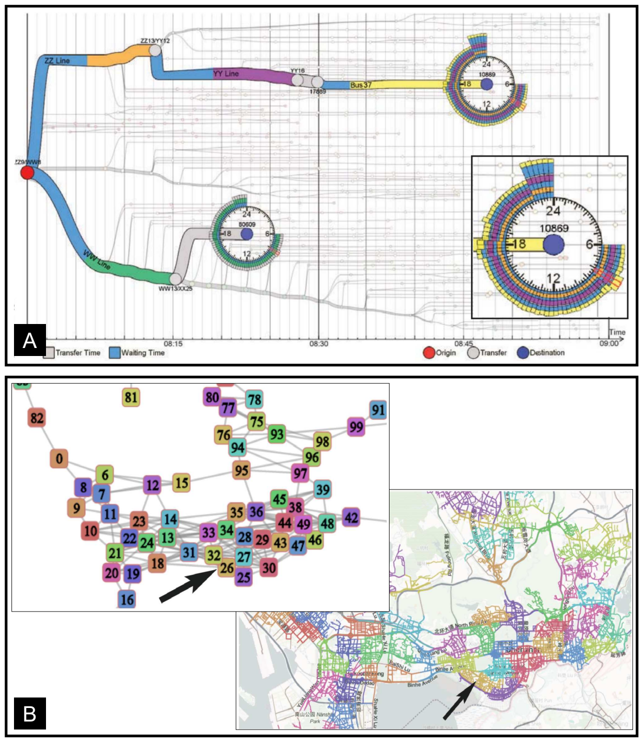

Fig. 6 In a public transportation system, journeys starting from a certain station are organized into a tree [196]. The timeline is from left to right. A graph visualization shows the traffic linkages of different regions [61]

target location [178, 196, 198]. The target location was viewed as the root nodes of the tree. For example, in Fig. 6A, Zeng et al. [196] employed a tree visualization to organize the journeys of a transportation system that started from the leftmost station.

These visualizations are usually coordinated with other multiple views and thus can be related back to the spatial context. Some approaches copied the graph visualizations directly from the map as a separate view, in which free layouts provide clear appearances for investigation [36, 61, 82, 167]. Fig. 6B shows the graph visualization by Huang et al. [61]. Each edge indicates the linkage between traffic regions.

If nodes are without geographic positions, they usually represent the relationships between process status or data attributes. For example, tree visualizations visualized and guided the processes of steerable data partitioning in [50, 96]. Fig. 11 is the interface of TPFlow [96]. In Qu et al.'s method [136], graph representations were adopted to present the correlation of different air pollutants.

Some studies transformed graph data into other forms, such as tabular [37, 38]. We would not discuss them here.

Scatterplot. The scatterplot is another basic chart. It can unveil interesting patters, such as, clusters, outliers, trends, and correlation [165].

In a traditional scatterplot, the x-axis and y-axis are assigned with semantic dimensions [128], for example, driving velocities and distances of vehicles [160] (Fig. 7A1). Similarities and correlations of the dimensions can be identified. In addition, some scatterplots are generated by reducing the dimensionality of high-dimensional or embedded data, e.g., in [214](Fig. 7A2). Both the x-axis and y-axis have no specific semantics. So, they are only used to show clusters and outliers. Nonetheless, due to its capability of summarizing massive high-dimensional data, such a visualization is increasingly popular in urban analytics encountering a sheer volume of data.

Some approaches replaced the projected dots with the glyphs encoding the original dimensions, e.g., in [95](Fig. 9D). In this way, scatterplots become more informative. Note that using glyphs in a scatterplot may cause occlusion and clutter issues. Therefore, it is suitable when the amount of data is small.

Parallel Coordinate Plot (PCP). High-dimensional attributes of urban data are commonly seen. They sometimes are neither spatial nor temporal, but are important in urban analysis.

PCP is the most widely used high-dimensional visualization besides the projection-based scatterplot in urban visual analytics. For example, every trajectory was described from non-geographical dimensions of PCPs [3, 32, 49, 51, 152, 177, 211]. Fig. 7B shows the example from [3]. Attributes that describe physical urban environments quantitatively were encoded with PCPs in [43, 203]. PCPs, together with the 3D Map, supported both quantitative and qualitative understanding of the urban environment.

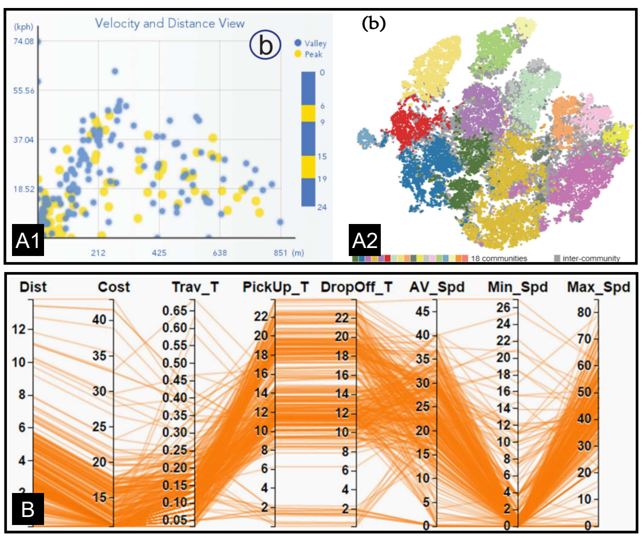

Fig. 7 Examples of scatterplot in [160](A1) and [214](A2). An example of PCP in [3].

PCPs can handle more dimensions with less space than scatterplot matrices. Besides, they are easy to learn and understand compared to glyphs.

Radar chart, glyph, and scatterplot matrix. These visualizations are also suitable for high-dimensional data. Radar charts can be seen as a circular PCP with radar metaphor but handle fewer dimensions than PCP. They usually do not support user interactions. But due the universality and artistic circular shape, radar charts can be seen in many studies [57, 100, 115, 214]. Glyph-based designs are usually placed on the map, as we have discussed before. Some glyphs are put outside the map, such as scatterplots after projected [95] or side by side [36, 76]. These layouts can support better comparisons. By contrast, the scatterplot matrix was less used in urban visual analytics [145], which may be due to the space inefficiency.

Matrix. Matrix visualizations are a well-arranged representation usually equipped with color encodings. In urban visual analytics, matrix visualizations can visualize location pair's relationships, where each row or column represents a location, and each cell encodes the relationship between the row and column. For example, each cell encodes the traffic volume from the row location to the column location [152, 172, 188] or the interconnection between the row and column locations [178, 218]. Besides, there are also some classic uses of matrix visualizations, such as showing classification performances (confusion matrix) [55] and statistics [49, 170].

Wordle. Wordle or word cloud [148] is usually used in text-based urban visual analytics [21, 26, 32, 74, 116]. It assists in inspecting massive amounts of original text information, usually combined with keyword extraction.

Video. Video is not essentially a visualization method. Nonetheless, in urban analytics, video is an important way to provide raw information. It is also the most familiar component for many urban experts. Some researchers will add a video component to their visual analytics approaches, allowing experts to verify the conclusions obtained [47, 55, 66, 72, 123, 133].

4 Computational analysis method

This section reviews the computational data analysis methods used in urban visual analytics. We intend

Tab. 4 Computational methods in urban visual analytics.

Computational methods		**Example Papers**
Learning-based	Clustering	[1, 6–8, 26, 49, 61, 71, 77, 82, 98, 100, 101, 110, 111, 117, 146, 149, 152, 159, 161, 178, 179, 191, 204, 212]
	Classification	[14, 47, 66, 92, 112, 132, 145, 191]
	Representation	[36, 47, 55, 127, 214, 218]
	Dimensionality reduction	[36, 76, 95, 184, 214]
	Regression	[12, 14, 38]
	Forecasting	[72, 144, 201]
Statistical	Kernel density estimation	[39, 42, 69, 73, 112, 113, 117, 118, 120–122, 126, 129, 146, 179, 214]
	Matrix/Tensor/Time series decomposition	[18, 21, 67, 96, 117, 119, 122, 195]
	Deviation-based anomaly detection	[18, 21, 80, 111, 118, 119, 146, 149, 163, 168]
	Keyword/Topic extraction	[4, 21, 26, 32, 63, 74]
	Frequent pattern mining	[1, 36, 76, 184, 198]
	Peak/Periodicity/Correlation of time series	[9, 38, 50, 120, 163]
	Metric/Indicator/Factors calculation	[32, 57, 61, 81, 98, 108, 112, 132, 136, 177, 211]
Rule-based	Association	[29, 32, 42, 63, 71, 76, 178, 184, 199]
	Summarization	[6, 12, 105, 117, 118, 133, 159, 169, 198, 201, 211]
	Map matching	[3, 32, 72, 98, 103, 107–111, 135, 146, 160, 167]
	Traffic modelling	[12, 72, 167, 168]
	Data cleaning	[25, 115]
	Object tracking	[123]
	Heuristic search	[37, 91, 95, 105, 172]
Simulation-based		[2, 36, 43, 69, 125, 134]
Mathematical programming		[80, 82]
Index		[3, 29, 63, 75, 95, 103, 108–111, 160, 161, 173, 184, 198–200]
Query		[3, 28, 44, 63, 73, 127, 145, 160, 161, 173, 200]

to reveal the commonly used methods and their purposes. Practitioners can build on these well-established methods to develop intelligent analysis solutions. Seven categories are identified in this survey. Table 4 summarizes the computational methods.

4.1 Learning-based

Learning-based methods' parameters are learned from the inherent distribution of given data. These methods do not require much prior knowledge to learn the intrinsic patterns of the data that provide useful insights or aid in predictions.

Clustering. Clustering is a basic data analysis operation, which divides data into multiple clusters by learning the similarities of data. The data items in the same cluster are similar in some aspects. Such an operation can reduce the amount of data to be analyzed and provide exploration guidance in visual analysis. As long as a similarity measure is well defined, clustering can flexibly apply to various kinds of data, such as events [117], regions [179], and trajectories [49]. Popular clustering techniques are DBSCAN (e.g., in [61, 146]) and k-means (e.g., in [117, 179]).

Classification. Classification refers to labeling data with a given set of categorical tags. In the early years, researchers tended to use traditional classification techniques. such as, Support Vector Machine for predicting non-visual city attributes [14], Conditional Random Field for classifying abnormal trajectories [92], and random-forest for classifying locations into work, home, and others [191]. As powerful neural networks are proposed and popularized, neural-network-based classifiers are increasingly applied to urban analytics, such as sentiment classification of geo-tagged social media [112], human-scale visual feature identification [145], traffic light detection [47], and autonomous driving decision [66].

Classification techniques rely on the acquisition and quality of labels. Active learning [92, 191] or semi-supervised learning [132] is an efficient mechanism for obtaining labels. Briefly, model trainers first label a portion of data. The model trained on these data is used to label another portion of data. Repeat this process until the number of training samples meets the requirements.

Representation. Representation learning embeds data into high-dimensional space by vectorization [217]. In this space, adversarial examples can also be generated for adversarial learning [47, 55]. The vectors also capture the inherent relationships between data. Reducing the dimensionality of the data vectors to a low-dimensional space can generate an overview for further exploration.

Representation learning can be based on autoencoder [47, 55] or word2vec [36, 214, 218]. Those word2vec-based make full use of the characteristics of spatiotemporal data. For example, in [218], each location is cast as a word, and each trajectory is cast as a sentence comprising the words it passes. Thus, the latent semantics of locations are learned like learning the semantics of words.

The above methods often take the input itself or the context of the input as the output, so avoid the labeling process to obtain labeled samples. Some approaches leveraged the models pre-trained with labels in the open-source community. For example, Miranda et al. [127] input street images into pre-trained image recognition models and obtained the latent vector representations for these images from the hidden layers.

Dimensionality reduction. Dimensionality reduction transforms the ubiquitous multi-dimensional data into low-dimensional data. It is usually combined with 2D scatterplots, which is useful and popular in urban visual analytics [95, 170]. Some high-dimensional embeddings also require dimensionality reduction to be visualized, as we mentioned before. Popular techniques include PCA [47], MDS [95, 184], and t-SNE [36, 76, 214]. Currently, t-SNE is the most effective dimensionality reduction method, particularly when dealing with high-dimensional and large-scale data.

Regression. Regression refers to estimating the relationships between variables [12, 14, 38]. For example, in [38], the Granger causality test based on vector autoregression was adopted to detect the causal relations between time series. In [12], a polynomial regression model was applied to capture the dependencies between the traffic intensity and velocity.

Forecasting. Forecasting refers to predicting the future status. Understanding future development trends is an essential prerequisite for wise urban planning. Time series forecasting based on deep learning has achieved great success. Hence, the corresponding visual analytics has also emerged [72, 144, 201]. Forecasting belongs to the supervised learning. Fortunately, training samples with labels (i.e., future situations) for forecasting can be constructed through sliding windows on the time dimension, avoiding the labeling process.

4.2 Statistical

Statistical methods mainly build on statistics to process data. Because of mathematical theories, such methods are often explainable and can guarantee the reliability of analyses.

Kernel density estimation (KDE). KDE estimates the probability density function of discrete data to characterize them in a continuous way.

KDE can be directly applied to spatial events [120], locations [121], geo-tagged social media posts [112], trajectories [95, 173], and bus schedules [129], generating clear overviews without clutters. Furthermore, KDE has been extended for many scenarios. Feng et al. [42] proposed a topology density map that considers topological conditions of the road network. Li et al. [73] proposed a peak-based KDE that avoids manually setting the band width.

KDEs are usually equipped with heatmap visualizations (e.g., in Fig. 1C1 and C2). There are also studies that use KDEs not for visualization. Data processed by KDE can facilitate further computation and analysis [118, 122], such as, sampling [214], reducing errors [146], prediction [117], flow map extraction [69] and topology feature extraction [39, 113, 126]. For example, Lukasczy et al. [113] used KDE to generate the scalar function of discrete spatial event data. The scalar function enabled the following topology analysis.

Matrix/Tensor/Time series decomposition. We classify matrix factorization, tensor decomposition, and time series decomposition into this category as they decompose data into multiple components.

Zanabria et al. [195] modeled crime data as a matrix in which each row represents a region and each column represents a time slice. Then, matrix factorization was applied to this matrix and extracted the spatial and temporal patterns represented by decomposed matrices. Tensor decomposition techniques in [18, 96] mathematically extended matrix factorization. In these two studies, spatiotemporal data were modeled as tensors. Latent spatial and temporal patterns were then extracted by decomposing the tensors.

In contrast, the time series decomposition, particularly the Seasonal-Trend Decomposition (STL), works on time series data. STL decomposes a time series into multiple ones with temporal patterns, such as, yearly seasonality, day-of-the-week effect, and global trend [21, 117, 119, 122]. Anomaly detection and prediction can be performed given the decomposed time series.

Deviation-based anomaly detection. For the temporal perspective, determining whether a data point is abnormal is generally based on how much it statistically deviates from other historical ones. To compare with historical data, many methods can be used, such as Local Outlier Factor (LOF) [18], cumulative summation (CUSUM) [118, 119],minDistort algorithm [168], and Extreme Value Theory [146]. From the sampling perspective, anomaly detection can also be accomplished by measuring the deviation of a data point in the samples [21, 80, 149].

Keyword/Topic extraction. This type of method extracts frequently occurring keywords or implicit topics from document corpora. The corpora can be directly from text [4, 21, 26, 74] or textualized trajectories after rule-based association [32, 63] (see Section 4.3). For example, Chen et al. [26] extracted frequent keywords for locations from the microblogs there, enabling semantic analysis of movement data. Another different example is Huang et al.'s approach [63]. Each trajectory was assigned with the names of the streets it passed and thereby was regarded as a sentence. This process generated a trajectory corpus. Afterward, Latent Dirichlet Allocation (LDA), a topic modeling technique, extracted the latent topic and related keywords for querying and reasoning mobility patterns.

Frequent pattern mining. Frequent pattern mining methods can extract significant (frequent) patterns with sufficient occurrences from various types of urban data, such as transactional [76, 184], and sequential [1, 198], and graph [36] data. Frequent patterns summarize massive data, which prevents users from being overwhelmed. For example, Deng et al. [36] adopted frequent subgraph mining to extract the propagation patterns from numerous propagation processes of air pollutants. These patterns were then organized and visualized for analyzing the air quality deterioration on a large spatiotemporal scale.

Peak/Periodicity/Correlation of time series. Time series data is one of the important data types in urban analysis. Correlation analysis and periodicity identification are common analysis tasks. Correlation analysis is usually based on Pearson's r [38, 120, 163]. Information theory can also be applied [50]. Periodicity can be identified using the Fourier Transform [9].

Metric/Indicator/Factors calculation. Some indicators need to be extracted through statistical methods. Entropy based on the information theory is popular because it can leverage the inherent data distribution to estimate the amount of information [32, 98, 112, 132, 177, 211]. The richer the amount of information, the more worthy of analysis. For example, Zheng et al. [211] used entropy to quantify the interestingness region by incorporating spatiotemporal and mobility-related attributes. Other calculations include but are not limited to estimating the road importance by the PageRank algorithm in a graph-represented road network [61] and deriving correlation metrics between dimensions in air pollution data [136].

4.3 Rule-based

Rule-based methods incorporate effective problem- or domain-related rules to guide data analysis. Similar to statistical methods, rule-based methods are explainable. Besides, they are more appreciated by domain experts due to their ability to incorporate domain knowledge.

Association. Many rule-based approaches can associate different urban data based on their co-occurring observations, which enriches data and discloses the latent relationships of various urban data.

Associating POIs with mobilities is the most common association operation in urban visual analytics. In [32, 63, 71, 199], geographic positions of movement records were assigned with a set of POIs spatially near the positions. Besides, POIs can be associated with the road network, enriching semantics to urban facilities for accessibility analysis [42].

The spatiotemporal co-occurrence rule can also apply to associating people with people [29, 178], value ranges with value ranges [184], and events with events [76].

Summarization. Summarization procedure is often demanded in visually analyzing massive urban data.

Adrienko et al. [10] demonstrated that movement summarization is an important step towards scalable and effective urban analyses. In many practical approaches, mobilities were aggregated by their origins and destinations and summarized as the flow between them [6, 12, 105, 159, 211]. Visiting sequences were aggregated and summarized as a visual summary according to their shared parts [198].

The aggregation of spatial geo-tagged data [117, 118, 169] and spatial regions [159, 201] in varied granularities are also seen in many approaches. These summarization procedures can reduce visual clutters on the map and sometimes protect data privacy.

Map matching. Map matching is a fundamental process for movement analysis. Position shifts exist in the movement data collected by GPS devices. Map matching maps these GPS records onto real road networks because vehicles and people can only move on the roads, generating clean and usable movement data. Given that, many applications using GPS data adopted map matching techniques (e.g., [3, 103, 167]).

Heuristic search. Optimal solutions to some urban analysis problems cannot be obtained directly and quickly. Thus, researchers incorporated some rules into heuristic search methods to solve the problems.

Greedy search is one of the most popular heuristic search methods. In the billboard selection [95], the billboard locations should cover as many trajectories as possible to have high exposure. Such a problem is essentially the K-cover problem [85, 86]. In the cascading pattern inference [37], extracting cascading patterns that best describe the observed event data is also the K-cover problem [87]. In these scenarios, greedy search methods were because of the NP-Hardness of the K-cover problem.

Heuristic search can also be used in improving a bus network. In Lorenzo et al.'s AllAboard [105], the public transportation system can be improved by minimizing users' travel times under constraints. Lorenzo et al. considered it a separable programming problem and adopted a heuristic procedure to approximately solve it. In Weng et al.'s BNVA [172], the bus network improvement was seen as a multi-objective optimization problem considering, such as cost, passenger volume, and directness. Weng et al. used a heuristic search called Monte Carlo tree search in this context. Besides, heuristic strategies can be used for OD-flow sampling [214] and majority voting of weather forecast calibration [91].

Traffic modeling, data cleaning, tracking. Domain-specific rules also guide traffic modeling, data cleaning, and tracking.

Vehicles must run on the road based on traffic rules under physical constraints. Therefore, traffic congestion propagation can be estimated based on the topology of the road network [72, 167]. Moreover, the road traffic was simulated based on the similar heuristic rules [12, 168]. Some data cleaning operations are also based on rules. Ma et al. [115] defined the Ping-Pong effects in the mobility data as frequently switching between different locations and proposed detection and elimination algorithms. Chen et al. [25] concluded five types of uncertainty in human behavior data and developed a semi-automatic processing framework. Finally, Meghdadi et al. [123] tracked people from consecutive video frames based on the continuity of movement.

4.4 Simulation-based

Simulation-based methods refer to those simulating real-world conditions based on physical phenomena. They effectively recover the real-world urban situation and analyze the development and evolution.

Ray tracing was utilized to compute the impacts of new buildings in terms of shadow [125] and visibility [43]. Gravity model was utilized by Kim et al. [69] to construct flow field from discrete event data. The particle advection technique helps generate and visualize the 2D vector fields of traffic flow [134]. The aforementioned methods have potential generalizability. Domain-specific models based on epidemiology and fluid mechanics were also employed for epidemic response evaluation [2] and air pollution propagation modelling [36], respectively.

4.5 Mathematical programming

Some analysis problems can be characterized as operation research problems and be solved via mathematical programming. Although mathematical programming methods can obtain globally optimal solutions, they are used less often, perhaps due to their inefficient performance on large amounts of data.

MaraVis [80] computed the optimal camera path for monitoring marathon games by solving a traveling salesman problem. SEEVis [82] also computed the optimal camera path but for exploring human movements during an emergency evacuation.

4.6 Index and query

Index techniques guarantee efficient data retrieval, computation, and visualization.

Many urban visual analytics approaches incorporated well-established spatiotemporal index methods, such as B^+ tree [184], quad tree [29, 108, 111], octree [145], k-d tree [44], locality sensitive hashing [127], and space-time cube [28]. In addition to these general indexes, there are indexes tailored for trajectory computation and visualizations. Location-trajectory indexes enable efficient trajectory retrieval by the passing locations [95, 103, 160, 161, 173, 198, 200]. Furthermore, trajectories can be indexed by text, after the locations they pass by are textualized by associating POIs [3, 63]. With these indexes, the system can support flexible semantic-based queries.

With the increasing amount of urban data, many advanced index techniques are proposed in the database community [83, 84]. Being aware of the uniqueness of visualization and visual analytics tasks, visualization researchers proposed big data indexing and management approaches for visualization purpose [34, 40, 73, 75, 93, 94, 124].

5 System

Due to the spatiotemporal, heterogeneous, uncertain, and dynamic characteristics of urban data, machines and human intelligence are often required to be integrated into the analysis process. This section discloses how computation and visualizations constitute a human-in-the-loop urban analytics system. Unlike the previous survey [210], our survey focuses on the role of computation and visualizations and how they interact each other in a system. Four categories are identified based on the combination ways of models and visualizations: *visual analytics without models*, *post-model visual analytics*, *model-integrated visual analytics*, and *visual analytics-assisted models*.

Note that the models mentioned in this section may comprise multiple computational methods in the last section. For example, although the location section model by Liu et al. [95] comprises map-matching, index, and optimization modules, it is viewed as one model from the perspective of an entire system.

5.1 Visual analytics without models

The systems in this category exclude models. They are suitable for those scenarios where the raw data does not need to undergo complex computational transformations. Urban analyses mainly rely on well-designed data visualizations.

Individual visualizations. For an emerging field and problem, data visualizations contribute enough to the visualization community and domains [4, 136, 139, 158, 197]. The most obvious evidence is in the study of human mobility visualization. In the early years when mobility data started to be collected, the presentation of mobility data provided sufficient insights [139, 158, 197]. Similar evidences can be observed in air pollution [136] and geo-tagged social media analyses [4].

Coordinated visualizations. As the data becomes increasingly abundant and the tasks become increasingly challenging, multiple coordinated visualizations become demanded [22, 29, 51, 77, 98, 109, 110, 115, 121, 133, 135, 154, 163, 169, 170, 179]. To handle large-scale data, these systems generally follow the workflow of "Overview first, zoom and filter, then details-on-demand [147]." For example, VisMate [77] firstly provided a spatiotemporal summary of climate data collected in all meteorological stations. Then,

Fig. 8 The views in AirVis [36] from left to right hierarchically organize numerous propagation patterns of air pollution.

users were allowed to zoom into a station cluster of interests and finally individual stations. Similarly, Liu et al.'s method [98] allowed users to analyze the numerous trajectories from the region, trip, and road perspectives.

Completing complex tasks also raises the same requirement of coordinated visualizations, although the data is not too large. For instance, Alvis [133] coordinated multiple visualizations to support tasks of tunnel surveillance, such as navigation, event prioritization, video retrieval, and situation predictions.

5.2 Post-model visual analytics

In this category of system, the role of computational models is to discover knowledge behind urban data in advance. Once the model has run, it rarely needs to be adjusted and run again. Visualizations are designed to organize and present the discovered knowledge. We further classify these systems based on the specific purposes of models into three types, namely, pattern extraction, item detection, and data enrichment.

Pattern extraction. Models first extract many patterns from original data. Visual analytics systems are developed for understanding these extracted patterns, such as co-occurrence [178, 184, 212], mobility [49, 69, 159, 198, 214], air pollution propagation [36], and social media distribution patterns [21, 74, 75, 112]. The huge amount of patterns may hinder identifying and reasoning valuable ones. A well-designed pattern organization is desirable to this end. For example, Wu et al. [178] extracted co-occurrence patterns in human mobility from trajectory data and designed visualizations for pattern exploration from multiple levels and perspectives. Deng et al. [36] extracted massive propagation patterns of air pollutants from the ubiquitous propagation processes. They used hierarchical organization for effective top-

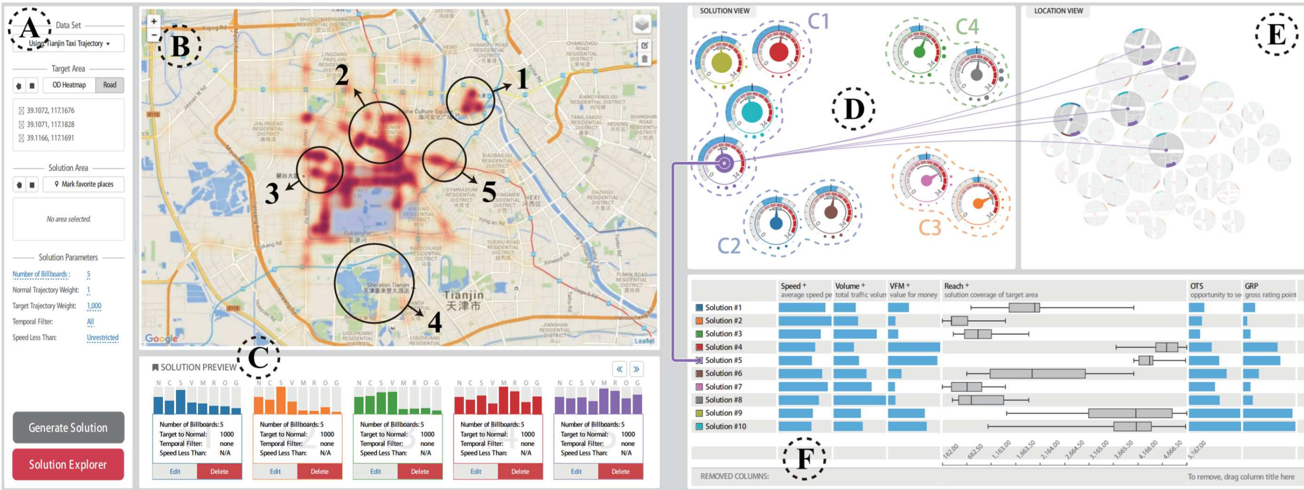

Fig. 9 Interface of SmartAdP [95], a visual analytics system for selecting billboard locations based on taxi trajectories.

down exploration (Fig. 8).

Item detection. Models can detect important ones from a collection of data items. In urban scenarios, it is common to detect important locations from the vast geographic space [7, 8], such as those with frequent movement [211], traffic jams [132], prominent pulses [126], and crime hotspots [113, 118]. Starting from these locations, users utilize visualizations to understand these locations and what happened there. For instance, in Pi et al.'s traffic congestion analyses [132], the roads detected with potential traffic jams and the inferred causes provided the analysis entrance.

Data enrichment. In many systems, models enrich original data by deriving the additional information [71, 123, 145, 177, 196, 199, 218], such as transportation system metrics [42, 61, 105, 203], clusters [32, 67, 149, 152, 204], and temporal prediction [117].

Take the transportation network diagnosis as an example. Accessibility [42] and centrality [61] of road networks were calculated and visually encoded together with the networks. Human mobilities were estimated and encoded into the public transportation networks to diagnose the system efficiency [196](Fig. 6A). The additional information provided important implications for understanding transportation systems, which cannot be supported only by visualizations.

5.3 Model-integrated visual analytics

The utility and intelligence of the system can be significantly improved by tightly integrating computational models into visual analytics. Visualizations and computational models tend to interact frequently. Here are some typical scenarios.

Problem-specific analysis. In many specific problems, users need to determine a subset of data of interest through spatial [1, 11, 37, 81, 95, 108, 172, 195], temporal [100, 107], spatiotemporal [38, 111, 122, 173], or other property visualizations [76]. The subset is fed into the models, and the outputs are visually displayed.

Unlike post-model visual analytics systems, these systems require visualizations to determine the inputs of the model. For instances, the billboards should be placed in a solution area and viewed by target audiences from a target area. Thus, SmartAdP [95] firstly required users to specify target and solution areas on a map based on a visual trajectory overview (Fig. 9B). Afterward, candidate billboard locations are extracted, given the areas and trajectories. The next step is to visually compare and evaluate these candidate locations (Fig. 9D, E, and F). In the study of BNVA [172], users were firstly required to conduct network-level analysis on the map and identify the bus routes with low efficiency. Then, an optimization method manipulated the target bus route and users can visually evaluate the optimized routes.

Multi-steps analysis. If an analysis workflow contains multiple steps, each step may involve models and visualizations. In the study of trajectory data cleaning [25], users iteratively interact with different modules of the model through visualizations to address the uncertainty of different dimensions.

Iterative analysis. Another case is that the analysis pipeline contains a loop where models can iteratively learn users' feedback and update accordingly [18, 92, 191]. Take anomaly detection as

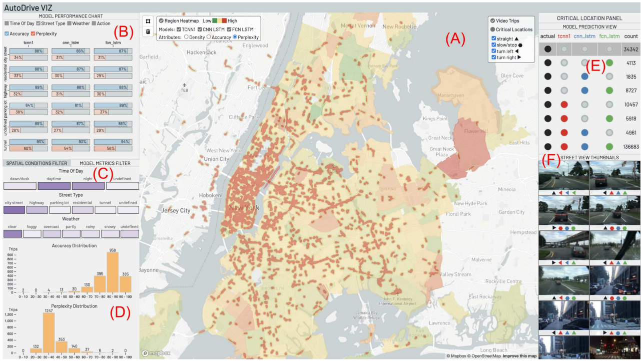

Fig. 10 Interface of the geo-context aware diagnosis systems for autonomous driving models [66]

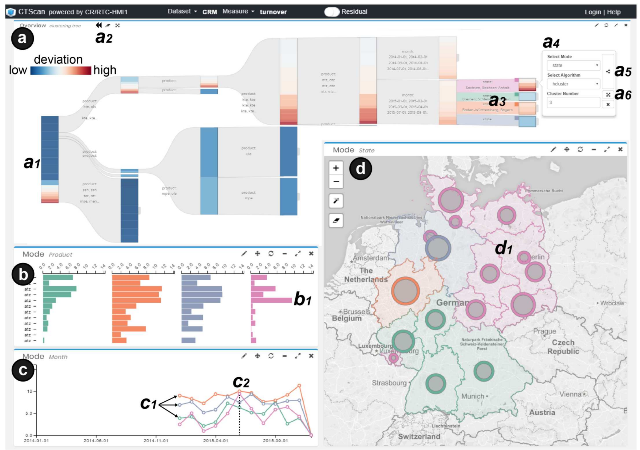

Fig. 11 Interface of TPFlow, a visual analytics system that supports steering tensor decomposition and analyzing spatiotemporal patterns [96].

an example. Liao et al.'s method [92] adopted active learning that can learn the samples labeled by users interactively. Cao et al.'s method [18] also accepted user feedback through Bayesian theory. In this way, models perform better under the supervision of human knowledge.

Query-based analysis. Visual query allows users to repeatedly modify query conditions and issue queries to obtain desired results [3, 44, 63, 68, 73, 127, 160, 161, 200]. For example, Ferreira et al. [44] allowed users to query taxi trips in different geographic areas. In [160], two queries were issued to identify a low-capacity watershed-like intersection.

Monitoring. Urban monitoring models should run continuously [18, 72, 80, 82, 181]. The time-varying state of the city is reflected by the data collected by the terminal sensors. To monitor cities in (near) real time, the models should continuously integrate the streaming and historical data and output new results, such as predicted traffic congestion [72] and detected anomalies [18].

5.4 Visual analytics-assisted models

Urban visual analytics can be a great aid if the models serving the cities need to be diagnosed, adjusted, or improved. In this context, models' better performances are the purpose.

The representative examples are model diagnoses [47, 55, 57, 66, 201] and steering [12, 50, 91, 96, 120]. For example, autonomous driving models and systems include artificial-intelligence-based decision-making models that run in the urban space. Diagnose and improve them demand visual analytics system with an urban context [47, 55, 57, 66]. Fig. 10 shows Jamonnak et al.'s system for autonomous driving model diagnosis [66]. In this system, users can assess model performances within the spatial context.

Spatiotemporal analysis models may require interactive visualizations to steer, such as the voting framework for weather forecast calibration [91], tensor decomposition for spatiotemporal pattern extraction [96], traffic analysis, and forecasting [12]. For example, Liu et al. [96] designed the tree in Fig. 11 that can be split and expanded interactively to steer the tensor decomposition. In a model steering process, every step of the model is transparent, controllable, and reliable to analysts. Models move towards better under the supervision of human knowledge.

6 Challenges of designing urban visual analytics

This section summarizes four high-level challenges of designing a visual analytic approach and discusses feasible solutions.

6.1 Problem characterization

Urban visual analytics is an application-driven field. After domain experts put forward their problems and requirements, visual analytics designers must disassemble and characterize them into problems that can be addressed by visualization techniques and computational methods if needed. Characterizing the problem well requires close and long-term collaboration with experts [143]. Such a challenge generally exists in the development of other visual analytics methods [41, 182].

6.2 High-dimensionality and heterogeneity

Urban data usually involves high-dimensional and heterogeneous attributes, such as spatiotemporal [76], text [75], image [127], and 3D physical models [43], which poses challenges for urban data visualization. Two strategies can solve these challenges.

The first one is to coordinate multiple visualizations in a visual analytics system. In this way, different visualizations show different and heterogeneous dimensions of urban data, for example, in TPFlow [96] (Fig. 11) and StreetVizor [145]. Although this strategy is effective and widely used, designers should keep the interaction and context switching costs in mind. The second one is to design integrated visualizations. Different aspects of urban data are encoded together intuitively. Urban data can be analyzed within the same view and context. For example, Sun et al. [153] embedded the temporal data into graphically broadened roads. Glyphs in Fig. 2A1 visualize spatial, temporal, and categorical attributes [199]. This strategy is more difficult than coordinating visualizations because it needs to leverage limited visual channels appropriately in a confined space.

6.3 Scalability

Ubiquitous sensors in urban environments monitor the city all the time, generating massive amounts of raw data. Furthermore, the patterns hidden behind sometimes can be numerous. For accommodating the sheer volume of such urban data, scalability issues must be addressed from visualization and computation perspectives.

Visualization. To avoid being overwhelmed, users expect the data and visualizations to be well organized. The most famous organization follows the information-seeking mantra [147]: "overview first, zoom and filter, then details-on-demand." Designing an overview and drill-down interactive exploration is what designers work on according to domain requirements. If an inherent hierarchy exists, the corresponding hierarchical organization is desirable, for example, in AirVis [36] (Fig. 8) and BNVA [172]. Such a mechanism of hierarchical exploration can be called the level-of-detail mechanism [184].

In addition, graphic optimizations can make the visualizations more readable by reducing visual clutters. For example, edge bundling techniques [56] can bundle massive trajectories [202]. Element ordering can reduce visual wiggles and crossings [104, 156]. Sampling techniques that consider visual perception can reduce visual occlusion [162, 193, 206, 207, 214–216].

Computation. Excessive amounts of data also slow down computation and prevent seamless interactions. Therefore, index techniques [34, 40, 73, 75, 93, 94, 124], progressive analysis [96, 150, 172], approximation [131], GPU rendering [78, 79], etc., can be incorporated into systems for accelerating the computation. Effective indexes ensure fast access to data. Approximation sacrifices a small amount of accuracy to improve computational speed. The progressive analysis returns the results continuously rather than the final result obtained after a time-consuming computation. Progressive analysis can be adjusted or stopped on the fly. GPU rendering focuses on efficient rendering of graphic elements rather than data computation.

6.4 Uncertainty

Urban data can be inherently uncertain due to insufficient spatial and temporal granularity or imprecise and errors at sensor terminals [25, 63, 107]. The uncertainties can propagate during data processing or transformations, which may lead to wrong results. Urban experts need to access these uncertainties and thereby perform reliable analysis and make informed decisions.

There are two feasible actions to alleviate the uncertainty issue. The first one is to design uncertainty-aware visualizations [46, 62, 99, 141] for those key steps in data processing or transformations. For example, in the cascading pattern inference [37], the occurrences of instances that contradict the pattern inference result were visually exposed to users. In the natural-language-based trajectory query [63], a "relevance tree" was designed to show the uncertain semantic matching between natural language and POIs. In movement semantics enrichment [71], a gradient colormap is applied to visualize the uncertainties of assigning a POI to a destination point of movement.

Another one is to allow users to inspect the raw data for validation. This action is naive but practical in many systems [37, 38].

7 Future directions and opportunities

Although urban visual analytics has made remarkable achievements towards smart cities, there are still gaps to be filled. The ever-changing urban life also raises new requirements for urban visual analytics. This section intends to set to the ball rolling by illustrating the gaps and requirements.

7.1 Domain problem

There are two potential domain problems that could be investigated in the future.

In-situ real-time decision-making. In the traffic domain, few tools support in-situ real-time decision-making based on effective urban online monitoring and offline diagnoses, although about half of the urban visual analytics studies were concerned with traffic problems. Improving traffic efficiency is the most important issue in the traffic domain. When congestion is detected in real time, what kind of diversion measures can the department take to alleviate the congestion and improve traffic efficiency, and which one is the best? Such a scenario requires traffic monitoring and prediction, mobility analysis, and comparative analysis to be supported in visual analytics.

City-wide disease spreading. Affected by COVID-19, the issue of public health safety has been deeply rooted in the hearts of the people. There are many visualization methods for studying the spread of diseases [2, 15, 117], but few focus on infectious diseases in a fine-grained way in urban space. The health department may need an uncertain-aware visual analytics system that integrates mobility patterns and disease transmission models to assess potential risks of geographic regions and identify vulnerable individuals. At the same time, privacy protection is an important issue in such public health analyses.

Finally, we consider that the rapid development of urban analytics in the traffic domain can partially be attributed to public traffic datasets. We call on practitioners of urban analytics, whether in industry or academia, to release more high-quality and interesting data [90, 209, 213] and thereby promote the urban analytics research in various domains.

7.2 Visualizations

Two types of visualization techniques are worth investigating in the future.

Integrated visualization. We reached a consensus with the previous surveys that urban data can always be classified into spatial, temporal, and other properties [27, 210]. Many spatial, temporal, and other information visualizations have been designed for these three kinds of data, respectively. It is non-trivial to propose something novel given these excellent visual designs. Nevertheless, it is still a promising research direction to effectively integrate spatial and temporal information together [153, 158, 185, 186].

AI4VIS. Besides, rich visualization datasets [35, 58] and powerful artificial intelligence (AI) models have yielded an emerging and promising research topic called AI4VIS [166, 176]. That is, researchers start to leverage AI to generate or recommend visualizations given an input dataset. The existing AI4VIS methods mainly focused on tabular data and basic charts (e.g., bar charts) for information visualization, while the visualizations of spatiotemporal data in urban space were ignored. We can imagine that in the future, a set of effective spatiotemporal visualizations can be automatically generated and coordinated for urban visual analysis only based on simple inputs like a dataset and tasks. Realizing such a vision is also challenging and requires long-term efforts from researchers.

7.3 Computational methods

Computational methods tend to be interpretable and tailored to scenarios.

Interpretable computation. Deep learning is still rarely used in urban visual analytics, although it is already in full swing in the field of artificial intelligence [192]. Obtaining deep insights into improved cities relies on a comprehensive understanding of urban data through such as frequent occurrences by frequent pattern mining, various indicators based on rules, and spatial distribution by KDE. However, deep learning currently mainly supports prediction, classification, and representation, which are only a small part of the analysis functions. Besides, its limited interpretability prevents its use in a user-center system. Based on these observations, we believe that interpretable computational methods, such as these statistical and rule-based methods, will still be prevailing in urban visual analytics in the foreseeable future. At the same time, practitioners should make full use of the powerful capabilities of deep learning in prediction, classification, and representation.

Scenario-tailored computation. Visualization researchers proposed many practical computational methods, some of which were extended by existing ones. Urban visual analytics research aims to design human-in-the-loop solutions for improved cities rather than developing innovative computational methods. Nonetheless, a user-center system may raise unique requirements on computational methods, such as being fast enough to support seamless interaction [172, 173] or being steerable [96, 131]. These requirements motivate visualization researchers to adapt state-of-the-art computational methods or couple them with visualizations, thereby improving their practicality and effectiveness in urban visual analytics.

7.4 System

Among the four types of systems, the type of **visual analytics-assisted model** has a better prospect as the other three are well established.

Many complicated models, particularly those based on deep learning, are increasingly applied in urban scenarios, such as autonomous driving, trajectory prediction [114], and urban flow analysis [88, 89, 130]. Despite the great performances they achieved, the dynamic and complex urban environment poses challenges for deploying these models in the real world [57, 66, 201]. Urban experts require testing and debugging them in real scenarios, i.e., model diagnosis. Besides, if users directly manipulate an advanced model, the model can achieve better performance than itself only. Such a model steering mechanism is an effective way to involve human knowledge into machine intelligence. Users and machines communicate closely through interactive visualizations. The obvious advantages of model diagnosis and steering will encourage this kind of urban visual analytics system, especially if models become more powerful but complex.

7.5 Analytics environments

The development of hardware and collaboration technologies has unveiled future analytics environments for urban visual analytics, such as immersive, collaborative, and mobile environments.

Immersive environment. In recent years, the rapid development of mixed reality technologies has made immersive urban visual analytics possible [30, 33, 64, 151, 189]. Mixed reality devices can significantly empower urban analysts with a sense of presence by integrating the critical 3D context [31]. Nonetheless, immersive urban visual analytics is still in a very early stage. We have not seen an immersive tool that supports the complex urban analysis tasks like those have been done in the desktop environment. This puts forward an urgent need for researchers to explore the effective spatiotemporal visualizations and visual analytics methods in an immersive environment.

Collaborative environment. A collaborative analytics environment is conducive to urban visual analysis. The investigation of urban phenomenons may involve multiple fields, and thus the collaboration among the experts in these fields is generally necessary. For example, air pollution can be caused by local traffic congestion or the air pollutants propagated from remote regions. In such a case, transportation experts and environmental experts need to cooperate in analyzing the problem. How to apply existing collaborative analysis methods [65, 142] to urban visual analytics and what challenges will be encountered during the application remain open problems.

Mobile environment. Nowadays, mobile devices, such as mobile phones and tablets, has become the most accessible analytics terminals. For example, a police officer could regulate traffic at a crossroad with a tablet, which empowers him/her to perform traffic analysis and take actions in real time, instead of relying on a remote command center. Unfortunately, no urban visual analytics method has been developed for the mobile environment. One potential reason could be the limited computing capability and screen size available for complex data visualization. Although researchers have made some preliminary attempts in mobile information visualization [16, 17, 174, 175], how to design and develop urban visual analytics systems on mobile devices needs further investigation.

8 Conclusion

Urban visual analytics has been an effective way towards smart cities. Developing an urban visual analytics system demands the domain problem characterization and combination of visualizations and computation. This paper reviews the research progress in urban visual analytics from four perspectives of domain problem, visualization, computational analysis, and system. Visualization and urban analysis practitioners can fully understand the state-of-the-art urban visual analytics, the development guidance of visualization systems, and future research directions and opportunities. We published an interactive tool for exploring the surveyed papers based on our proposed typologies: `https://urban-va-survey.github.io/`.

9 Declarations

Availability of data and materials. All papers investigated in this survey are accessible via the corresponding publishers' websites.

Competing interests. No competing interests.

Funding. This work was supported by NSFC (62072400) and partially the Collaborative Innovation Center of Artificial Intelligence by MOE and Zhejiang Provincial Government (ZJU).

Authors' contributions. All authors are involved in the survey and paper writing.

Acknowledgement. We thank all reviewers, editors, and the funding agencies mentioned above.

References

[1] P. Accorsi, N. Lalande, M. Fabrègue, A. Braud, P. Poncelet, A. Sallaberry, S. Bringay, M. Teisseire, F. Cernesson, and F. L. Ber. HydroQual: Visual analysis of river water quality. In *Proceedings of IEEE VAST*, pages 123–132, 2014.

[2] S. Afzal, R. Maciejewski, and D. S. Ebert. Visual analytics decision support environment for epidemic modeling and response evaluation. In *Proceedings of IEEE VAST*, pages 191–200, 2011.

[3] S. Al-Dohuki, Y. Wu, F. Kamw, J. Yang, X. Li, Y. Zhao, X. Ye, W. Chen, C. Ma, and F. Wang. SemanticTraj: A new approach to interacting with massive taxi trajectories. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):11–20, 2017.

[4] G. L. Andrienko, N. V. Andrienko, H. Bosch, T. Ertl, G. Fuchs, P. Jankowski, and D. Thom. Thematic patterns in georeferenced tweets through space-time visual analytics. *Computing in Science and Engineering*, 15(3):72–82, 2013.

[5] G. L. Andrienko, N. V. Andrienko, W. Chen, R. Maciejewski, and Y. Zhao. Visual analytics of mobility and transportation: State of the art and further research directions. *IEEE Transactions on Intelligent Transportation Systems*, 18(8):2232–2249, 2017.

[6] G. L. Andrienko, N. V. Andrienko, G. Fuchs, and J. Wood. Revealing patterns and trends of mass mobility through spatial and temporal abstraction of origin-destination movement data. *IEEE Transactions on Visualization and Computer Graphics*, 23(9):2120–2136, 2017.

[7] G. L. Andrienko, N. V. Andrienko, C. Hurter, S. Rinzivillo, and S. Wrobel. From movement tracks through events to places: Extracting and characterizing significant places from mobility data. In *Proceedings of IEEE VAST*, pages 161–170, 2011.

[8] G. L. Andrienko, N. V. Andrienko, C. Hurter, S. Rinzivillo, and S. Wrobel. Scalable analysis of movement data for extracting and exploring significant places. *IEEE Transactions on Visualization and Computer Graphics*, 19(7):1078–1094, 2013.

[9] G. L. Andrienko, N. V. Andrienko, M. Mladenov, M. Mock, and C. Pölitz. Discovering bits of place histories from people's activity traces. In *Proceedings of IEEE VAST*, pages 59–66, 2010.

[10] N. V. Andrienko and G. L. Andrienko. Spatial generalization and aggregation of massive movement data. *IEEE Transactions on Visualization and Computer Graphics*, 17(2):205–219, 2011.

[11] N. V. Andrienko, G. L. Andrienko, F. Patterson, and H. Stange. Visual analysis of place connectedness by public transport. *IEEE Transactions on Intelligent Transportation Systems*, 21(8):3196–3208, 2020.

[12] N. V. Andrienko, G. L. Andrienko, and S. Rinzivillo. Leveraging spatial abstraction in traffic analysis and forecasting with visual analytics. *Information Systems*, 57:172–194, 2016.

[13] V. P. Araya, A. Bezerianos, and E. Pietriga. A comparison of geographical propagation visualizations. In *Proceedings of ACM CHI*, pages 1–14, 2020.

[14] S. M. Arietta, A. A. Efros, R. Ramamoorthi, and M. Agrawala. City forensics: Using visual elements to predict non-visual city attributes. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2624–2633, 2014.

[15] T. Baumgartl, M. Petzold, M. Wunderlich, M. Höhn, D. Archambault, M. Lieser, A. Dalpke, S. Scheithauer, M. Marschollek, V. Eichel, N. T. Mutters, H. Consortium, and T. von Landesberger. In search of patient zero: Visual analytics of pathogen transmission pathways in hospitals. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):711–721, 2021.

[16] M. Brehmer, B. Lee, P. Isenberg, and E. K. Choe. Visualizing ranges over time on mobile phones: A task-based crowdsourced evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):619–629, 2019.

[17] M. Brehmer, B. Lee, P. Isenberg, and E. K. Choe. A comparative evaluation of animation and small multiples for trend visualization on mobile phones. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):364–374, 2020.

[18] N. Cao, C. Lin, Q. Zhu, Y. Lin, X. Teng, and X. Wen. Voila: Visual anomaly detection and monitoring with streaming spatiotemporal data. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):23–33, 2018.

[19] N. Cao, Y. Lin, X. Sun, D. Lazer, S. Liu, and H. Qu. Whisper: Tracing the spatiotemporal process of information diffusion in real time. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2649–2658, 2012.

[20] G. Carenini and J. Loyd. ValueCharts: analyzing linear models expressing preferences and evaluations. In *Proceedings of the working conference on Advanced visual interfaces*, pages 150–157. ACM Press, 2004.

[21] J. Chae, D. Thom, H. Bosch, Y. Jang, R. Maciejewski, D. S. Ebert, and T. Ertl. Spatiotemporal social media analytics for abnormal event detection and examination using seasonal-trend decomposition. In *Proceedings of IEEE VAST*, pages 143–152, 2012.

[22] R. Chen, X. Shu, J. Chen, D. Weng, J. Tang, S. Fu, , and Y. Wu. Nebula: A coordinating grammar of graphics. *IEEE Transactions on Visualization and Computer Graphics*, 2021.

[23] S. Chen, S. Chen, Z. Wang, J. Liang, X. Yuan, N. Cao, and Y. Wu. D-Map: Visual analysis of ego-centric information diffusion patterns in social media. In *Proceedings of IEEE VAST*, pages 41–50, 2016.

[24] S. Chen, L. Lin, and X. Yuan. Social media visual analytics. *Computer Graphics Forum*, 36(3):563–587, 2017.

[25] S. Chen, Z. Wang, J. Liang, and X. Yuan. Uncertainty-aware visual analytics for exploring human behaviors from heterogeneous spatial temporal

data. *Journal of Visual Languages and Computing*, 48:187–198, 2018.

[26] S. Chen, X. Yuan, Z. Wang, C. Guo, J. Liang, Z. Wang, X. L. Zhang, and J. Zhang. Interactive visual discovering of movement patterns from sparsely sampled geo-tagged social media data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):270–279, 2016.

[27] W. Chen, F. Guo, and F. Wang. A survey of traffic data visualization. *IEEE Transactions on Intelligent Transportation Systems*, 16(6):2970–2984, 2015.

[28] W. Chen, Z. Huang, F. Wu, M. Zhu, H. Guan, and R. Maciejewski. VAUD: A visual analysis approach for exploring spatio-temporal urban data. *IEEE Transactions on Visualization and Computer Graphics*, 24(9):2636–2648, 2018.

[29] W. Chen, J. Xia, X. Wang, Y. Wang, J. Chen, and L. Chang. RelationLines: Visual reasoning of egocentric relations from heterogeneous urban data. *ACM Transactions on Intelligent Systems and Technology*, 10(1):2:1–2:21, 2019.

[30] Z. Chen, Y. Su, Y. Wang, Q. Wang, H. Qu, and Y. Wu. MARVisT: Authoring glyph-based visualization in mobile augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 26(8):2645–2658, 2020.

[31] Z. Chen, Y. Wang, T. Sun, X. Gao, W. Chen, Z. Pan, H. Qu, and Y. Wu. Exploring the design space of immersive urban analytics. *Visual Informatics*, 1(2):132–142, 2017.

[32] D. Chu, D. A. Sheets, Y. Zhao, Y. Wu, J. Yang, M. Zheng, and G. Chen. Visualizing hidden themes of taxi movement with semantic transformation. In *Proceedings of IEEE Pacific Visualization Symposium*, pages 137–144, 2014.

[33] X. Chu, X. Xie, S. Ye, H. Lu, H. Xiao, Z. Yuan, Z. Chen, H. Zhang, and Y. Wu. TIVEE: visual exploration and explanation of badminton tactics in immersive visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):118–128, 2022.

[34] C. A. de Lara Pahins, S. A. Stephens, C. Scheidegger, and J. L. D. Comba. Hashedcubes: Simple, low memory, real-time visual exploration of big data. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):671–680, 2017.

[35] D. Deng, Y. Wu, X. Shu, J. Wu, M. Xu, S. Fu, W. Cui, and Y. Wu. VisImages: A corpus of visualizations in the images of visualization publications, 2021.

[36] Z. Deng, D. Weng, J. Chen, R. Liu, Z. Wang, J. Bao, Y. Zheng, and Y. Wu. AirVis: Visual analytics of air pollution propagation. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):800–810, 2020.

[37] Z. Deng, D. Weng, Y. Liang, J. Bao, Y. Zheng, T. Schreck, M. Xu, and Y. Wu. Visual cascade analytics of large-scale spatiotemporal data. *IEEE Transactions on Visualization and Computer Graphics*, 2021.

[38] Z. Deng, D. Weng, X. Xie, J. Bao, Y. Zheng, M. Xu, W. Chen, and Y. Wu. Compass: Towards better causal analysis of urban time series. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):1051–1061, 2022.

[39] H. Doraiswamy, N. Ferreira, T. Damoulas, J. Freire, and C. T. Silva. Using topological analysis to support event-guided exploration in urban data. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2634–2643, 2014.

[40] H. Doraiswamy, H. T. Vo, C. T. Silva, and J. Freire. A gpu-based index to support interactive spatio-temporal queries over historical data. In *Proceedings of IEEE International Conference on Data Engineering*, pages 1086–1097, 2016.

[41] J. Eirich, J. Bonart, D. Jäckle, M. Sedlmair, U. Schmid, K. Fischbach, T. Schreck, and J. Bernard. IRVINE: A design study on analyzing correlation patterns of electrical engines. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):11–21, 2022.

[42] Z. Feng, H. Li, W. Zeng, S. Yang, and H. Qu. Topology density map for urban data visualization and analysis. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):828–838, 2021.

[43] N. Ferreira, M. Lage, H. Doraiswamy, H. T. Vo, L. Wilson, H. Werner, M. Park, and C. T. Silva. Urbane: A 3d framework to support data driven decision making in urban development. In *Proceedings of IEEE VAST*, pages 97–104, 2015.

[44] 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 Transactions on Visualization and Computer Graphics*, 19(12):2149–2158, 2013.

[45] J. Gautier, M. Brédif, and S. Christophe. Co-visualization of air temperature and urban data for visual exploration. In *Proceedings of IEEE VIS (Short Papers)*, pages 71–75, 2020.

[46] J. Görtler, C. Schulz, D. Weiskopf, and O. Deussen. Bubble treemaps for uncertainty visualization. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):719–728, 2018.

[47] L. Gou, L. Zou, N. Li, M. Hofmann, A. K. Shekar, A. Wendt, and L. Ren. VATLD: A visual analytics system to assess, understand and improve traffic light detection. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):261–271, 2021.

[48] S. Gratzl, A. Lex, N. Gehlenborg, H. Pfister, and M. Streit. LineUp: Visual analysis of multi-attribute rankings. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2277–2286, 2013.

[49] T. Gu, M. Zhu, W. Chen, Z. Huang, R. Maciejewski, and L. Chang. Structuring mobility transition with an adaptive graph representation. *IEEE Transactions on*

Computational Social Systems, 5(4):1121–1132, 2018.

[50] F. Guo, T. Gu, W. Chen, F. Wu, Q. Wang, L. Shi, and H. Qu. Visual exploration of air quality data with a time-correlation-partitioning tree based on information theory. *ACM Transactions on Interactive Intelligent Systems*, 9(1):4:1–4:23, 2019.

[51] 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 *Proceedings of IEEE Pacific Visualization Symposium*, pages 163–170, 2011.

[52] Y. Guo, S. Guo, Z. Jin, S. Kaul, D. Gotz, and N. Cao. A survey on visual analysis of event sequence data. *IEEE Transactions on Visualization and Computer Graphics*, 2021.

[53] S. Han, S. Ye, and H. Zhang. Visual exploration of internet news via sentiment score and topic models. *Computional Visual Media*, 6(3):333–347, 2020.

[54] T. He, J. Bao, S. Ruan, R. Li, Y. Li, H. He, and Y. Zheng. Interactive bike lane planning using sharing bikes' trajectories. *IEEE Transactions on Knowledge and Data Engineering*, 32(8):1529–1542, 2020.

[55] W. He, L. Zou, A. K. Shekar, L. Gou, and L. Ren. Where can we help? A visual analytics approach to diagnosing and improving semantic segmentation of movable objects. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):1040–1050, 2022.

[56] D. Holten. Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):741–748, 2006.

[57] Y. Hou, C. Wang, J. Wang, X. Xue, X. L. Zhang, J. Zhu, D. Wang, and S. Chen. Visual evaluation for autonomous driving. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):1030–1039, 2022.

[58] K. Z. Hu, S. N. S. Gaikwad, M. Hulsebos, M. A. Bakker, E. Zgraggen, C. A. Hidalgo, T. Kraska, G. Li, A. Satyanarayan, and Ç. Demiralp. VizNet: Towards a large-scale visualization learning and benchmarking repository. In *Proceedings of ACM CHI*, page 662, 2019.

[59] M. Hu, S. Liu, F. Wei, Y. Wu, J. T. Stasko, and K. Ma. Breaking news on twitter. In *Proceedings of ACM CHI*, pages 2751–2754. ACM, 2012.

[60] K. Huang. Mapping the hazard: Visual analysis of flood impact on urban mobility. *IEEE Computer Graphics and Applications*, 41(1):26–34, 2021.

[61] X. Huang, Y. Zhao, C. Ma, J. Yang, X. Ye, and C. Zhang. TrajGraph: A graph-based visual analytics approach to studying urban network centralities using taxi trajectory data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):160–169, 2016.

[62] Z. Huang, Y. Lu, E. A. Mack, W. Chen, and R. Maciejewski. Exploring the sensitivity of choropleths under attribute uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 26(8):2576–2590, 2020.

[63] Z. Huang, Y. Zhao, W. Chen, S. Gao, K. Yu, W. Xu, M. Tang, M. Zhu, and M. Xu. A natural-language-based visual query approach of uncertain human trajectories. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):1256–1266, 2020.

[64] C. Hurter, N. H. Riche, S. M. Drucker, M. Cordeil, R. Alligier, and R. Vuillemot. FiberClay: Sculpting three dimensional trajectories to reveal structural insights. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):704–714, 2019.

[65] P. Isenberg, D. Fisher, S. A. Paul, M. R. Morris, K. Inkpen, and M. Czerwinski. Co-located collaborative visual analytics around a tabletop display. *IEEE Transactions on visualization and Computer Graphics*, 18(5):689–702, 2011.

[66] S. Jamonnak, Y. Zhao, X. Huang, and M. Amiruzzaman. Geo-context aware study of vision-based autonomous driving models and spatial video data. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):1019–1029, 2022.

[67] Z. Jin, N. Cao, Y. Shi, W. Wu, and Y. Wu. EcoLens: Visual analysis of ecological regions in urban contexts using traffic data. *Journal of Visualization*, 24(2):349–364, 2021.

[68] F. Kamw, S. Al-Dohuki, Y. Zhao, T. Eynon, D. A. Sheets, J. Yang, X. Ye, and W. Chen. Urban structure accessibility modeling and visualization for joint spatiotemporal constraints. *IEEE Transactions on Intelligent Transportation Systems*, 21(1):104–116, 2020.

[69] S. Kim, S. Jeong, I. Woo, Y. Jang, R. Maciejewski, and D. S. Ebert. Data flow analysis and visualization for spatiotemporal statistical data without trajectory information. *IEEE Transactions on Visualization and Computer Graphics*, 24(3):1287–1300, 2018.

[70] J. Knittel, S. Koch, T. Tang, W. Chen, Y. Wu, S. Liu, and T. Ertl. Real-time visual analysis of high-volume social media posts. *IEEE Transactions on Visualization and Computer Graphics*, 2022.

[71] R. Krüger, D. Thom, and T. Ertl. Semantic enrichment of movement behavior with foursquare-a visual analytics approach. *IEEE Transactions on Visualization and Computer Graphics*, 21(8):903–915, 2015.

[72] C. Lee, Y. Kim, S. Jin, D. Kim, R. Maciejewski, D. S. Ebert, and S. Ko. A visual analytics system for exploring, monitoring, and forecasting road traffic congestion. *IEEE Transactions on Visualization and Computer Graphics*, 26(11):3133–3146, 2020.

[73] C. Li, G. Baciu, Y. Wang, J. Chen, and C. Wang. DDLVis: Real-time visual query of spatiotemporal data distribution via density dictionary learning. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):1062–1072, 2022.

[74] C. Li, X. Dong, and X. Yuan. Metro-Wordle: An interactive visualization for urban text distributions based on wordle. *Visual Informatics*, 2(1):50–59, 2018.

[75] J. Li, S. Chen, W. Chen, G. L. Andrienko, and N. V. Andrienko. Semantics-space-time cube: A conceptual framework for systematic analysis of texts in space and time. *IEEE Transactions on Visualization and Computer Graphics*, 26(4):1789–1806, 2020.

[76] J. Li, S. Chen, K. Zhang, G. L. Andrienko, and N. V. Andrienko. COPE: interactive exploration of co-occurrence patterns in spatial time series. *IEEE Transactions on Visualization and Computer Graphics*, 25(8):2554–2567, 2019.

[77] J. Li, K. Zhang, and Z. Meng. Vismate: Interactive visual analysis of station-based observation data on climate changes. In *Proceedings of IEEE VAST*, pages 133–142, 2014.

[78] J. K. Li and K. Ma. P4: portable parallel processing pipelines for interactive information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 26(3):1548–1561, 2020.

[79] J. K. Li and K. Ma. P5: portable progressive parallel processing pipelines for interactive data analysis and visualization. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):1151–1160, 2020.

[80] Q. Li, H. Lin, X. Wei, Y. Huang, L. Fan, J. Du, X. Ma, and T. Chen. MaraVis: Representation and coordinated intervention of medical encounters in urban marathon. In *Proceedings of ACMCHI*, pages 1–12, 2020.

[81] Q. Li, Q. Q. Liu, C. F. Tang, Z. W. Li, S. C. Wei, X. R. Peng, M. H. Zheng, T. J. Chen, and Q. Yang. WarehouseVis: A visual analytics approach to facilitating warehouse location selection for business districts. *Computer Graphics Forum*, 39(3):483–495, 2020.

[82] Q. Li, Y. J. Liu, L. Chen, X. C. Yang, Y. Peng, X. Yuan, and M. L. L. Wijerathne. SEEVis: A smart emergency evacuation plan visualization system with data-driven shot designs. *Computer Graphics Forum*, 39(3):523–535, 2020.

[83] R. Li, H. He, R. Wang, Y. Huang, J. Liu, S. Ruan, T. He, J. Bao, and Y. Zheng. JUST: JD urban spatio-temporal data engine. In *Proceedings of IEEE International Conference on Data Engineering*, pages 1558–1569, 2020.

[84] R. Li, S. Ruan, J. Bao, Y. Li, Y. Wu, L. Hong, and Y. Zheng. Efficient path query processing over massive trajectories on the cloud. *IEEE Transactions on Big Data*, 6(1):66–79, 2020.

[85] Y. Li, J. Bao, Y. Li, Y. Wu, Z. Gong, and Y. Zheng. Mining the most influential k-location set from massive trajectories. In *Proceedings of ACM SIGSPATIAL*, pages 51:1–51:4, 2016.

[86] Y. Li, J. Bao, Y. Li, Y. Wu, Z. Gong, and Y. Zheng. Mining the most influential k-location set from massive trajectories. *IEEE Transactions on Big Data*, 4(4):556–570, 2018.

[87] Y. Liang, Z. Jiang, and Y. Zheng. Inferring traffic cascading patterns. In *Proceedings of ACM SIGSPATIAL*, pages 2:1–2:10. ACM, 2017.

[88] Y. Liang, K. Ouyang, L. Jing, S. Ruan, Y. Liu, J. Zhang, D. S. Rosenblum, and Y. Zheng. UrbanFM: Inferring fine-grained urban flows. In *Proceedings of ACM SIGKDD*, pages 3132–3142, 2019.

[89] Y. Liang, K. Ouyang, J. Sun, Y. Wang, J. Zhang, Y. Zheng, D. S. Rosenblum, and R. Zimmermann. Fine-grained urban flow prediction. In *Proceedings of WWW*, pages 1833–1845. ACM / IW3C2, 2021.

[90] B. Liao, J. Zhang, C. Wu, D. McIlwraith, T. Chen, S. Yang, Y. Guo, and F. Wu. Deep sequence learning with auxiliary information for traffic prediction. In *Proceedings of ACM SIGKDD*, pages 537–546, 2018.

[91] H. Liao, Y. Wu, L. Chen, T. M. Hamill, Y. Wang, K. Dai, H. Zhang, and W. Chen. A visual voting framework for weather forecast calibration. In *IEEE SciVis*, pages 25–32, 2015.

[92] Z. Liao, Y. Yu, and B. Chen. Anomaly detection in GPS data based on visual analytics. In *Proceedings of IEEE VAST*, pages 51–58, 2010.

[93] L. D. Lins, J. T. Klosowski, and C. E. Scheidegger. Nanocubes for real-time exploration of spatiotemporal datasets. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2456–2465, 2013.

[94] C. Liu, C. Wu, H. Shao, and X. Yuan. SmartCube: An adaptive data management architecture for the real-time visualization of spatiotemporal datasets. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):790–799, 2020.

[95] D. Liu, D. Weng, Y. Li, J. Bao, Y. Zheng, H. Qu, and Y. Wu. SmartAdP: Visual analytics of large-scale taxi trajectories for selecting billboard locations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):1–10, 2017.

[96] D. Liu, P. Xu, and L. Ren. TPFlow: Progressive partition and multidimensional pattern extraction for large-scale spatio-temporal data analysis. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):1–11, 2019.

[97] H. Liu, X. Chen, Y. Wang, B. Zhang, Y. Chen, Y. Zhao, and F. Zhou. Visualization and visual analysis of vessel trajectory data: A survey. *Visual Informatics*, 5(4):1–10, 2021.

[98] H. Liu, Y. Gao, L. Lu, S. Liu, H. Qu, and L. M. Ni. Visual analysis of route diversity. In *Proceedings of IEEE VAST*, pages 171–180, 2011.

[99] M. Liu, S. Liu, X. Zhu, Q. Liao, F. Wei, and S. Pan. An uncertainty-aware approach for exploratory microblog retrieval. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):250–259, 2016.

[100] Q. Q. Liu, Q. Li, C. F. Tang, H. Lin, X. Ma, and T. Chen. A visual analytics approach to scheduling customized shuttle buses via perceiving passengers'

travel demands. In *Proceedings of IEEE VIS (Short Papers)*, pages 76–80, 2020.

[101] Q. Q. Liu, Q. Li, C. F. Tang, H. B. Lin, Z. Peng, Z. W. Li, and T. Chen. Visual analysis of car-hailing reimbursement data for overtime. In *Proceedings of EuroVis (Posters)*, pages 21–23. Eurographics Association, 2020.

[102] S. Liu, W. Cui, Y. Wu, and M. Liu. A survey on information visualization: Recent advances and challenges. *Visual Computer*, 30(12):1373–1393, 2014.

[103] S. Liu, J. Pu, Q. Luo, H. Qu, L. M. Ni, and R. Krishnan. VAIT: A visual analytics system for metropolitan transportation. *IEEE Transactions on Intelligent Transportation Systems*, 14(4):1586–1596, 2013.

[104] S. Liu, Y. Wu, E. Wei, M. Liu, and Y. Liu. Storyflow: Tracking the evolution of stories. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2436–2445, 2013.

[105] G. D. Lorenzo, M. L. Sbodio, F. Calabrese, M. Berlingerio, F. Pinelli, and R. Nair. AllAboard: Visual exploration of cellphone mobility data to optimise public transport. *IEEE Transactions on Visualization and Computer Graphics*, 22(2):1036–1050, 2016.

[106] Y. Lou, C. Zhang, Y. Zheng, X. Xie, W. Wang, and Y. Huang. Map-matching for low-sampling-rate GPS trajectories. In *Proceedings of ACM SIGSPATIAL*, pages 352–361, 2009.

[107] L. Lu, N. Cao, S. Liu, L. M. Ni, X. Yuan, and H. Qu. Visual analysis of uncertainty in trajectories. In *Proceedings of PAKDD*, volume 8443 of *Lecture Notes in Computer Science*, pages 509–520. Springer, 2014.

[108] M. Lu, C. Lai, T. Ye, J. Liang, and X. Yuan. Visual analysis of multiple route choices based on general GPS trajectories. *IEEE Transactions on Big Data*, 3(2):234–247, 2017.

[109] M. Lu, J. Liang, Z. Wang, and X. Yuan. Exploring OD patterns of interested region based on taxi trajectories. *Journal of Visualization*, 19(4):811–821, 2016.

[110] M. Lu, Z. Wang, J. Liang, and X. Yuan. OD-Wheel: Visual design to explore OD patterns of a central region. In *Proceedings of IEEE Pacific Visualization Symposium*, pages 87–91, 2015.

[111] M. Lu, Z. Wang, and X. Yuan. TrajRank: Exploring travel behaviour on a route by trajectory ranking. In *Proceedings of IEEE Pacific Visualization Symposium*, pages 311–318, 2015.

[112] Y. Lu, X. Hu, F. Wang, S. Kumar, H. Liu, and R. Maciejewski. Visualizing social media sentiment in disaster scenarios. In *Proceedings of WWW Companion*, pages 1211–1215. ACM, 2015.

[113] J. Lukasczyk, R. Maciejewski, C. Garth, and H. Hagen. Understanding hotspots: A topological visual analytics approach. In *Proceedings of ACM SIGSPATIAL*, pages 36:1–36:10, 2015.

[114] P. Lv, H. Wei, T. Gu, Y. Zhang, X. Jiang, B. Zhou, and M. Xu. Trajectory distributions: A new description of movement for trajectory prediction. *Computional Visual Media*, 8(2):213–224, 2022.

[115] Y. Ma, T. Lin, Z. Cao, C. Li, F. Wang, and W. Chen. Mobility viewer: An eulerian approach for studying urban crowd flow. *IEEE Transactions on Intelligent Transportation Systems*, 17(9):2627–2636, 2016.

[116] A. M. MacEachren, A. R. Jaiswal, A. C. Robinson, S. Pezanowski, A. Savelyev, P. Mitra, X. Zhang, and J. I. Blanford. SensePlace2: Geotwitter analytics support for situational awareness. In *Proceedings of IEEE VAST*, pages 181–190, 2011.

[117] R. Maciejewski, R. Hafen, S. Rudolph, S. G. Larew, M. A. Mitchell, W. S. Cleveland, and D. S. Ebert. Forecasting hotspots - A predictive analytics approach. *IEEE Transactions on Visualization and Computer Graphics*, 17(4):440–453, 2011.

[118] R. Maciejewski, S. Rudolph, R. Hafen, A. M. Abusalah, M. Yakout, M. Ouzzani, W. S. Cleveland, S. J. Grannis, and D. S. Ebert. A visual analytics approach to understanding spatiotemporal hotspots. *IEEE Transactions on Visualization and Computer Graphics*, 16(2):205–220, 2010.

[119] R. Maciejewski, B. Tyner, Y. Jang, C. Zheng, R. V. Nehme, D. S. Ebert, W. S. Cleveland, M. Ouzzani, S. J. Grannis, and L. T. Glickman. LAHVA: linked animal-human health visual analytics. In *Proceedings of IEEE VAST*, pages 27–34, 2007.

[120] A. Malik, R. Maciejewski, N. Elmqvist, Y. Jang, D. S. Ebert, and W. Huang. A correlative analysis process in a visual analytics environment. In *Proceedings of IEEE VAST*, pages 33–42, 2012.

[121] A. Malik, R. Maciejewski, B. Maule, and D. S. Ebert. A visual analytics process for maritime resource allocation and risk assessment. In *Proceedings of IEEE VAST*, pages 221–230, 2011.

[122] A. Malik, R. Maciejewski, S. Towers, S. McCullough, and D. S. Ebert. Proactive spatiotemporal resource allocation and predictive visual analytics for community policing and law enforcement. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1863–1872, 2014.

[123] A. H. Meghdadi and P. Irani. Interactive exploration of surveillance video through action shot summarization and trajectory visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2119–2128, 2013.

[124] H. Mei, W. Chen, Y. Wei, Y. Hu, S. Zhou, B. Lin, Y. Zhao, and J. Xia. RSATree: Distribution-aware data representation of large-scale tabular datasets for flexible visual query. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):1161–1171, 2020.

[125] F. Miranda, H. Doraiswamy, M. Lage, L. Wilson, M. Hsieh, and C. T. Silva. Shadow accrual maps:

Efficient accumulation of city-scale shadows over time. *IEEE Transactions on Visualization and Computer Graphics*, 25(3):1559–1574, 2019.

[126] F. Miranda, H. Doraiswamy, M. Lage, K. Zhao, B. Gonçalves, L. Wilson, M. Hsieh, and C. T. Silva. Urban pulse: Capturing the rhythm of cities. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):791–800, 2017.

[127] F. Miranda, M. Hosseini, M. Lage, H. Doraiswamy, G. Dove, and C. T. Silva. Urban mosaic: Visual exploration of streetscapes using large-scale image data. In *Proceedings of ACM CHI*, pages 1–15, 2020.

[128] Q. V. Nguyen, N. Miller, D. Arness, W. Huang, M. L. Huang, and S. Simoff. Evaluation on interactive visualization data with scatterplots. *Visual Informatics*, 4(4):1–10, 2020.

[129] C. Palomo, Z. Guo, C. T. Silva, and J. Freire. Visually exploring transportation schedules. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):170–179, 2016.

[130] Z. Pan, Y. Liang, W. Wang, Y. Yu, Y. Zheng, and J. Zhang. Urban traffic prediction from spatio-temporal data using deep meta learning. In *Proceedings of ACM SIGKDD*, pages 1720–1730, 2019.

[131] N. Pezzotti, B. P. F. Lelieveldt, L. van der Maaten, T. Höllt, E. Eisemann, and A. Vilanova. Approximated and user steerable tSNE for progressive visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 23(7):1739–1752, 2017.

[132] M. Pi, H. Yeon, H. Son, and Y. Jang. Visual cause analytics for traffic congestion. *IEEE Transactions on Visualization and Computer Graphics*, 27(3):2186–2201, 2021.

[133] H. Piringer, M. Buchetics, and R. Benedik. AlVis: Situation awareness in the surveillance of road tunnels. In *Proceedings of IEEE VAST*, pages 153–162, 2012.

[134] J. Poco, H. Doraiswamy, H. T. Vo, J. L. D. Comba, J. Freire, and C. T. Silva. Exploring traffic dynamics in urban environments using vector-valued functions. *Computer Graphics Forum*, 34(3):161–170, 2015.

[135] J. Pu, S. Liu, Y. Ding, H. Qu, and L. M. Ni. T-Watcher: A new visual analytic system for effective traffic surveillance. In *Proceedings of IEEE International Conference on Mobile Data Management*, pages 127–136, 2013.

[136] H. Qu, W. Chan, A. Xu, K. Chung, A. K. Lau, and P. Guo. Visual analysis of the air pollution problem in hong kong. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1408–1415, 2007.

[137] H. Qu, H. Wang, W. Cui, Y. Wu, and M. Chan. Focus+context route zooming and information overlay in 3d urban environments. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1547–1554, 2009.

[138] P. S. Quinan and M. D. Meyer. Visually comparing weather features in forecasts. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):389–398, 2016.

[139] R. Scheepens, N. Willems, H. van de Wetering, G. L. Andrienko, N. V. Andrienko, and J. J. van Wijk. Composite density maps for multivariate trajectories. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2518–2527, 2011.

[140] S. Schöttler, Y. Yang, H. Pfister, and B. Bach. Visualizing and interacting with geospatial networks: A survey and design space. *Computer Graphics Forum*, 40(6):5–33, 2021.

[141] C. Schulz, A. Nocaj, J. Görtler, O. Deussen, U. Brandes, and D. Weiskopf. Probabilistic graph layout for uncertain network visualization. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):531–540, 2017.

[142] M. Schwab, D. Saffo, Y. Zhang, S. Sinha, C. Nita-Rotaru, J. Tompkin, C. Dunne, and M. A. Borkin. Visconnect: Distributed event synchronization for collaborative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):347–357, 2021.

[143] M. Sedlmair, M. D. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2431–2440, 2012.

[144] Q. Shen, Y. Wu, Y. Jiang, W. Zeng, A. K. Lau, A. Vianova, and H. Qu. Visual interpretation of recurrent neural network on multi-dimensional time-series forecast. In *Proceedings of IEEE Pacific Visualization Symposium*, pages 61–70, 2020.

[145] Q. Shen, W. Zeng, Y. Ye, S. M. Arisona, S. Schubiger, R. Burkhard, and H. Qu. StreetVizor: Visual exploration of human-scale urban forms based on street views. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):1004–1013, 2018.

[146] L. Shi, C. Huang, M. Liu, J. Yan, T. Jiang, Z. Tan, Y. Hu, W. Chen, and X. Zhang. UrbanMotion: Visual analysis of metropolitan-scale sparse trajectories. *IEEE Transactions on Visualization and Computer Graphics*, 27(10):3881–3899, 2021.

[147] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of IEEE Symposium on Visual Languages*, pages 336–343, 1996.

[148] X. Shu, J. Wu, X. Wu, H. Liang, W. Cui, Y. Wu, and H. Qu. DancingWords: Exploring animated word clouds to tell stories. *Journal of Visualization*, 24(1):85–100, 2021.

[149] M. Steptoe, R. Krüger, R. Garcia, X. Liang, and R. Maciejewski. A visual analytics framework for exploring theme park dynamics. *ACM Transactions on Interactive Intelligent Systems*, 8(1):4:1–4:27, 2018.

[150] C. D. Stolper, A. Perer, and D. Gotz. Progressive visual analytics: User-driven visual exploration

of in-progress analytics. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1653–1662, 2014.

[151] C. Su, C. Yang, Y. Chen, F. Wang, F. Wang, Y. Wu, and X. Zhang. Natural multimodal interaction in immersive flow visualization. *Visual Informatics*, 5(4):56–66, 2021.

[152] G. Sun, B. Chang, L. Zhu, H. Wu, K. Zheng, and R. Liang. TZVis: Visual analysis of bicycle data for traffic zone division. *Journal of Visualization*, 22(6):1193–1208, 2019.

[153] G. Sun, R. Liang, H. Qu, and Y. Wu. Embedding spatio-temporal information into maps by route-zooming. *IEEE Transactions on Visualization and Computer Graphics*, 23(5):1506–1519, 2017.

[154] G. Sun, R. Liang, F. Wu, and H. Qu. A web-based visual analytics system for real estate data. *Science China Information Sciences*, 56(5):1–13, 2013.

[155] G. Sun, Y. Liu, W. Wu, R. Liang, and H. Qu. Embedding temporal display into maps for occlusion-free visualization of spatio-temporal data. In *Proceedings of IEEE Pacific Visualization Symposium*, pages 185–192, 2014.

[156] G. Sun, Y. Wu, S. Liu, T. Peng, J. J. H. Zhu, and R. Liang. EvoRiver: Visual analysis of topic coopetition on social media. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1753–1762, 2014.

[157] J. Tang, Y. Zhou, T. Tang, D. Weng, B. Xie, L. Yu, H. Zhang, and Y. Wu. A visualization approach for monitoring order processing in e-commerce warehouse. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):857–867, 2022.

[158] C. Tominski, H. Schumann, G. L. Andrienko, and N. V. Andrienko. Stacking-based visualization of trajectory attribute data. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2565–2574, 2012.

[159] T. von Landesberger, F. Brodkorb, P. Roskosch, N. V. Andrienko, G. L. Andrienko, and A. Kerren. MobilityGraphs: Visual analysis of mass mobility dynamics via spatio-temporal graphs and clustering. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):11–20, 2016.

[160] F. Wang, W. Chen, F. Wu, Y. Zhao, H. Hong, T. Gu, L. Wang, R. Liang, and H. Bao. A visual reasoning approach for data-driven transport assessment on urban roads. In *Proceedings of IEEE VAST*, pages 103–112, 2014.

[161] F. Wang, W. Chen, Y. Zhao, T. Gu, S. Gao, and H. Bao. Adaptively exploring population mobility patterns in flow visualization. *IEEE Transactions on Intelligent Transportation Systems*, 18(8):2250–2259, 2017.

[162] G. Wang, J. Guo, M. Tang, J. F. de Queiroz Neto, C. Yau, A. Daghistani, M. Karimzadeh, W. G. Aref, and D. S. Ebert. STULL: unbiased online sampling for visual exploration of large spatiotemporal data. In *[Proceedings of IEEE VAST*, pages 72–83, 2020.

[163] H. Wang, Y. Lu, S. T. Shutters, M. Steptoe, F. Wang, S. Landis, and R. Maciejewski. A visual analytics framework for spatiotemporal trade network analysis. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):331–341, 2019.

[164] H. Wang, Y. Ni, L. Sun, Y. Chen, T. Xu, X. Chen, W. Su, and Z. Zhou. Hierarchical visualization of geographical areal data with spatial attribute association. *Visual Informatics*, 5(3):82–91, 2021.

[165] J. Wang, X. Cai, J. Su, Y. Liao, and Y. Wu. What makes a scatterplot hard to comprehend: Data size and pattern salience matter. *Journal of Visualization*, pages 1–17, 2021.

[166] Q. Wang, Z. Chen, Y. Wang, and H. Qu. A survey on ML4VIS: Applying machinelearning advances to data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 2021.

[167] Z. Wang, M. Lu, X. Yuan, J. Zhang, and H. van de Wetering. Visual traffic jam analysis based on trajectory data. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2159–2168, 2013.

[168] Z. Wang, T. Ye, M. Lu, X. Yuan, H. Qu, J. Yuan, and Q. Wu. Visual exploration of sparse traffic trajectory data. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1813–1822, 2014.

[169] D. Wei, C. Li, H. Shao, Z. Tan, Z. Lin, X. Dong, and X. Yuan. SensorAware: visual analysis of both static and mobile sensor information. *Journal of Visualization*, 24(3):597–613, 2021.

[170] D. Weng, R. Chen, Z. Deng, F. Wu, J. Chen, and Y. Wu. SRVis: Towards better spatial integration in ranking visualization. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):459–469, 2019.

[171] D. Weng, R. Chen, J. Zhang, J. Bao, Y. Zheng, and Y. Wu. Pareto-optimal transit route planning with multi-objective monte-carlo tree search. *IEEE Transactions on Intelligent Transportation Systems*, 22(2):1185–1195, 2021.

[172] D. Weng, C. Zheng, Z. Deng, M. Ma, J. Bao, Y. Zheng, M. Xu, and Y. Wu. Towards better bus networks: A visual analytics approach. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):817–827, 2021.

[173] D. Weng, H. Zhu, J. Bao, Y. Zheng, and Y. Wu. HomeFinder revisited: Finding ideal homes with reachability-centric multi-criteria decision making. In *Proceedings of ACM CHI*, page 247, 2018.

[174] M. Whitlock, K. Wu, and D. A. Szafir. Designing for mobile and immersive visual analytics in the field. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):503–513, 2020.

[175] A. Wu, W. Tong, T. Dwyer, B. Lee, P. Isenberg, and H. Qu. MobileVisFixer: Tailoring web visualizations for mobile phones leveraging an

explainable reinforcement learning framework. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):464–474, 2021.

[176] A. Wu, Y. Wang, X. Shu, D. Moritz, W. Cui, H. Zhang, D. Zhang, and H. Qu. AI4VIS: Survey on artificial intelligence approaches for data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 2021.

[177] F. Wu, M. Zhu, X. Zhao, Q. Wang, W. Chen, and R. Maciejewski. Visualizing the time-varying crowd mobility. In *Proceedings of ACM SIGGRAPH Asia Visualization in High Performance Computing*, pages 15:1–15:4, 2015.

[178] W. Wu, J. Xu, H. Zeng, Y. Zheng, H. Qu, B. Ni, M. Yuan, and L. M. Ni. TelCoVis: Visual exploration of co-occurrence in urban human mobility based on telco data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):935–944, 2016.

[179] W. Wu, Y. Zheng, H. Qu, W. Chen, M. E. Gröller, and L. M. Ni. BoundarySeer: Visual analysis of 2d boundary changes. In *Proceedings of IEEE VAST*, pages 143–152, 2014.

[180] Y. Wu, N. Cao, D. Gotz, Y. Tan, and D. A. Keim. A survey on visual analytics of social media data. *IEEE Transactions on Multimedia*, 18(11):2135–2148, 2016.

[181] Y. Wu, Z. Chen, G. Sun, X. Xie, N. Cao, S. Liu, and W. Cui. StreamExplorer: A multi-stage system for visually exploring events in social streams. *IEEE Transactions on Visualization and Computer Graphics*, 24(10):2758–2772, 2018.

[182] Y. Wu, J. Lan, X. Shu, C. Ji, K. Zhao, J. Wang, and H. Zhang. ittvis: Interactive visualization of table tennis data. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):709–718, 2018.

[183] Y. Wu, S. Liu, K. Yan, M. Liu, and F. Wu. OpinionFlow: Visual analysis of opinion diffusion on social media. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1763–1772, 2014.

[184] Y. Wu, D. Weng, Z. Deng, J. Bao, M. Xu, Z. Wang, Y. Zheng, Z. Ding, and W. Chen. Towards better detection and analysis of massive spatiotemporal co-occurrence patterns. *IEEE Transactions on Intelligent Transportation Systems*, 22(6):3387–3402, 2021.

[185] Y. Wu, X. Xie, J. Wang, D. Deng, H. Liang, H. Zhang, S. Cheng, and W. Chen. ForVizor: Visualizing spatio-temporal team formations in soccer. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):65–75, 2019.

[186] X. Xie, J. Wang, H. Liang, D. Deng, S. Cheng, H. Zhang, W. Chen, and Y. Wu. PassVizor: Toward better understanding of the dynamics of soccer passes. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1322–1331, 2021.

[187] P. Xu, Y. Wu, E. Wei, T. Peng, S. Liu, J. J. H. Zhu, and H. Qu. Visual analysis of topic competition on social media. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2012–2021, 2013.

[188] Y. Yang, T. Dwyer, S. Goodwin, and K. Marriott. Many-to-many geographically-embedded flow visualisation: An evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):411–420, 2017.

[189] S. Ye, Z. Chen, X. Chu, Y. Wang, S. Fu, L. Shen, K. Zhou, and Y. Wu. ShuttleSpace: Exploring and analyzing movement trajectory in immersive visualization. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):860–869, 2021.

[190] L. Ying, T. Tangl, Y. Luo, L. Shen, X. Xie, L. Yu, and Y. Wu. Glyphcreator: Towards example-based automatic generation of circular glyphs. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):400–410, 2022.

[191] L. Yu, W. Wu, X. Li, G. Li, W. S. Ng, S. Ng, Z. Huang, A. Arunan, and H. M. Watt. iVizTRANS: Interactive visual learning for home and work place detection from massive public transportation data. In *Proceedings of IEEE VAST*, pages 49–56, 2015.

[192] J. Yuan, C. Chen, W. Yang, M. Liu, J. Xia, and S. Liu. A survey of visual analytics techniques for machine learning. *Computional Visual Media*, 7(1):3–36, 2021.

[193] J. Yuan, S. Xiang, J. Xia, L. Yu, and S. Liu. Evaluation of sampling methods for scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1720–1730, 2021.

[194] J. Yuan, Y. Zheng, C. Zhang, X. Xie, and G. Sun. An interactive-voting based map matching algorithm. In *Proceedings of IEEE International Conference on Mobile Data Management*, pages 43–52, 2010.

[195] G. G. Zanabria, J. Silveira, J. Poco, A. Paiva, M. B. Nery, C. T. Silva, S. Adorno, and L. G. Nonato. CrimAnalyzer: Understanding crime patterns in são paulo. *IEEE Transactions on Visualization and Computer Graphics*, 27(4):2313–2328, 2021.

[196] W. Zeng, C. Fu, S. M. Arisona, A. Erath, and H. Qu. Visualizing mobility of public transportation system. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1833–1842, 2014.

[197] W. Zeng, C. Fu, S. M. Arisona, and H. Qu. Visualizing interchange patterns in massive movement data. *Computer Graphics Forum*, 32(3):271–280, 2013.

[198] W. Zeng, C. Fu, S. M. Arisona, S. Schubiger, R. Burkhard, and K. Ma. A visual analytics design for studying rhythm patterns from human daily movement data. *Visual Informatics*, 1(2):81–91, 2017.

[199] W. Zeng, C. Fu, S. M. Arisona, S. Schubiger, R. Burkhard, and K. Ma. Visualizing the relationship between human mobility and points of interest. *IEEE Transactions on Intelligent Transportation Systems*, 18(8):2271–2284, 2017.

[200] W. Zeng, P. C. Fu, S. M. Arisona, A. Erath, and H. Qu. Visualizing waypoints-constrained origin-destination patterns for massive transportation data. *Computer Graphics Forum*, 35(8):95–107, 2016.

[201] W. Zeng, C. Lin, J. Lin, J. Jiang, J. Xia, C. Turkay,

and W. Chen. Revisiting the modifiable areal unit problem in deep traffic prediction with visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):839–848, 2021.

[202] W. Zeng, Q. Shen, Y. Jiang, and A. C. Telea. Route-aware edge bundling for visualizing origin-destination trails in urban traffic. *Computer Graphics Forum*, 38(3):581–593, 2019.

[203] W. Zeng and Y. Ye. VitalVizor: A visual analytics system for studying urban vitality. *IEEE Computer Graphics and Applications*, 38(5):38–53, 2018.

[204] J. Zhang, E. Yanli, J. Ma, Y. Zhao, B. Xu, L. Sun, J. Chen, and X. Yuan. Visual analysis of public utility service problems in a metropolis. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1843–1852, 2014.

[205] Y. Zhang and R. Maciejewski. Quantifying the visual impact of classification boundaries in choropleth maps. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):371–380, 2017.

[206] Y. Zhao, H. Jiang, Q. Chen, Y. Qin, H. Xie, Y. Wu, S. Liu, Z. Zhou, J. Xia, and F. Zhou. Preserving minority structures in graph sampling. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1698–1708, 2021.

[207] F. Zheng, J. Wen, X. Zhang, Y. Chen, X. Zhang, Y. Liu, T. Xu, X. Chen, Y. Wang, W. Su, and Z. Zhou. Visual abstraction of large-scale geographical point data with credible spatial interpolation. *Journal of Visualization*, 24(6):1303–1317, 2021.

[208] Y. Zheng, L. Capra, O. Wolfson, and H. Yang. Urban computing: Concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology*, 5(3):38:1–38:55, 2014.

[209] Y. Zheng, F. Liu, and H. Hsieh. U-Air: When urban air quality inference meets big data. In *Proceedings of ACM SIGKDD*, pages 1436–1444, 2013.

[210] Y. Zheng, W. Wu, Y. Chen, H. Qu, and L. M. Ni. Visual analytics in urban computing: An overview. *IEEE Transactions on Big Data*, 2(3):276–296, 2016.

[211] Y. Zheng, W. Wu, H. Qu, C. Ma, and L. M. Ni. Visual analysis of bi-directional movement behavior. In *Proceedings of IEEE International Conference on Big Data*, pages 581–590, 2015.

[212] Y. Zheng, W. Wu, H. Zeng, N. Cao, H. Qu, M. Yuan, J. Zeng, and L. M. Ni. TelcoFlow: Visual exploration of collective behaviors based on telco data. In *Proceedings of IEEE International Conference on Big Data*, pages 843–852, 2016.

[213] Y. Zheng, X. Yi, M. Li, R. Li, Z. Shan, E. Chang, and T. Li. Forecasting fine-grained air quality based on big data. In *Proceedings of ACM SIGKDD*, pages 2267–2276, 2015.

[214] Z. Zhou, L. Meng, C. Tang, Y. Zhao, Z. Guo, M. Hu, and W. Chen. Visual abstraction of large scale geospatial origin-destination movement data. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):43–53, 2019.

[215] Z. Zhou, C. Shi, X. Shen, L. Cai, H. Wang, Y. Liu, Y. Zhao, and W. Chen. Context-aware sampling of large networks via graph representation learning. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1709–1719, 2021.

[216] Z. Zhou, X. Zhang, Z. Yang, Y. Chen, Y. Liu, J. Wen, B. Chen, Y. Zhao, and W. Chen. Visual abstraction of geographical point data with spatial autocorrelations. In *Proceedings of IEEE VAST*, pages 60–71, 2020.

[217] H. Zhu, M. Zhu, Y. Feng, D. Cai, Y. Hu, S. Wu, X. Wu, and W. Chen. Visualizing large-scale high-dimensional data via hierarchical embedding of knn graphs. *Visual Informatics*, 5(2):51–59, 2021.

[218] M. Zhu, W. Chen, J. Xia, Y. Ma, Y. Zhang, Y. Luo, Z. Huang, and L. Liu. Location2vec: A situation-aware representation for visual exploration of urban locations. *IEEE Transactions on Intelligent Transportation Systems*, 20(10):3981–3990, 2019.

Zikun Deng received his B.S. degree in Transportation Engineering from Sun Yat-Sen University in 2018. He is currently pursuing the doctoral degree with the State Key Lab of CAD & CG, Zhejiang University. His research interests mainly include spatiotemporal data mining, visualization, and urban visual analytics. For more information, please visit `https://zkdeng.org`.

Dr. Di Weng is a researcher at Microsoft Research Asia. He received his Ph.D. degree in Computer Science from the State Key Lab of CAD & CG, Zhejiang University in 2021 and his B.S. degree in Computer Science from Taishan Honors College, Shandong University in 2016. His research interests mainly include the data mining, visualization, and visual analytics of large-scale urban data. For more information, please visit `https://dweng.org`.

Shuhan Liu received her B.S.degree in Computer Science from Chu Kochen Honors College, Zhejiang University in 2021. She is currently pursuing the doctoral degree with the State Key Lab of CAD & CG, Zhejiang University. Her research interests mainly include spatiotemporal data mining, visualization, and industrial data visual analytics.

Yuan Tian is currently an undergraduate in the State Key Lab of CAD & CG, Zhejiang University. Her research interests mainly include visualization and visual analytics.

Dr. Mingliang Xu is a professor in the School of Information Engineering of Zhengzhou University, China, and currently is the director of CIISR (Center for Interdisciplinary Information Science Research) and the vice General Secretary of ACM SIGAI China. He received his Ph.D. degree in computer science and technology from the State Key Lab of CAD & CG at Zhejiang University, Hangzhou, China. His current research interests include computer graphics and artificial intelligence. He has authored more than 80 journal and conference papers in these areas, including ACM TOG, ACM TIST, IEEE TPAMI, IEEE TIP, IEEE TCYB, IEEE TCSVT, IEEE TAC, IEEE TCIAIG, ACM SIGGRAPH (Asia), ACM MM, IJCAI, etc.

Dr. Yingcai Wu is a Professor at the State Key Lab of CAD & CG, Zhejiang University. His main research interests are in information visualization and visual analytics, with focuses on urban computing, sports science, immersive visualization, and narrative visualization. He received his Ph.D. degree in Computer Science from the Hong Kong University of Science and Technology. Prior to his current position, Dr. Wu was a postdoctoral researcher in the University of California, Davis from 2010 to 2012, a researcher in Microsoft Research Asia from 2012 to 2015, and a ZJU100 Young Professor at Zhejiang University from 2015 to 2020. For more information, please visit `http://www.ycwu.org`.