

Frontier of Information Visualization and Visual Analytics in 2016

Min Lu · Siming Chen · Chufan Lai · Lijing Lin ·
Xiaoru Yuan ·

Received: date / Accepted: date

Abstract Visualization have evolved into a flourishing research field in recent 30 years. There are substantial visualization methodologies and applications published every year. Most of literature surveys focus on reviewing the state-of-art techniques in a certain direction in-depth. In this work, we conduct a cross-section survey by taking all the latest literatures as a whole, to obtain insights into the ecology of Information Visualization and Visual Analytics field in 2016. Center around 70 related publications in the IEEE VIS, we perform a mixed quantitative and qualitative analysis to report the current research progress, including statistical overview as well as detailed research topics.

Keywords Information Visualization · Visual Analytics · Survey

1 Introduction

Visualization is the discipline that transforms data into images and leverages human visual capability into the analysis, which helps to see the unseen. In 1987, the NSF (U.S. National Science Foundation) held a panel to discuss the potential of visualization as a new technology in scientific data analysis [McCormick, 1988]. Later, the visualization field has evolved into three branches, i.e., Scientific Visualization, Information Visualization and Visual Analytics. Scientific Visualization typically deals with three-dimensional phenomena, such as medical, biological data, etc. Information Visualization [Chen, 2005] concerns nonnumeric, nonspatial, and high-dimensional data, e.g., textural data, network, etc. Visual Analytics [Thomas, 2005] is the science which supports users to perform analytical reasoning via interactive visual interfaces.

In the last 30 years, numerous visualization methods and applications have been published in the field of visualization. In such a flourishing research field, many surveys summarize the variety of the visualization methods. Early in 1998, Geisler [Geisler, 1998] discusses information visualization applications and techniques according to data type, such as spatial, temporal, hierarchical, etc. Zudalova et al. [Zudilova-Seinstra et al., 2008] discuss a wide range of topics in interactive visualization, including data representation, novel user interface, etc. Liu et al. [Liu et al., 2014] perform a survey of the major research trends in the four categories of information visualization, namely empirical methodologies, user interactions, visualization frameworks, and applications. Other surveys focus on specific directions of visualization, such as graph visualization [Von Landesberger et al., 2011], visualization construction tools [Grammel et al., 2013], software visualization [Caserta and Zendra, 2011], network security events [Shiravi et al., 2012], etc.

As opposed to surveys of the state-of-art approaches in a certain direction in-depth, in this work, we aim to give a cross-section survey by taking all the latest literatures published in one year in the field of Information Visualization and Visual Analytics, in order to take the freshest snapshot of the academic ecology. We are more concerned with the research landscape of the whole Information Visualization and Visual Analytics field in 2016, such as *what challenges have been covered recently, how the research works are distributed, e.g., over different data types, etc., what new research trends there are*, etc.

Min Lu, Siming Chen, Chufan Lai, Lijing Lin, Xiaoru Yuan
Key Laboratory of Machine Perception (Ministry of Education), and School of EECS, Peking University
E-mail: xiaoru.yuan@pku.edu.cn

70 highly related publications are collected from the top visualization conference venues, IEEE VIS InfoVis (Information Visualization) and VAST (Visual Analytics Science and Technology) 2016. A set of descriptors are extracted from an array of well-known visualization frameworks [Shneiderman, 1996], challenges [Chen, 2005, Keim et al., 2008] to depict a research publication from multiple aspects. Then a peer review among all authors is performed to code literatures with the descriptors. With the structural literature collection, a mixed quantitative and qualitative analysis is conducted to gain insights into the current research progress, including:

- Provide a statistical overview of the research progress
- Summary the research topics of interest
- Identify the newly emerging research directions and discuss the future trends

In the rest of paper, we introduce the top challenges identified in the field in Section 2. Then we explain how the literature in 2016 is analyzed in Section 3. In Section 4, we give the statistical overview of the research in 2016. The research topics are introduced in detail in Section 5. In Section 6 we discuss the observations in research trends. Lastly, we conclude the report in Section 7.

2 Top Research Problems in Visualization

In the past years, researchers identified the top unsolved problems in Information Visualization and Visual Analytics. It is these challenges that drive the field of information visualization to continue evolving. Aligning the research progress with the proposed challenges provides a good awareness of current status of the field, namely what has been achieved and what has not. Conversely, research progress not aligned with proposed challenges indicates the new research interest.

In 2005, Chen [Chen, 2005] outlined the top 10 unsolved problems in information visualization, including those caused by technical barrier, those from the user-centric perspective, and those at the disciplinary level. Keim et al. [Keim et al., 2008] presented the top 10 most significant challenges in visual analytics from the perspectives of application and technology respectively. In 2014, Liu [Liu et al., 2014] also addressed five major technical challenges which hinder perfect information visualization.

We consolidated unsolved problems through a combination of similar ones. The resulting set contains 15 unsolved problems. (1) *Usability* asks for low-cost, ready-to-use information visualization systems and techniques. (2) *Assessment* of information visualization systems is essential for the science of visualization, including the understanding of elementary perceptual-cognitive tasks, measurement of visual quality, etc. (3) *Prior Knowledge* requires to adapt information visualization systems to the accumulated knowledge of their users. (4) *Education and Training* refers to the need to spread and communicate the knowledge of visualization inside as well as outside the field. (5) *Scalability* is one of the long-lasting challenges which requires continual performance as the scale increases. (6) *Aesthetic* issues ask for insightful and visually appealing information visualizations. (7) *Dynamics* needs to deal with the changes over time. (8) *Causality, Visual Inference and Predictions* require to understand the technology and comprehend the logic, reasoning and common sense. (9) *Semantics* requires to recognize complex coherences with human beings. (10) *Data Quality and Uncertainty* poses the challenge of analyzing data with quality problems or uncertainty. (11) *Data Provenance* asks for the understanding where data come from. (12) *Data Stream* requires to deal with the streaming data. (13) *Integration* requires integration with automatic analysis, database, statistics, etc. (14) *Knowledge Domain Visualization* is a synthesized challenge which requires conveying of information structures with knowledge. (15) *Synthesis of Problems* requires the solution to a series of heterogeneous problems.

3 Literature Analysis

To have an overview of the latest research progress of Information Visualization and Visual Analytics in 2016, we systematically reviewed 70 related literatures in IEEE VIS. Specifically, we covered all journal-track papers in IEEE VAST and InfoVis, as well as related papers in EuroVis, PacificVis partially. We also included some related papers in the array of workshops, panels, etc.

3.1 Analysis Workflow

Figure 1 illustrates the workflow of literature analysis. Firstly, a unified set of descriptors is proposed to depict the major features and challenges of a work. Referred to a collection of classical literatures about frameworks of visualization and visual analytics, the set includes five types of descriptors, which will be introduced in detail in Section 3.2. Secondly, all authors manually code the literature with descriptors in two rounds. In the first round, each literature is assigned to two authors and coded. In the second round, authors resolve the literature with conflict descriptors. Finally, a mixed quantitative and qualitative approach is performed based on the descriptors. Quantitatively, the insight is obtained from the statistics of descriptors (Section 4). Qualitatively, all authors derive the research topics of interest based on their opinions and experiences (Section 5).

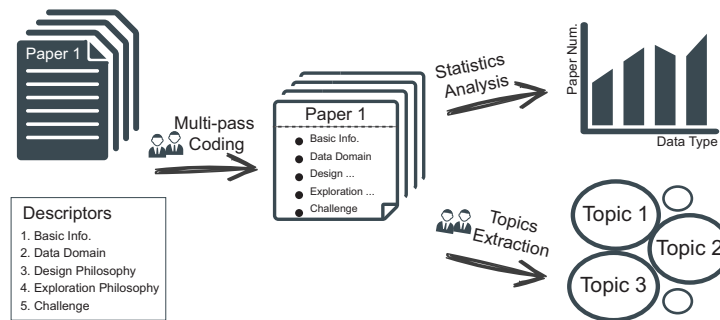


Fig. 1 Workflow of Literature Analysis: a unified set of descriptors is derived for information visualization and visual analytics; literature is multi-pass coded with descriptors by authors; insights are obtained from a mixed quantitative and qualitative analysis.

3.2 Literature Descriptors

Each paper is depicted with five descriptors covering from the basic information to detailed features. Specifically, they are defined as following.

- **Basic Information** includes the title, the major affiliation, venue where the paper is published, i.e., VAST, InfoVis, PacificVis, EuroVis
- **Data Domain** depicts what type of data the paper works on. Based on the category by Shneiderman [Shneiderman, 1996], we refine the data domain by following types, i.e., *Textual*, *Spatial*, *Temporal*, *Multi-dimensional*, *Hierarchical*, *Network*, *Hybrid*, and *General*, where if one method can be applied to the general data type, which means it has no specific requirement on the data type.
- **Visual Design Philosophy** describes what the visual interface is like. There are five types. *Stand-alone* design emphasizes one major diagram and others serve as auxiliaries. *Multi-view* design refers to the interface with multiple coordinate visual components. *Mixed-in* is to design one hybrid visualization based on two or more existing visualizations. *Add-on* design is to add visual enhancement while preserving the design of original one. *Physical* design takes the objects in reality as the medium of visualization.
- **Exploring Philosophy** depicts how the exploration works in the visual analytic system. There are seven types. *Overview-detail* as the most well-known exploration mantra [Shneiderman, 1996], is used to explore globally first and then perform detail analysis on demand. *Brush-link* provides the connecting exploration among multiple views. *Exploration-recommendation* takes the exploration (e.g., labelling) of users as input and responses users with feedback accordingly. *Query* implies those systems based on information retrieval. *Progressive* exploration updates the result iteratively. *Interaction enhanced exploration* improves exploring experience by interaction recording and recovering. *Immersive* exploration emphasizes on embedding users in the visualization environment.
- **Challenge** depicts what problem the work claims to tackle. The list of challenges is introduced in Section 2.

4 Statistics

Table 1 Publication Distribution over Challenges

Challenge	Literature
Usability	Data-oriented [Bach et al., 2017][Cordeil et al., 2017][Kwon et al., 2017][Liu et al., 2017c][Schulz et al., 2017][Yang et al., 2017a][Van et al., 2017][van der Zwan et al., 2016][Veras and Collins, 2017]
	General Visualization Authoring [Amini et al., 2017][Bigelow et al., 2017][Bryan et al., 2017][Kim et al., 2017b] [Satyanarayan et al., 2017][Yu and Silva, 2017]
	Others [Loorak et al., 2017][Saket et al., 2017][Sarvghad et al., 2017]
Assessment	Intrinsic Visual Measures [Beecham et al., 2017][Behrisch et al., 2017][Gramazio et al., 2017][Meulemans et al., 2017][Netzel et al., 2017][Padilla et al., 2017][Wu et al., 2017][Zhang and Maciejewski, 2017]
	Human Factor Measures [Cordeil et al., 2017][Crouser et al., 2017][Dasgupta et al., 2017][Dimara et al., 2017][Tam et al., 2017]
Scalability	Visual Scalability [Alsallakh and Ren, 2017][Veras and Collins, 2017]
	Interactive Scalability [Kim et al., 2017a][Turkay et al., 2017]
	Data Efficiency [Pahins et al., 2017][Xie et al., 2017][Wang et al., 2017]
Causality, Visual Inference and Predictions	Urban Related [Al-Dohuki et al., 2017][Liu et al., 2017a][Yang et al., 2017b]
	Text Related [Felix et al., 2017][Kim et al., 2017a][Berger et al., 2017][Shen et al., 2017]
	Social Media Related [Chen et al., 2016][Hu et al., 2017][Wang et al., 2016]
	Others [Fu et al., 2017][Li et al., 2017][Liu et al., 2017c][Xu et al., 2017]
Education and Training	Theory [Ceneda et al., 2017][Crouser et al., 2017][Dabek and Caban, 2017][Goodwin et al., 2017]
	Landscape Construction [Isenberg et al., 2017][Sacha et al., 2017]
	Education for the Public [He and Adar, 2017][Willett et al., 2017]
Integration	Machine Learning [Liu et al., 2017b][Rauber et al., 2017][Ren et al., 2017]

As the top visualization venue, IEEE VIS 2016 accepts 70 full journal-track papers on Information Visualization and Visual Analytics, 37 from InfoVis and 33 from VAST.

We count the publication distribution over different data types. General data ranks first, 21 publications in total. 16 of them are InfoVis papers about general visualization methodology, evaluation or theory model. Without the limitation to a certain type of data, this kind of general works are consolidating the foundation of the visualization field. On the other hand, visual analytics keeps being applied in various application domains. 11 out of 33 VAST papers model the data in the visual application as multi-dimensional data. Spatial, temporal and network also stimulate some research publications. Fusing heterogeneous data sets from multiple sources is a trend in the recent years. In 2016, 6 systems handle the hybrid data sets.

Figure 2 shows publication distributions over different design and exploration philosophies respectively. 42 out of 70 publications propose certain visualization or visual analytic methodologies. The other 28 publications concern more either about evaluation, theory or performance improvement. As Figure 2 shows, the majority of visual designs adopt the typical multiple or stand-alone view and provide exploration in overview-detail and brush-link manner. With the marriage of visualization and machine learning, the exploration and then recommendation is one of the hot exploration philosophies, which make the best of interactions (such as tagging, annotation, etc.) and automated algorithms (such as cluster, classification, etc.). There are 6 related works published in 2016. On the other hand, as the increase of data scale and task complexity, some papers turn visual exploration into the progressive one, to support users to explore the data and complete tasks iteratively. As found in Figure 2, there are some works concerning about new design philosophies in 2016, such as add-on, physical design, and exploration philosophies, such as interaction enhancement and immersive exploration. More detailed discussion will be included in Section 6.

Table 1 gives an overview of the publication distribution over different challenges. There are several main challenges where the research efforts go to in 2016. *Usability* issue is the top challenge tackled in 2016, with the largest number of publications. The majority of works improve the usability of visualization from two aspects, namely the efficient visualizations of specific data types such as graph (Section 5.1),

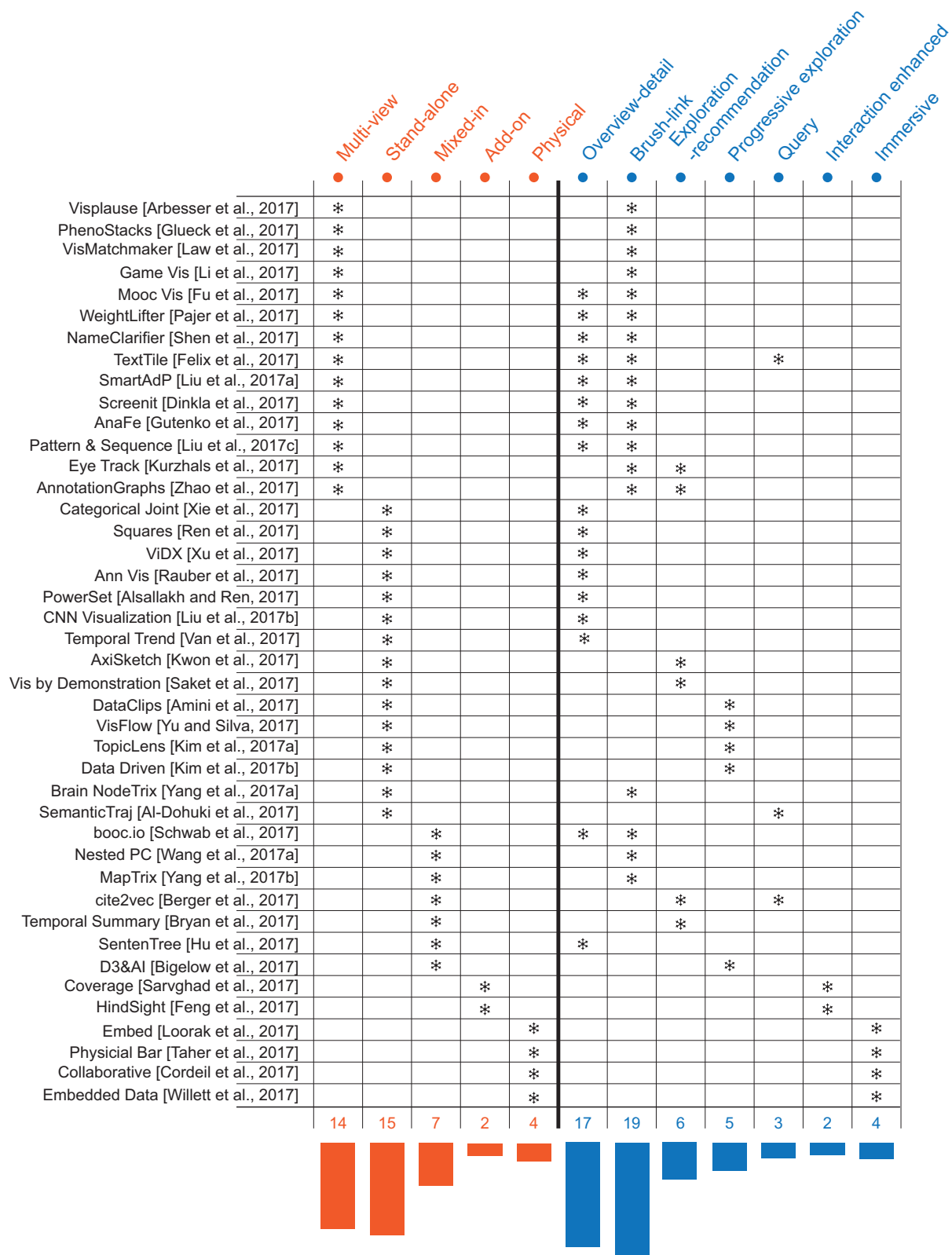


Fig. 2 The Distribution of 42 Publications over Different Design and Exploration Philosophies: the philosophies of multiple and stand-alone view dominate in visual design, accompanied with overview-design and brush-link exploration philosophies. There are some new philosophies emerging in 2016 (such as physical, add-on, etc.) which will be introduced in detail in Section 6.

high-dimensional data (Section 5.2) as well as visual authoring tools to provide ready-to-use visualization service for general data (Section 5.3). Assessing the quality of visualization and visual analytics is another challenge (Section 5.4), including *Intrinsic Visual Measures* and *Human Factor Measures*. As the data scales up, *Scalability* is a notable challenge, which attracts more or less research every year. The latest progress on *Scalability* is introduced in Section 5.5. New visual analytic systems are developed to support the *Causality*, *Visual Inference and Predictions* analysis in various application domains (Section 5.6). For better *Education and Train*, visualization theory is continuously developed to consolidate the foundation of the field (Section 5.7). Moreover, one new topic is the *Integration* of visualization with machine learning (Section 5.8), to make the best of human and computer algorithm. For other challenges, such as *Data Quality and Uncertainty*, and *Data Stream*, are included in some works but not dominant so that we will cover them along with the detailed explanation of the major challenges.

5 Research Topics

In this section, we summary the research progress of break-through topics in Table 1 in 2016.

5.1 Graph Visual Analytics

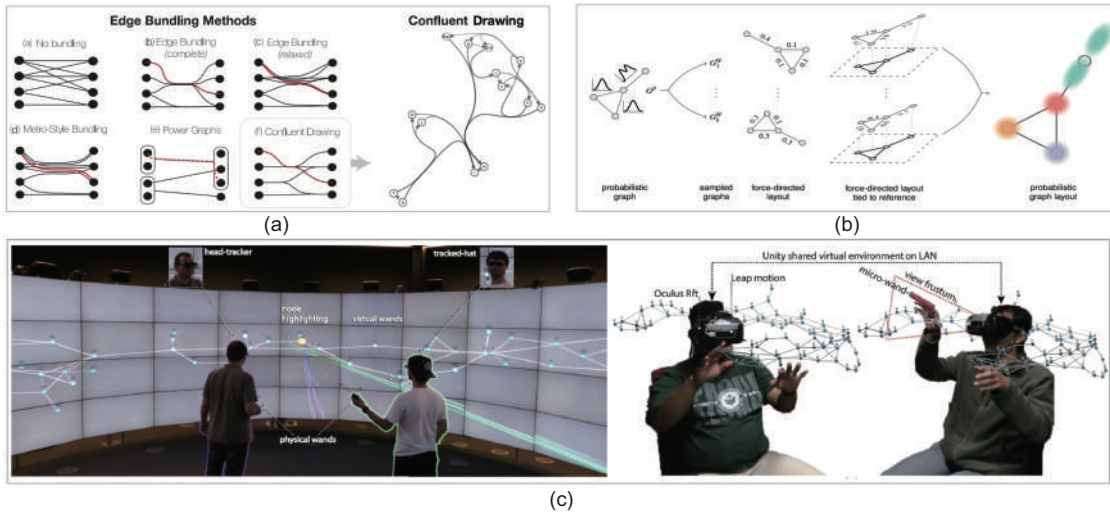


Fig. 3 Selected Graph Visual Analytics Work: (a) confluent edge drawings, towards unambiguous edge bundling [Bach et al., 2017]; (b) a novel uncertain network visualization technique [Schulz et al., 2017]; (c) a study, comparing the effectiveness of immersive collaborative analysis in CAVE-style and head-mounted display [Cordeil et al., 2017].

As one of the most important data models in visualization field, graph is widely used in various problems and applications. In 2016, graph papers can be categorized into three parts based on the problem they focus, namely graph layout, graph evaluation and graph application.

There are several works about graph layout. As Figure 3 (a) shows, Bach et al. [Bach et al., 2017] propose a novel technique, confluent edge drawings, for bundling edges in node-link diagrams based on network connectivity. Their method calculates graph layout with less ambiguous bundling edges, which is usually caused by traditional bundling techniques as they only focus on spatial proximity when bundling. CUBu [van der Zwan et al., 2016] presents a GPU-based framework that addresses several challenges, including speed, clutter, level-of-detail and parameter control, in visualizing large graph with edge bundling. Different from the certain graph data, Schulz et al. [Schulz et al., 2017] present a novel uncertain network visualization technique to visualize the distribution of possible realizations of a probabilistic graph (Figure 3 (b)). Vehlow et al. [Vehlow et al., 2016] deal with dynamic graph and design the matrix sequence to analyze the dynamic topology and hierarchical group structures.

Several works focus on evaluating the performance of large graph visualization in specific condition. With the same large graph, different sampling strategies would get very different resulting visualizations.

Wu et al. [Wu et al., 2017] investigate how sampling strategies influence the node-link graphs. Kwon et al. [Kwon et al., 2016] visualize large graph in an immersive environment and conduct a user study to compare their method to traditional 2D graph visualization. The study concludes that users would have better performance in an immersive graph visualization than traditional 2D one when facing difficult tasks and large graph. Cordeil et al. [Cordeil et al., 2017] evaluate the effectiveness of collaborative visualization in CAVE-style and head-mounted display (Figure 3 (c)). In the study, they find that participants using head-mounted displays are faster than those in CAVE2 condition.

There are several applications using graph to model data and analyze patterns. Annotation Graphs is a dynamic graph visualization of user-authored annotations [Zhao et al., 2017], which enables users to externalize their thoughts. Shi et al. [Yang et al., 2017a] adapt the NodeTrix representation to compare human brain networks.

As the network data becomes larger and more dynamic, graph visual techniques keep evolving. Techniques for small static graph are not effective enough due to increasing limitation of traditional display space and human cognition burden on dynamics. Analysing large dynamic graph keeps attracting attentions in research fields.

5.2 High-dimensional Data Visual Analytics

High-dimensional/Multivariate data visualization has long been an important and popular research direction. Numerous works concerning various topics are published in 2016.

5.2.1 Interaction with Projection

Dimension-reduced projection is an essential technique in high-dimensional data analysis. Different from the black box of dimension reduction (DR) in data mining, the visualization field enhances the interactions in the projection to provide intuitive DR controls and promote understanding. InterAxis [Kim et al., 2016] allows users to define a projection direction by assigning data items to its two ends. AxiSketcher [Kwon et al., 2017] extends the idea to non-linear projections, enabling users to create a non-linear axis by drafting a curve through the projected data. Zhou et al. [Zhou et al., 2016] extend the same idea in a different way. They allow users to draw a line on the projection to pick up the corresponding compound direction. More generally, Sacha et al. [Sacha et al., 2017] perform an in-depth survey in interactive DR, summarizing related literatures from the past decade.

5.2.2 Visual Analysis of Parameter/Solution Space

Optimization problems are very common in data analysis and decision making scenarios. But the solution space is usually too large to go through. Sampling in the space is a means to solve this issue. These samples are assessed by the optimization rule and then visualized to give the overview and details. The Grassmannian Atlas [Liu et al., 2016] is one such example, where the linear projection space is fully sampled and visualized as an atlas. Scatterplot assessments like the Scagnostics [Wilkinson and Wills, 2008] are used as rules to indicate projections with the most prominent features. As Figure 4(a) shows, the work by Xie et al. [Xie et al., 2017] shows a slightly different scenario where all alternatives are equally good. Users have to carefully inspect the solutions to make the final judgment. The WeightLifter [Pajer et al., 2017], on the other hand, deals with a similar problem in weight tuning (Figure 4(b)). The equally good weights are visualized as a tolerance area, allowing users to make decisions while being aware of parameter sensitivity.

Though the designs are different, the methodologies behind these works are very much alike. This is also applicable to many other situations, such as the ordering space of parallel coordinates or matrixes, the combination space of set-type data, etc. There has not been much research in this topic, but we can see a growing interest in it.

There are also other notable trends in high-dimensional data visualization. Time-varying multivariate data has gathered much interest in recent years. Researchers have tried various visual designs including parallel coordinates [Gruendl et al., 2016], dimension reduction [Bach et al., 2016][Jäckle et al., 2016], and even Self-Organizing Map [Bernard et al., 2016]. But there still lacks a good scheme to nicely integrate the two kinds of analysis. Progressive high-dimensional visualization is also a rising topic [Pezzotti et al., 2016, Turkay et al., 2017]. We'll simply leave this part to the subsequent section about scalability.

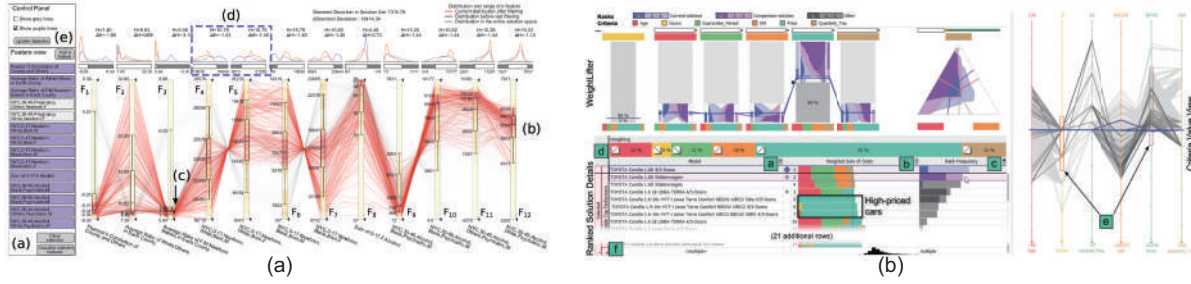


Fig. 4 Two Notable Works in Visual Analysis of Parameter/Solution Space in 2016: (a) a visual interface [Xie et al., 2017] supports users to iteratively narrow down the solution space in joint distribution reconstruction; (b) WeightLifter [Pajer et al., 2017] helps with the exploration of high-dimensional weight space in multi-criteria decision making.

5.3 Visualization Authoring

Recent years, there is a soaring boom of visualization authoring toolkits to facilitate visualization construction, from interactive online tool such as iVisDesigner [Ren et al., 2014] for non-programmers to powerful program libraries like D3.js [Bostock et al., 2011] for programmers. In 2016, there are more works in this direction - tools for constructing and editing visualization and tools for presenting visualization for story-telling.

5.3.1 Authoring for Visualization Construction and Editing

How to quickly and efficiently build a visualization system is always a challenge. There are two types of solutions, one for the programmer and the other for the general public. Vega-Lite [Satyanarayan et al., 2017] is a representative work in providing the visualization library, as one step further from D3.js. With Vega-Lite, Programmers can build up fully interactive visualizations. On the other hand, researchers provide What-You-See-Is-What-You-Get (WYSIWYG) visualization authoring tools. VisFlow [Yu and Silva, 2017] is a good example that users can flexibly construct different visualization views in the drag-and-drop manner. Kim et al. [Kim et al., 2017b] provide a data-driven interface enabling users to interactively construct infographics. As Figure 5(a) shows, Bigelow et al. [Bigelow et al., 2017] provide an add-on to Adobe Illustrator, which enables users to iteratively edit a visualization between D3 and Adobe Illustrator. Though it is a prototype, it sheds light on exploring the crossover in construction between the code generation and graphical editing.



Fig. 5 Selected Visualization Authoring Work: (a) combination of Adobe Illustrator with D3 [Bigelow et al., 2017]; (b) DataClips [Amini et al., 2017] for Story Telling Authoring.

5.3.2 Authoring for Story Telling

Presenting and communicating information is one of the important goals of visualization. In 2016, Bryan et al. [Bryan et al., 2017] proposed Temporal Summary Images to easily layout the annotations and tell the

stories of analytical findings. Besides static visualization results, motion graphics are also suitable to tell the stories with carefully designed animation. However, current methods to create motion graphics with visualization require lots of effort. To reduce the workload in good motion graphics construction, Amini et al. [Amini et al., 2017] summarize the existing visualization and motion graphics primitives and construct a data clip library. With the library, they provide DataClips to help users quickly construct motion graphics (Figure 5 (b)). Compared with the traditional motion graphics, it achieves good or even better quality but takes less effort.

We can see that visualization authoring is a trend of visualization research to reuse, construct visualizations easily and tell good stories of data. Such newly developed methods can help the public users and analysts make better use of visualization.

5.4 Assessment

There are a number of evaluation papers in 2016. One type focuses on the metrics or guideline to access the intrinsic visual quality of various options in the design space. The other type takes human factor in visual analytics into consideration and models them with models. The latter one is an increasing research topic emerging in the direction of assessment, which not only evaluates the visual design but also involves users as an essential part.

5.4.1 Intrinsic Visual Measures

In 2016, there is a particular interest in the visual quality assessment of geospatial visualization. Zhang et al. [Zhang and Maciejewski, 2017] quantify how critical the elements in choropleth map are in enabling the identification of spatial clusters. With the quantification, users are guided to pay special attention to the highly critical elements. Again spatial visualization, Meulemans et al. [Meulemans et al., 2017] construct a suite of metrics to quantify the optimization of various grid maps in the design space, including vertical alignment, false neighbors, etc. Similar to the gridded map, Padilla et al. [Padilla et al., 2017] study the impact of 2D scalar field in different binning conditions on analytics tasks. Beecham et al. [Beecham et al., 2017] perform crowdsourced experiments to study how graphical inference varies with spatial structure in geospatial visualization. Netzel et al. [Netzel et al., 2017] perform a controlled eye-tracking study to compare the performance of four different variants of map annotation in the visual searches. On the other hand, there are two works on evaluating the visual interface of graph visualization. Wu et al. [Wu et al., 2017] study the impact of five different sampling strategies on the eight visual factors in graph visual perception, including cluster quality, high degree nodes, edge linking, etc. Magnostics [Behrisch et al., 2017] covers image-based 30 feature descriptors to quantify the interest of matrix diagrams. Choosing color is one of the essential challenges in visualization. Gramazio et al. [Gramazio et al., 2017] propose Colorgorical which supports users to customize the color palette with three color-scoring functions to balance between aesthetic and discriminability.

5.4.2 Human Factor Measures

Other works are concerned with measurement of the human factors impacting the human-machine analysis loop. Dasgupta et al. [Dasgupta et al., 2017] study the relationships between domain experts trust and familiarity with the analysis medium. They evaluate the level-of-trust of 34 experts in a controlled user study to distill trust-augmented design. Dimara et al. [Dimara et al., 2017] examine whether one of well-studied human cognitive bias, i.e., attraction effect, appears in visualizations as the same presentation formats such as text, numerical tables, etc. Tam et al. [Tam et al., 2017] propose a common theoretic basis to model and compare visual analytics approach with machine-learning approach. Modeling human contributions to an algorithm as queries to a Human Oracle, Crouser et al. [Crouser et al., 2017] propose theoretical tool to reason the balance of human-machine collaboration. Besides the stand-alone analysis, Cordell et al. [Cordeil et al., 2017] perform the first study to evaluate the effectiveness of immersive collaboration between two displays, i.e., CAVE-style and Head-Mounted, in the context of network connectivity analysis.

5.5 Scalability

Scalability of information visualization and visual analytics methods is always a critical issue. We review the scaling methods in 2016 from three perspectives, namely visually scalable methods, interactively scalable methods, and data-efficient methods.

5.5.1 Visual Scalability

Visualizations with good scalability works well when the scale of underlying data increases. For example, as Figure 6(a) shows, PowerSet [Alsallakh and Ren, 2017] is a treemap-based visualization method to provide a scalable visual representation for intersection data. They argue that existing works can not easily show large amount of intersection. Each cell in the treemap represents one intersection among the set. From the top to bottom, the number of set in the intersection increase. With this approach, they can provide a scalable overview for the intersections among set. Other than set relation, Veras et al. [Veras and Collins, 2017] focus on the hierarchical data. They propose a scalable visualization method based on Minimum Description Length Principle, which can help users gain a balanced tree to increase the scalability. Figure 6(b) demonstrates three scalable display-tailored visualizations.

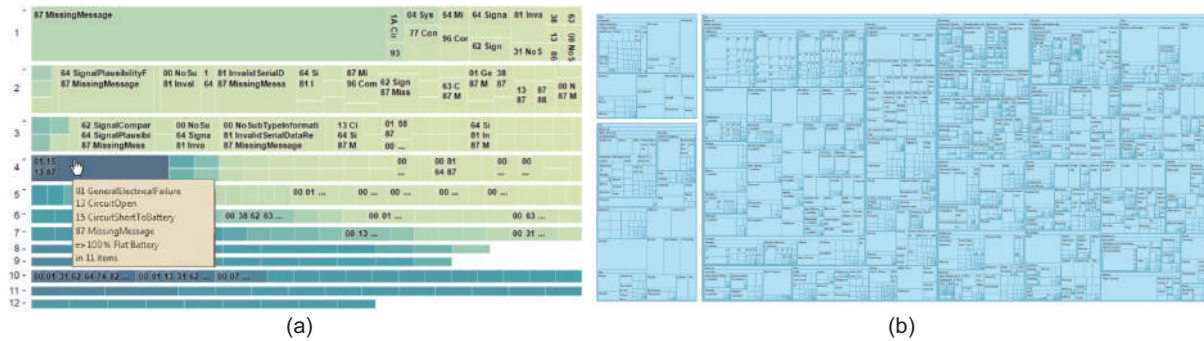


Fig. 6 Visual Scalability: (a) PowerSet [Alsallakh and Ren, 2017] presents each intersection as a cell in treemap; (b) three visualizations for large scale of hierarchical data in three displays with different sizes [Veras and Collins, 2017].

5.5.2 Interactive Scalability

Dealing with large data sets, researchers meet the scalability issues by designing interactions to achieve real-time changes. Turkay et al. [Turkay et al., 2017] design progressive visualizations for high-dimensional data, which iteratively update results while users can conduct interactions during the process. This method increases the usability of the visualization system. To design a scalable interaction, not only the real-time feedback is important but also the focus should be on preserving the semantic context. As Figure 7(b) shows, Kim et al. [Kim et al., 2017a] deploy the metaphor of lenses and design TopicLens which allows for progressive topic modeling on the fly. In their lens interaction, the detailed projection of the highlighting data are preserved in the context of overview, which maintains the scalability of such interactions. Scalable interactions often require the real time feedback with context preserved.

5.5.3 Data Efficiency

An efficient data organization can provide a solid foundation for designing scalable visualization and visual analytics systems. Pahins et al. propose Hashedcube [Pahins et al., 2017] with low memory requirements, low query latencies, and simple implementation as a scalable visualization system. Their methods can greatly reduce the query time and support the streaming data analysis. Besides the large amount, high dimensionality of data increases the complexity and becomes a scalability issue. Sampling is a way to improve the scalability. Wang et al. [Wang et al., 2017] propose Gaussian Cube, which pre-computes the best multivariate Gaussian for the respective data subsets in the preprocessing stage. It greatly enhances the efficiency of data exploration that it can handle one hundred million data points with 5-10 dimensions.

5.6 VAST Application Fields

In 2016, visual analytics is continuously applied in various application fields, to help domain experts perform causality analysis. We summarize the VAST application fields from the following perspectives, including urban data visual analytics, textual data visual analytics, social media visual analytics and the new application fields in VAST.

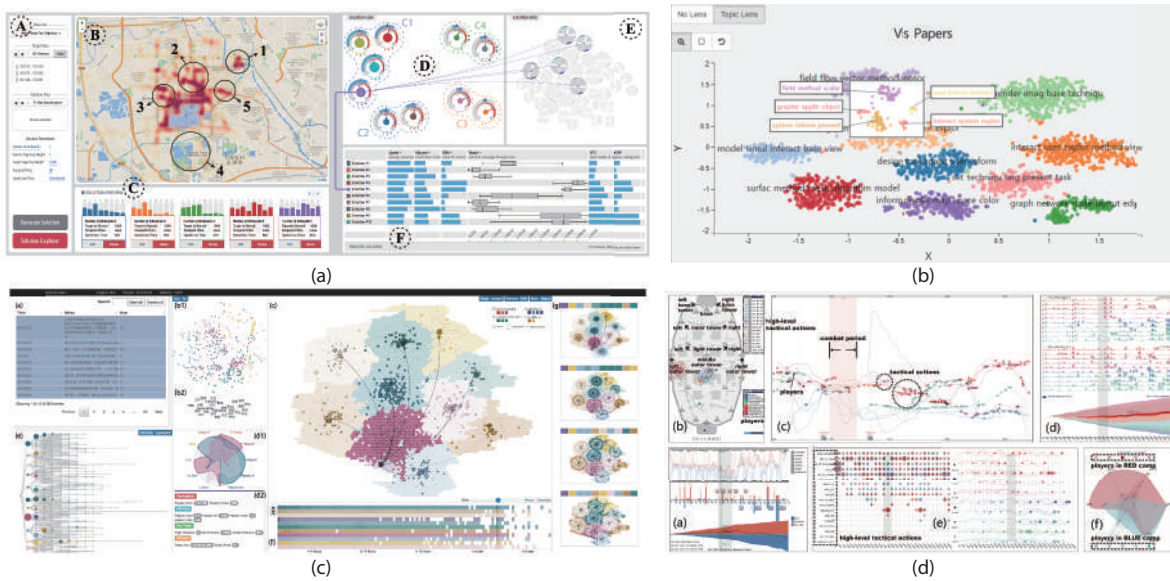


Fig. 7 Selected Visual Analytics Application Systems: (a) SmartADP, urban visual analytics to support billboard position selection [Liu et al., 2017a]; (b) TopicLens, textual data visual analytics with an interactive lens [Kim et al., 2017a]; (c) D-Map, a social media visual analytics system supporting ego-centric information diffusion [Chen et al., 2016]; (d) MOBA game visual analytics system [Li et al., 2017].

5.6.1 Urban Data Visual Analytics

In 2016, there are continual research works dealing with urban related data. Deriving semantics from spatial data, SemanticTraj [Al-Dohuki et al., 2017] leverages the processing methods from text mining and management to deal with the traffic data. They transform the GPS logs into a text database and build up a novel visual analytics system which supports text searching and semantic deriving functions. Cases from the experts confirm the capability in exploration of traffic patterns. With a similar goal but aiming to derive OD patterns, Yang et al. propose MapTrix [Yang et al., 2017b], a pair-wised OD matrix visualization to show uncluttered OD flows. More than investigating the traffic patterns itself, Liu et al. [Liu et al., 2017a] make use of the large scale of taxi data to help users make decisions on the billboard position selection. They provide a dual-view visual analytics system, including a spatial temporal view and a decision making view (Figure 7 (a)). By combining the multiple dimensions of the traffic data, they enable users to explore the solution space and find suitable solutions.

5.6.2 Textual Data Visual Analytics

In text analysis, researchers mainly focus on deriving the semantics from large corpus in many applications, such as news, social media content, research paper, etc. Felix et al. [Felix et al., 2017] summarize a data model for both structured and unstructured data. They provide a visualization tool to support filtering, splitting and summarizing in the data model. Besides the visual design, researchers also design novel interactions to explore large scale of documents. Florian et al. [Heimerl et al., 2016] propose an interactive lens called DocuCompass. They provide a Focus+Context method to explore the detailed keywords distribution with a lens. More than keywords, Kim et al. [Kim et al., 2017a] design a topic lens to support semantic exploration of the topic distribution (Figure 7 (b)). Their method can keep the context of the surrounding distribution of topics when zooming into the detail clusters. In this year, investigation of research

papers gains more focus. As opposed to the original approach to describe the paper with its own keywords, cite2vec [Berger et al., 2017] uses the keywords of how other papers cite it to represent the original paper. When investigating the research paper, name ambiguity is a common problem where researchers might share the same name. To conquer this problem, Shen et al. [Shen et al., 2017] propose a visual analytics system to solve name ambiguity problem with the hints in three keys, i.e., co-authorship, publication venues, and temporal information. With the specific designed hints and interactions, users can differentiate the authors with the same name.

5.6.3 Social Media Data Visual Analytics

Different from the textual data, social media emphasizes the people relationship, geo tags and more fruitful features. Social media is continuously a hot research focus in recent years. Hu et al. [Hu et al., 2017] propose SentenTree to visualize the correlated keywords in the social events. Besides the keyword and textual data research, researchers focus more on the information diffusion process this year. Chen et al. [Chen et al., 2016] propose D-Map to visualize the ego-centric information diffusion process with a map-based visual metaphor (Figure 7 (c)). They have found many interesting new social interaction patterns, such as dual-center social network, strong center diffusion patterns, etc. Different from the ego-centric perspective, Wang et al. [Wang et al., 2016] propose ideaFlow, to visualize how ideas (i.e. general information, topics, etc.) diffuse across multiple groups in different time periods. They also support the overview to detail exploration, showing the main themes of each idea. Besides information diffusion, Liu et al. [Liu et al., 2017a] also investigate what are publics' opinions towards different social brands based on social media data.

5.6.4 New VAST Application Field

In 2016, visual analytics has been applied in a wider range of new domain fields. Li et al. [Li et al., 2017] design visual analytics system to help game designers find key events and game parameters resulting in snowballing or comeback occurrences in MOBA game data (Figure 7 (d)). Fu et al. [Fu et al., 2017] present a system which allows for effectively discovering and understanding temporal patterns in MOOC forums. By collecting the data from users, the visual analytics systems help game develops and teachers better re-design game or courses. Xu et al. [Xu et al., 2017] extend the Mareys graph by introducing a time-aware outlier-preserving visual aggregation technique to support effective troubleshooting in manufacturing processes. It allows for both real-time and historical analysis. Liu et al. [Liu et al., 2017c] perform visual mining in the modern web clickstream data and propose the analytic pipeline consisting of three stages: pattern mining, pattern pruning and coordinated exploration between patterns and sequences.

5.7 Education and Training

The challenge of education and training refers to build up the theoretic foundation and language of visualization to facilitate the communication and sharing in the field. Early in 2004, the **Theory of Visualization** has been proposed as one of the top research problems in the visualization field [Johnson, 2004]. Lots of research effort has been put into building up the theoretic foundation of visualization and substantial amount of research works have been published, covering a wide range of taxonomies, principles, concept models and quantitative laws. It is estimated that hundreds of principles and guidelines are recommended by visualization books [Ware, 2012][Ward et al., 2010]. Meanwhile, the top visualization venues organize series of focused events on this topic in the past few years [Caroline et al., 2010][Min et al., 2016][Çağatay et al., 2011].

5.7.1 Theory

In 2016, there are considerable amount of publications on this topic. Different from classical discussion on the principles of visual design [Sedlmair et al., 2013][Strobelt et al., 2016] and taxonomies about visual analytics tasks [Jansen and Dragicevic, 2013][Munzner, 2009], some of them focus more on the theory of human-machine collaboration. On one hand, some propose conceptual frameworks to model the human factors in the loop of visual analytic. For example, Crouser et al. [Crouser et al., 2017] present the theoretical tool, Human Oracle Model, to evaluate the workload balance between human and machine. It models human contributions to an algorithm as queries to an Oracle with human-level intelligence. On the other hand, some focus on the concept model of how machines better interact with human. Ceneda et al. [Ceneda et al., 2017]

focus on the general model of guidances in the process of visual analytics, which is very essential for effective data analysis. They model guidelines with three main aspects, i.e., knowledge gap about *what does the user need to know to make progress*, input and output about *what is the basis for generating the guidance*, and the guidance degree about *how much guidance is provided*. Inspired by automaton, Dabek et al. [Dabek and Caban, 2017] propose a grammar-based approach to model and learn user interactions, which helps with the detection of common exploration patterns and recommendation of the end-users with useful data exploration in turn.

5.7.2 Landscape Construction

As the literatures accumulated, some researchers build up the landscape of visualization in finer granularity. One topic of interest in 2016 is the emergence of bottom-top literature survey method. Different from the top-bottom surveying approach which is based on the opinion and experiences of the authors, bottom-top takes every single publication as input data and performs analysis to draw the conceptual map of domain. One of the first steps in this direction is taken by Isenberg et al. [Isenberg et al., 2017]. They code each publication with a list of keywords by multi-pass analysis of visualization papers published in the IEEE Visualization conference series, and then derive major research topics and hot keywords in the whole field of visualization. Their work serves as a starting point to facilitate the construction of common vocabulary in the visualization field. Similarly, Sacha et al. [Sacha et al., 2017] perform a semi-automatic process on all the literatures related to dimension reduction in the visual analytics of high-dimensional data to derive the model of interactive dimension reduction.

5.7.3 Education for the Public

With the maturation of the visualization field, visualization has come into the public sight. Promoting the education of the public has been recognized in recent years. There are some related works published in 2016. As Figure 8 (a) shows, He et al. [He and Adar, 2017] introduce a card-driven workshop developed during their graduate infovis class, to demonstrate how InfoVis can be taught. Based on the collaborative-learning principle, VIZITCARDS stimulates positive collaborations and helps learners to make high-quality designs. Willett et al. [Willett et al., 2017] explore the conceptual framework which formalizes the notion of *embedded data representations*. They connect the existing research on visualization to the physical world via physically embedded data representations. Figure 8 (b) demonstrates the examples of the physical data representations, such as visualizing the wind flow by augmented reality in urban landscapes.



Fig. 8 Education for the Public: (a) VIZITCARDS [He and Adar, 2017] in Progress; (b) conceptual demonstration of embedded data representations [Willett et al., 2017].

5.8 Machine Learning

Machine learning is well known for its great analytic power in various data mining tasks, such as clustering, classification, prediction and so on. Machine learning is often the major means to promote analysis, while visualization helps people comprehend the results and tune the underlying model. One example is the newly published work by Dabek et al. [Dabek and Caban, 2017], where deterministic finite automaton (DFA) is adopted to build an interaction model and predict users subsequent moves.

In recent years, many academic activities have been held trying to bridge the gap between the two fields. One example is the MLVis in EuroVis 2016, a half-day tutorial that introduces machine learning methods

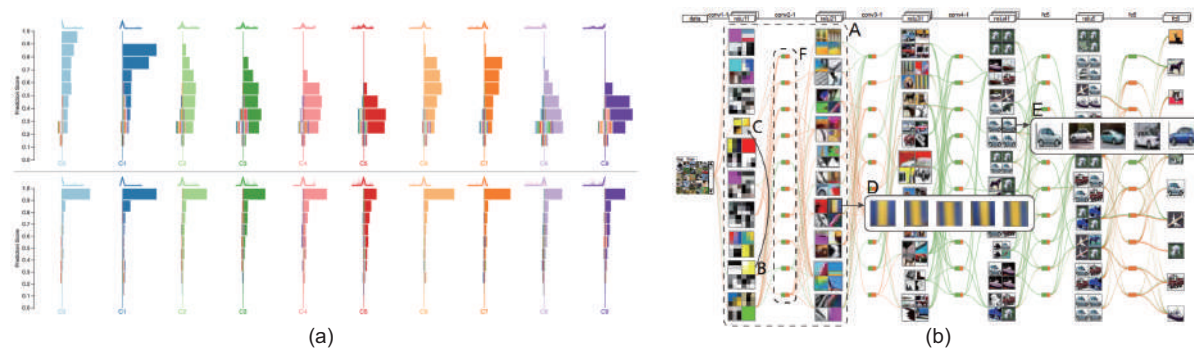


Fig. 9 Two Examples of Interplay between Machine Learning and Visualization: (a) Squares [Ren et al., 2017] visually evaluates a classifier; (b) convolutional neural network visualization helps to comprehend a machine learning process [Liu et al., 2017b].

to visualization researchers. The same event is also planned for EuroVis 2017, implying a growing trend to use machine learning to support interactive visual analysis.

More recently, this collaboration is proved reciprocal, as machine learning experts find them also gaining insights from the visualization field. Ren et al. [Ren et al., 2017] design a parallel-coordinates-style interface to help users evaluate the performance of a classifier (Figure 9 (a)). Liu et al. [Liu et al., 2017b] propose a graph-based visualization to show the overall structure and the learned knowledge in a convolutional neural network (Figure 9 (b)). Knowledge on each neuron is shown as featured image patches, allowing users to perceive and diagnose intuitively. Besides the connections, we can also analyze a neural network regarding the evolution of its activation vectors [Rauber et al., 2017].

Visual analysis of machine learning could be a promising new trend. It can deepen understanding of existing approaches, and boost the development of new techniques. However, a direct visualization may not be enough. It is important to introduce interactivity and seek for more in-depth topics. These are not easy tasks and require close cooperation with domain experts.

6 Discussion

During the literature analysis, we also review the visual design and exploration philosophy for each work to seek for potential future research of interest. In this section, we discuss our findings.

6.1 Design Philosophy

The majority of visual interfaces mainly consider either stand-alone view or linked multiple views. In 2016, three new thoughts emerge in visual design, namely add-on mechanism, mixed-in design and physical design space.

Add-on mechanism means to improve the visual capability by enhancing the existing approaches with auxiliaries, while maintains the familiarity of existing approaches to the max. Add-on design philosophy is not totally new in 2016. Back to 2012, Kong and Agrawala [Kong and Agrawala, 2012] had proposed five overlays to cover over charts to facilitate the chart reading tasks, such as reference lines, etc. However, more related designs are published in 2016. One example is HEDA proposed by Loran et al. [Loran et al., 2017]. HEDA extends the familiar multi-dimensional visualizations by embedding a tabular visualization. As Figure 10 (a) illustrates, two more variables are added to the line chart to make the x-axis more explorable. HindSight [Feng et al., 2017] augments existing visualizations with visual indicators of user exploration history. Figure 10 (b) shows two examples of HindSight, which encode the interaction history via opacity and color channel respectively. Similarly in the augmentation of interaction history, Sarvghad et al. [Sarvghad et al., 2017] add a scent widget into the interaction widgets to visualize the dimension coverage in the exploration space of high-dimensional data. As a new perspective for exploring the existing approaches, add-on design philosophy has great promise for the future.

Mixed-in philosophy is to design hybrid visualization by combining two or more visualizations. One of the earliest examples is the NodeTrix [Henry et al., 2007], a visualization combining the node-link diagram to show the global structure and matrices to visualize the local communities. In 2016, there are several

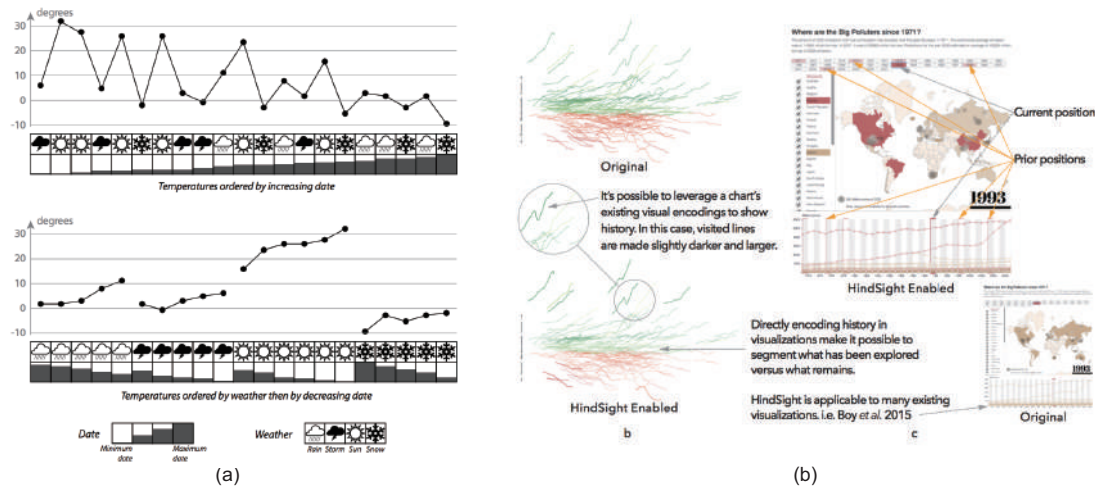


Fig. 10 Examples Adopting Add-on Machansim: (a) a line chart extended with HEDA [Loorak et al., 2017]; (b) HindSight [Feng et al., 2017] enhances visual systems with visual coding of exploring history.

mixed-in visualizations. For example, Yang et al. [Yang et al., 2017b] combine the matrix with the map view, providing MapTriX, to enable users to investigate the Origin-Destination Flow (Figure 11 (a)). Figure 11 (b) shows SentenTree [Hu et al., 2017], which adopts the node-link diagram where nodes are replaced with words and links indicate word co-occurrence within the same sentence. booc.io [Schwab et al., 2017] visualizes the concept map combining hierarchical circular layout of concepts and linear path of learning path (Figure 11 (c)). Moreover, the visualization authoring tool [Bigelow et al., 2017] integrates D3 into the well-known graphical editor Adobe Illustrator to improve the generative capability. Instead of designing from scratch, it would be an easy and efficient way to design with mixed-in design philosophy. However, the mixed-in visualization set high bar in the efficient combination to make the best of each part.

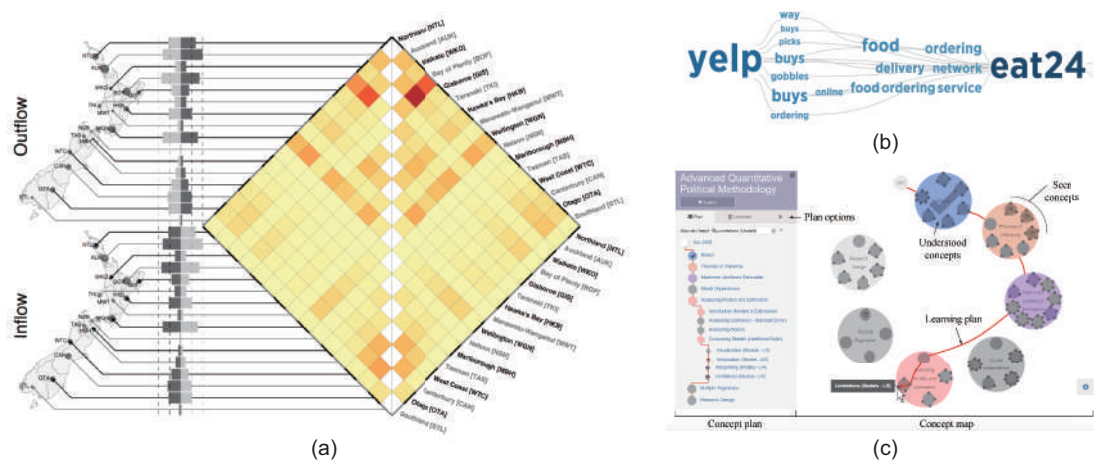


Fig. 11 Examples Adopting Mixed-in Machansim: (a) MapTriX [Yang et al., 2017b] combines OD Flow with map view; (b) SentenTree [Hu et al., 2017] displays a node-link diagram where nodes are words; (c) booc.io [Schwab et al., 2017] visualizes the conceptual map with hierarchical concept map and linear learning path.

In 2016, there is an interesting observation that visualizations in the virtual space (i.e., computer) are being externalized in the reality to enhance the visual communication, i.e., physical design. As Figure 12 (a) shows, Taher et al. [Taher et al., 2017] create the matrix with physical bars and allow users to create a bunch of true interactions. Stoppel et al. [Stoppel and Bruckner, 2017] present a printable tangible wheel chart, Vol²velle, which supports users to explore different parameter settings in volume rendering. Moreover, the general concept of embedding data representation in physical world is discussed by Willett et al. [Loorak et al., 2017]. Figure 12 (c) illustrates the embedding example of different communities informa-

tion visualization attached to the nametag of a group at a conference. In the near future, if the visualization can be naturally embedded in the physical space and enhance the understanding of the surroundings, the impact of visualization will be much larger for human lives. At the same time, it will induce a large design space, including design of physical world, perception tasks, and occlusion problems, etc.

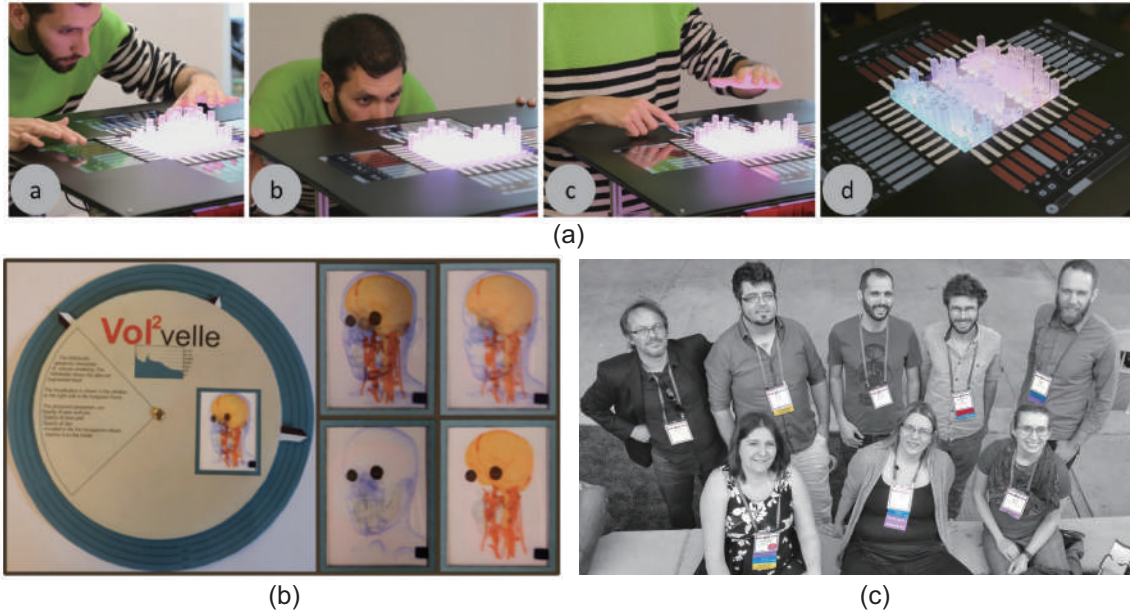


Fig. 12 Examples Adopting Physical Design: (a) visualization and exploration with physical bar chart [Taher et al., 2017]; (b) Vol²velle [Stoppel and Bruckner, 2017] provides physical exploration in different configurations in volume rendering; (c) an embedded data representation [Loorak et al., 2017].

6.2 Exploratory Philosophy

Interactive exploration is an important part in visualization design. In the recent decades, the overview to detail method from Shneiderman is well known and widely applied to many visualizations [Shneiderman, 1996]. Besides it, we follow two young explorations in the latest works, i.e., exploration-recommendation mechanism and progressive exploration.

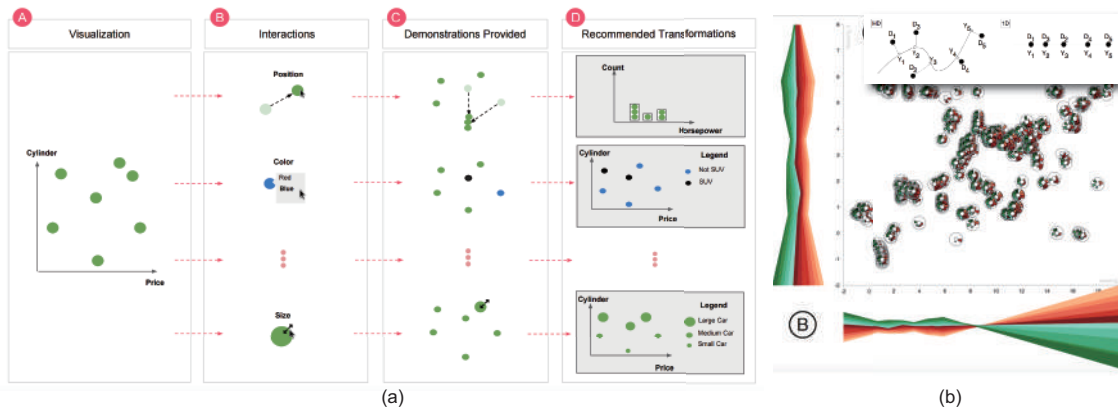


Fig. 13 Examples in Exploration-Recommendation Manner: (a) estimating the transformation in visualization by demonstration [Saket et al., 2017]; (b) interface of AxiSketcher [Kwon et al., 2017], which forms the axes of scatterplot based on users' sketching input.

The first one is the exploration-recommendation mechanism, where the system responds according to users' exploration. In 2016, researchers fulfill its meaning with many new ideas. One of typical exploration-recommendation systems is to suggest options after modeling and computation, such as recommendation of annotating positions' candidates in layout [Bryan et al., 2017], interaction suggestion by Dabek et al. [Dabek and Caban, 2017]. There are two special exploration-recommendation systems recommending according to users' demonstration. As Figure 13 (a) shows, Saket et al. [Saket et al., 2017] propose a visualization system which estimates the desirable transformation by analyzing the users' demonstration in example. The other is the AxiSketcher [Kwon et al., 2017], as Figure 13 (b) shows. In AxiSketcher, users sketch lines which reflects the desirable increasing pattern in the high-dimensional data. Then AxiSketcher creates scatterplot axes which reflect their intended model. Exploration-recommendation mechanism emphasizes on the human-in-the-loop in visual analytics, which would support users to have true discourse with the system.

The second one is the progressive exploration. Progressive computation improves the computation result iteratively, which continuously engage users in interpreting the responses from computer without any interruption. As the scale of data and computation increases, progressive computation is getting increasing interest. There are some related work in 2016. Turkay et al. [Turkay et al., 2017] propose three levels of operations to control the pace of progressive computation in high-dimensional data analysis, i.e., unit task completion, human-computer dialogue and visualization update. Their method computes the result within a limited time constraint and responses to the user. Another example is TopicLens [Kim et al., 2017a] which adopts the progressive computation of topic modelling. In the era of big data, progressive analytics potentially plays a key role in interactive analysis systems dealing with large scale of dataset.

7 Conclusion

In this work, we survey the 2016 research progress in Information Visualization and Visual Analytics. 70 recent publications are collected and coded with five descriptors, including basic information, challenges to tackle, etc. Aligning with the top challenges in the field, we identify the ongoing research topics and the future research interest. Classical topics keep being solved, such as graph visualization, multi-dimensional visual analytics. Visual analytics is applied in more and more domains, with advanced machine learning integrated. The communication and story-telling capability of visualizations is under exploration. Theories and models of visualization field are proposed to consolidate the foundation of the discipline. Moreover, researchers are active in exploring new design and exploration philosophies, such as visualization by demonstration. Immersive analytics and physical visualization emerge as a new research of interest in recent years.

Acknowledgements The authors wish to thank the anonymous reviewers for their valuable comments. This work is supported by NSFC No. 61672055. This work is also partially supported by NSFC Key Project No. 61232012 and the National Program on Key Basic Research Project (973 Program) No. 2015CB352503. This work is also funded by PKU-Qihoo Joint Data Visual Analytics Research Center.

References

- [Al-Dohuki et al., 2017] Al-Dohuki, S., Wu, Y., Kamw, F., Yang, J., Li, X., Zhao, Y., Ye, X., Chen, W., Ma, C., and Wang, F. (2017). SemanticTraj: A new approach to interacting with massive taxi trajectories. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):11–20.
- [Alsallakh and Ren, 2017] Alsallakh, B. and Ren, L. (2017). PowerSet : A comprehensive visualization of set intersections. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):361–370.
- [Amini et al., 2017] Amini, F., Riche, N. H., Lee, B., Monroy-Hernandez, A., and Irani, P. (2017). Authoring data-driven videos with dataclips. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):501–510.
- [Bach et al., 2017] Bach, B., Riche, N. H., Hurter, C., Marriott, K., and Dwyer, T. (2017). Towards unambiguous edge bundling: Investigating confluent drawings for network visualization. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):541–550.
- [Bach et al., 2016] Bach, B., Shi, C., Heulot, N., Madhyastha, T. M., Grabowski, T. J., and Dragicevic, P. (2016). Time Curves: Folding time to visualize patterns of temporal evolution in data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):559–568.
- [Beecham et al., 2017] Beecham, R., Dykes, J., Meulemans, W., Slingsby, A., Turkay, C., and Wood, J. (2017). Map LineUps: Effects of spatial structure on graphical inference. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):391–400.
- [Behrisch et al., 2017] Behrisch, M., Bach, B., Hund, M., Delz, M., Von Rden, L., Fekete, J.-D., and Schreck, T. (2017). Magnostics: Image-based search of interesting matrix views for guided network exploration. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):31–40.

- [Berger et al., 2017] Berger, M., McDonough, K., and Seversky, L. M. (2017). cite2vec: Citation-driven document exploration via word embeddings. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):691–700.
- [Bernard et al., 2016] Bernard, J., Sessler, D., Steiger, M., Spott, M., and Kohlhammer, J. (2016). Visual-interactive exploration of relations between time-oriented data and multivariate data. In *EuroVis Workshop on Visual Analytics (EuroVA)*, pages 49–53.
- [Bigelow et al., 2017] Bigelow, A., Drucker, S., Fisher, D., and Meyer, M. (2017). Iterating between tools to create and edit visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):481–490.
- [Bostock et al., 2011] Bostock, M., Ogievetsky, V., and Heer, J. (2011). D3: data-driven documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2301–2309.
- [Bryan et al., 2017] Bryan, C., Ma, K. L., and Woodring, J. (2017). Temporal summary images: An approach to narrative visualization via interactive annotation generation and placement. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):511–520.
- [Caroline et al., 2010] Caroline, Z., Peter, K., Robert, K., Jock, M., Bernice, R., and Ji, S. Y. (2010). Visualization theory: Putting the pieces together. Technical report, In IEEE VisWeek Panel.
- [Caserta and Zendra, 2011] Caserta, P. and Zendra, O. (2011). Visualization of the static aspects of software: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 17(7):913–933.
- [Çağatay et al., 2011] Çağatay, D., David, H. L., Jarke, J. v. W., and Colin, W. (2011). Theories of visualization – are there any? Technical report, IEEE VisWeek Panel.
- [Ceneda et al., 2017] Ceneda, D., Gschwandtner, T., May, T., Miksch, S., Schulz, H.-J., Streit, M., and Tominski, C. (2017). Characterizing guidance in visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):111–120.
- [Chen, 2005] Chen, C. (2005). Top 10 unsolved information visualization problems. *IEEE Computer Graphics and Applications*, 25(4):12–16.
- [Chen et al., 2016] Chen, S., Chen, S., Wang, Z., Liang, J., Yuan, X., Cao, N., and Wu, Y. (2016). D-Map: Visual analysis of ego-centric information diffusion patterns in social media. In *Proceedings of IEEE Visual Analytics Science and Technology (VAST)*, pages 41–50.
- [Cordeil et al., 2017] Cordeil, M., Dwyer, T., Klein, K., Laha, B., Marriott, K., and Thomas, B. H. (2017). Immersive collaborative analysis of network connectivity: Cave-style or head-mounted display? *IEEE Transactions on Visualization and Computer Graphics*, 23(1):441–450.
- [Crouser et al., 2017] Crouser, R. J., Franklin, L., Endert, A., and Cook, K. (2017). Toward theoretical techniques for measuring the use of human effort in visual analytic systems. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):121–130.
- [Dabek and Caban, 2017] Dabek, F. and Caban, J. J. (2017). A grammar-based approach for modeling user interactions and generating suggestions during the data exploration process. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):41–50.
- [Dasgupta et al., 2017] Dasgupta, A., Lee, J.-Y., Wilson, R., Lafrance, R. A., Cramer, N., Cook, K., and Payne, S. (2017). Familiarity vs Trust: A comparative study of domain scientists’ trust in visual analytics and conventional analysis methods. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):271–280.
- [Dimara et al., 2017] Dimara, E., Bezerianos, A., and Dragicevic, P. (2017). The attraction effect in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):471–480.
- [Felix et al., 2017] Felix, C., Pandey, A. V., and Bertini, E. (2017). TextTile: An interactive visualization tool for seamless exploratory analysis of structured data and unstructured text. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):161–170.
- [Feng et al., 2017] Feng, M., Deng, C., Peck, E. M., and Harrison, L. (2017). HindSight: Encouraging exploration through direct encoding of personal interaction history. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):351–360.
- [Fu et al., 2017] Fu, S., Zhao, J., Cui, W., and Qu, H. (2017). Visual analysis of mooc forums with iforum. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):201–210.
- [Geisler, 1998] Geisler, G. (1998). Making information more accessible: a survey of information visualization applications and techniques. URL: <https://www.cs.nmt.edu/viz/papers/GeislerMarrero.pdf> (Accessed Jan 2017).
- [Goodwin et al., 2017] Goodwin, S., Mears, C., Dwyer, T., de la Banda, M. G., Tack, G., and Wallace, M. (2017). What do constraint programming users want to see? Exploring the role of visualisation in profiling of models and search. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):281–290.
- [Gramazio et al., 2017] Gramazio, C. C., Laidlaw, D. H., and Schloss, K. B. (2017). Colorgorical: Creating discriminable and preferable color palettes for information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):521–530.
- [Gammel et al., 2013] Gammel, L., Bennett, C., Tory, M., and Storey, M.-A. (2013). A survey of visualization construction user interfaces. *EuroVis-Short Papers*, pages 19–23.
- [Gruendl et al., 2016] Gruendl, H., Riehmann, P., Pausch, Y., and Froehlich, B. (2016). Time-series plots integrated in parallel-coordinates displays. *Computer Graphics Forum*, 35(3):321–330.
- [He and Adar, 2017] He, S. and Adar, E. (2017). VizItCards : A card-based toolkit for infovis design education. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):561–570.
- [Heimerl et al., 2016] Heimerl, F., John, M., Han, Q., Koch, S., and Ertl, T. (2016). DocuCompass: Effective exploration of document landscapes. In *Proceedings of IEEE Visual Analytics Science and Technology (VAST)*, pages 11–20.
- [Henry et al., 2007] Henry, N., Fekete, J.-D., and McGuffin, M. J. (2007). NodeTriX: a hybrid visualization of social networks. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1302–1309.
- [Hu et al., 2017] Hu, M., Wongsuphasawat, K., and Stasko, J. (2017). Visualizing social media content with sententree. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):621–630.
- [Isenberg et al., 2017] Isenberg, P., Isenberg, T., Sedlmair, M., Chen, J., and Möller, T. (2017). Visualization as seen through its research paper keywords. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):771–780.
- [Jäckle et al., 2016] Jäckle, D., Fischer, F., Schreck, T., and Keim, D. A. (2016). Temporal MDS plots for analysis of multivariate data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):141–150.

- [Jansen and Dragicevic, 2013] Jansen, Y. and Dragicevic, P. (2013). An interaction model for visualizations beyond the desktop. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2396–2405.
- [Johnson, 2004] Johnson, C. (2004). Top scientific visualization research problems. *IEEE Computer Graphics and Applications*, 24(4):13–17.
- [Keim et al., 2008] Keim, D. A., Mansmann, F., Schneidewind, J., Thomas, J., and Ziegler, H. (2008). Visual analytics: Scope and challenges. In *Visual Data Mining*, pages 76–90.
- [Kim et al., 2016] Kim, H., Choo, J., Park, H., and Endert, A. (2016). InterAxis: Steering scatterplot axes via observation-level interaction. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):131–140.
- [Kim et al., 2017a] Kim, M., Kang, K., Park, D., Choo, J., and Elmqvist, N. (2017a). TopicLens: Efficient multi-level visual topic exploration of large-scale document collections. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):151–160.
- [Kim et al., 2017b] Kim, N. W., Schweickart, E., Liu, Z., Dontcheva, M., Li, W., Popovic, J., and Pfister, H. (2017b). Data-Driven Guides: Supporting expressive design for information graphics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):491–500.
- [Kong and Agrawala, 2012] Kong, N. and Agrawala, M. (2012). Graphical overlays: Using layered elements to aid chart reading. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2631–2638.
- [Kwon et al., 2017] Kwon, B. C., Kim, H., Wall, E., Choo, J., Park, H., and Endert, A. (2017). AxiSketcher: Interactive nonlinear axis mapping of visualizations through user drawings. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):221–230.
- [Kwon et al., 2016] Kwon, O.-H., Muelder, C., Lee, K., and Ma, K.-L. (2016). A study of layout, rendering, and interaction methods for immersive graph visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(7):1802–1815.
- [Li et al., 2017] Li, Q., Xu, P., Chan, Y. Y., Wang, Y., Wang, Z., Qu, H., and Ma, X. (2017). A visual analytics approach for understanding reasons behind snowballing and comeback in moba games. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):211–220.
- [Liu et al., 2017a] Liu, D., Weng, D., Li, Y., Bao, J., Zheng, Y., Qu, H., and Wu, Y. (2017a). SmartAdP: Visual analytics of large-scale taxi trajectories for selecting billboard locations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):1–10.
- [Liu et al., 2017b] Liu, M., Shi, J., Li, Z., Li, C., Zhu, J., and Liu, S. (2017b). Towards better analysis of deep convolutional neural networks. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):91–100.
- [Liu et al., 2016] Liu, S., Bremer, P., Jayaraman, J. J., Wang, B., Summa, B., and Pascucci, V. (2016). The Grassmannian Atlas: A general framework for exploring linear projections of high-dimensional data. *Computer Graphics Forum*, 35(3):1–10.
- [Liu et al., 2014] Liu, S., Cui, W., Wu, Y., and Liu, M. (2014). A survey on information visualization: recent advances and challenges. *The Visual Computer*, 30(12):1373–1393.
- [Liu et al., 2017c] Liu, Z., Wang, Y., Dontcheva, M., Hoffman, M., Walker, S., and Wilson, A. (2017c). Patterns and sequences: Interactive exploration of clickstreams to understand common visitor paths. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):321–330.
- [Loorak et al., 2017] Loorak, M. H., Perin, C., Collins, C., and Carpendale, S. (2017). Exploring the possibilities of embedding heterogeneous data attributes in familiar visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):581–590.
- [McCormick, 1988] McCormick, B. H. (1988). Visualization in scientific computing. *SIGBIO Newsletter*, 10(1):15–21.
- [Meulemans et al., 2017] Meulemans, W., Dykes, J., Slingsby, A., Turkay, C., and Wood, J. (2017). Small multiples with gaps. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):381–390.
- [Min et al., 2016] Min, C., Georges, G., Chris, R. J., Jessie, K., Tamara, M., and Melanie, T. (2016). Pathways for theoretical advances in visualization. Technical report, IEEE VIS Panel.
- [Munzner, 2009] Munzner, T. (2009). A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):921–928.
- [Netzel et al., 2017] Netzel, R., Hlawatsch, M., Burch, M., Balakrishnan, S., Schmauder, H., and Weiskopf, D. (2017). An evaluation of visual search support in maps. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):421–430.
- [Padilla et al., 2017] Padilla, L., Quinan, P. S., Meyer, M., and Creem-Regehr, S. H. (2017). Evaluating the impact of binning 2D scalar fields. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):431–440.
- [Pahins et al., 2017] Pahins, C. A. L., Stephens, S. A., Scheidegger, C., and Comba, J. L. D. (2017). Hashedcubes: Simple, low memory, real-time visual exploration of big data. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):671–680.
- [Pajer et al., 2017] Pajer, S., Streit, M., Torsney-Weir, T., Spechtenhauser, F., Miller, T., and Piringer, H. (2017). WeightLifter: Visual weight space exploration for multi-criteria decision making. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):611–620.
- [Pezzotti et al., 2016] Pezzotti, N., Hilt, T., Lelieveldt, B. P. F., Eisemann, E., and Vilanova, A. (2016). Hierarchical Stochastic Neighbor Embedding. *Computer Graphics Forum*, 35(3):021–030.
- [Rauber et al., 2017] Rauber, P. E., Fadel, S. G., Falcao, A. X., and Telea, A. C. (2017). Visualizing the hidden activity of artificial neural networks. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):101–110.
- [Ren et al., 2017] Ren, D., Amershi, S., Lee, B., Suh, J., and Williams, J. D. (2017). Squares: Supporting interactive performance analysis for multiclass classifiers. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):61–70.
- [Ren et al., 2014] Ren, D., Hollerer, T., and Yuan, X. (2014). iVisDesigner: Expressive interactive design of information visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2092–2101.
- [Sacha et al., 2017] Sacha, D., Zhang, L., Sedlmair, M., Lee, J. A., Peltonen, J., Weiskopf, D., North, S. C., and Keim, D. A. (2017). Visual interaction with dimensionality reduction: A structured literature analysis. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):241–250.
- [Saket et al., 2017] Saket, B., Kim, H., Brown, E. T., and Endert, A. (2017). Visualization by demonstration: An interaction paradigm for visual data exploration. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):331–340.

- [Sarvghad et al., 2017] Sarvghad, A., Tory, M., and Mahyar, N. (2017). Visualizing dimension coverage to support exploratory analysis. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):21–30.
- [Satyanarayan et al., 2017] Satyanarayan, A., Moritz, D., Wongsuphasawat, K., and Heer, J. (2017). Vega-Lite: A grammar of interactive graphics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):341–350.
- [Schulz et al., 2017] Schulz, C., Nocaj, A., Goertler, J., Deussen, O., Brandes, U., and Weiskopf, D. (2017). Probabilistic graph layout for uncertain network visualization. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):531–540.
- [Schwab et al., 2017] Schwab, M., Strobelt, H., Tompkin, J., Fredericks, C., Huff, C., Higgins, D., Strezhnev, A., Komisarich, M., King, G., and Pfister, H. (2017). booc.io: An education system with hierarchical concept maps and dynamic non-linear learning plans. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):571–580.
- [Sedlmair et al., 2013] Sedlmair, M., Munzner, T., and Tory, M. (2013). Empirical guidance on scatterplot and dimension reduction technique choices. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2634–2643.
- [Shen et al., 2017] Shen, Q., Wu, T., Yang, H., Wu, Y., Qu, H., and Cui, W. (2017). NameClarifier: A visual analytics system for author name disambiguation. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):141–150.
- [Shiravi et al., 2012] Shiravi, H., Shiravi, A., and Ghorbani, A. A. (2012). A survey of visualization systems for network security. *IEEE Transactions on visualization and computer graphics*, 18(8):1313–1329.
- [Shneiderman, 1996] Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages*, pages 336–343.
- [Stoppel and Bruckner, 2017] Stoppel, S. and Bruckner, S. (2017). Vol²velle: Printable interactive volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):861–870.
- [Strobelt et al., 2016] Strobelt, H., Oelke, D., Kwon, B. C., Schreck, T., and Pfister, H. (2016). Guidelines for effective usage of text highlighting techniques. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):489–498.
- [Taher et al., 2017] Taher, F., Jansen, Y., Woodruff, J., Hardy, J., Hornbk, K., and Alexander, J. (2017). Investigating the use of a dynamic physical bar chart for data exploration and presentation. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):451–460.
- [Tam et al., 2017] Tam, G. K., Kothari, V., and Chen, M. (2017). An analysis of machine-and human-analytics in classification. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):71–80.
- [Thomas, 2005] Thomas, J. J. (2005). *Illuminating the path the research and development agenda for visual analytics*. IEEE Computer Society.
- [Turkay et al., 2017] Turkay, C., Kaya, E., Balcisoy, S., and Hauser, H. (2017). Designing progressive and interactive analytics processes for high-dimensional data analysis. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):131–140.
- [Van et al., 2017] Van, G. A., Staals, F., Lffler, M., Dykes, J., and Speckmann, B. (2017). Multi-granular trend detection for time-series analysis. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):661–670.
- [van der Zwan et al., 2016] van der Zwan, M., Codreanu, V., and Telea, A. (2016). CUBu: universal real-time bundling for large graphs. *IEEE Transactions on Visualization and Computer Graphics*, 22(12):2550–2563.
- [Vehlow et al., 2016] Vehlow, C., Beck, F., and Weiskopf, D. (2016). Visualizing dynamic hierarchies in graph sequences. *IEEE Transactions on Visualization and Computer Graphics*, 22(10):2343–2357.
- [Veras and Collins, 2017] Veras, R. and Collins, C. (2017). Optimizing hierarchical visualizations with the minimum description length principle. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):631–640.
- [Von Landesberger et al., 2011] Von Landesberger, T., Kuijper, A., Schreck, T., Kohlhammer, J., van Wijk, J. J., Fekete, J.-D., and Fellner, D. W. (2011). Visual analysis of large graphs: state-of-the-art and future research challenges. In *Computer Graphics Forum*, volume 30, pages 1719–1749.
- [Wang et al., 2016] Wang, X., Liu, S., Chen, Y., Peng, T.-Q., Su, J., Yang, J., and Guo, B. (2016). How ideas flow across multiple social groups. In *Proceedings of IEEE Visual Analytics Science and Technology (VAST)*, pages 51–60.
- [Wang et al., 2017] Wang, Z., Ferreira, N., Wei, Y., Bhaskar, A. S., and Scheidegger, C. (2017). Gaussian Cubes: Real-time modeling for visual exploration of large multidimensional datasets. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):681–690.
- [Ward et al., 2010] Ward, M. O., Grinstein, G., and Keim, D. (2010). *Interactive data visualization: foundations, techniques, and applications*. CRC Press.
- [Ware, 2012] Ware, C. (2012). *Information visualization: perception for design*. Elsevier.
- [Wilkinson and Wills, 2008] Wilkinson, L. and Wills, G. (2008). Scagnostics distributions. *Journal of Computational and Graphical Statistics*, 17(2):473–491.
- [Willett et al., 2017] Willett, W., Jansen, Y., and Dragicevic, P. (2017). Embedded data representations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):461–470.
- [Wu et al., 2017] Wu, Y., Cao, N., Archambault, D., Shen, Q., Qu, H., and Cui, W. (2017). Evaluation of graph sampling: A visualization perspective. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):401–410.
- [Xie et al., 2017] Xie, C., Zhong, W., and Mueller, K. (2017). A visual analytics approach for categorical joint distribution reconstruction from marginal projections. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):51–60.
- [Xu et al., 2017] Xu, P., Mei, H., Ren, L., and Chen, W. (2017). ViDX: Visual diagnostics of assembly line performance in smart factories. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):291–300.
- [Yang et al., 2017a] Yang, X., Shi, L., Daianu, M., Tong, H., Liu, Q., and Thompson, P. (2017a). Blockwise human brain network visual comparison using nodetrix representation. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):181–190.
- [Yang et al., 2017b] Yang, Y., Dwyer, T., Goodwin, S., and Marriott, K. (2017b). Many-to-many geographically-embedded flow visualisation: An evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):411–420.
- [Yu and Silva, 2017] Yu, B. and Silva, C. T. (2017). VisFlow - Web-based visualization framework for tabular data with a subset flow model. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):251–260.
- [Zhang and Maciejewski, 2017] Zhang, Y. and Maciejewski, R. (2017). Quantifying the visual impact of classification boundaries in choropleth maps. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):371–380.
- [Zhao et al., 2017] Zhao, J., Glueck, M., Breslav, S., Chevalier, F., and Khan, A. (2017). Annotation Graphs: A graph-based visualization for meta-analysis of data based on user-authored annotations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):261–270.

- [Zhou et al., 2016] Zhou, F., Li, J., Huang, W., Zhao, Y., Yuan, X., Liang, X., and Shi, Y. (2016). Dimension reconstruction for visual exploration of subspace clusters in high-dimensional data. In *Proceedings of IEEE Pacific Visualization Symposium (PacificVis 2016)*, pages 128–135.
- [Zudilova-Seinstra et al., 2008] Zudilova-Seinstra, E., Adriaansen, T., and Liere, R. v. (2008). *Trends in Interactive Visualization: State-of-the-Art Survey*. Springer Publishing Company, Incorporated, 1 edition.