ACM TIST Special Issue on Visual Analytics

With the advance of technology, profound data are being produced and acquired daily in a wide range of settings spanning from urban computing, online social network, health care, and e-commerce to scientific simulation. Insights from data can help people carry out comprehensive assessments and make informed decisions. However, data are not readily intelligible. Recent advances in data mining shed light on data-intrinsic patterns and improve the analysis and understanding of large, complex datasets by applying automatic algorithms. Although huge successes have been achieved, automatic approaches have their limitations in regard to detecting expected patterns with fuzzy features and discovering unexpected patterns from large amounts of dynamic, heterogeneous, ambiguous, and even conflicting data from multiple sources. Even when patterns can be captured automatically, the intuitive explanation and validation of the patterns can pose another major challenge. Moreover, underspecified complex tasks with a high-dimensional space of input variables and parameters cannot be simply managed without the inclusion of human expertise and knowledge.

Visual analytics, which aim to support interactive analytical reasoning, has been introduced to address these challenges by combining computational analysis algorithms (to incorporate machine intelligence) and interactive data visualization (to incorporate human intelligence) via interactive visual interfaces. Visual analytics empowers analysts to use highly efficient computational algorithms for filtering out irrelevant data, simplifying complex data, narrowing down the search space, and detecting expected patterns interactively via visual interfaces. Furthermore, visual analytics can show analysts the data and analysis results in an intuitive visual context such that they can readily interpret and validate the results through their high-bandwidth visual channel. The integral combination of human knowledge, intuition, and expertise with powerful computational algorithms is the inherent strength of visual analytics and allows analysts to steer and supervise the process of analyzing and exploring large complex data effectively. Thus, users can formulate and test hypotheses and perform and explain reasoning in a quick, iterative manner. Visual analytics coupling machine and human intelligences via interactive visual interfaces have proven their effectiveness in various applications, such as in the fields of urban computing, bioinformatics, banking and finance, and health care.

This special issue introduces the readership to the most recent innovations in visual analytics, thereby casting a spotlight on a promising interdisciplinary research direction. The call for papers received a positive response, with a total of 40 article submissions. Through a rigid peer review, 10 high-quality articles were accepted for this special issue. A brief introduction is provided to each of the articles as follows.

Visual Analytics for Urban Computing

The first three articles introduce techniques for visually analyzing complex urban data. In "RelationLines: Visual Reasoning of Egocentric Relations from Heterogeneous Urban Data," Chen

2157-6904/2018/12-ART1
https://doi.org/10.1145/3277019

et al. introduce a novel visual analysis system for detecting egocentric mobility and communication patterns of a focal individual based on an urban dataset. The proposed system applies a progressive detect-and-filter technique that supports the visual reasoning and inspection of an ego-network. The system first identifies the focal individual by querying on a set of constraints, such as geographical locations, taxi trajectories, and geotags in microblog posts. Then, a temporal history of the focal individual is produced indicating his or her trajectory. Social ties between the focal individual and others are thus formed on the basis of their phone records or geographical co-occurrences. The resulting ego-network is refined gradually and progressively via an interactive procedure, which helps with the detection of the hidden mobility and communication patterns. The proposed technique has been evaluated on the basis of the real-world urban dataset.

The article entitled "Traffic Simulation and Visual Verification in Smog" describes a smog simulation model to estimate the current stadia. This method further inputs into a new smog full velocity difference model (SMOG-FVDM) to obtain a realistic simulation of traffic in smog weather conditions. This work also introduces a visual analysis system to allow users to search for a set of proper parameters of SMOG-FVDM effectively from a large parameter space. Specifically, exploring the parameter space, examining and comparing parameter correlations, and iteratively selecting proper parameters can help users improve the modeling results.

In "Visual Analytics of Heterogeneous Data using Hypergraph Learning," Cong et al. use a hypergraph that captures high-order relations to help fuse the heterogeneous data from various sources into a unified structure. A novel hypergraph learning algorithm extended from existing techniques, which predicts the missing labels in the fused datasets interactively guided by the users' assessment in accordance with their domain knowledge, is introduced in the paper. A set of visualizations is also provided in the system to help users construct the hypergraph and tune the learning parameters interactively. The power of the system is evaluated via two real-world cases in the field of urban computing.

Visual Analytics for Machine Learning

The next three articles present novel visual analysis approaches for analyzing and exploring machine learning algorithms. In "BayesPiles: Visualization Support for Bayesian Network Structure Learning," Vogogias et al. present a visualization approach that supports the analysis of large collections of Bayesian networks. They describe their approach, which combines the results of computational methods with interactive visual analysis for helping users develop a consensus network structure from the set of input networks. Incorporating interactive visualizations allows users to drive a consensus building process and produce improved results through iterative refinement. In addition, the article demonstrates the utility of their approach with applications to computational biology.

In "DeepTracker Visualizing the Training Process of Convolutional Neural Networks," Liu et al. introduce DeepTracker to enable users to understand the training process of convolutional neural networks (CNNs). DeepTracker, with its clever cube style overview visualization, supports drill-down operations into the vast information contained in training logs of CNNs and helps in understanding the correlations of training iterations, neuron weights, and validation images on different levels of visual abstraction. The approach constitutes another important step on the way to understand and solve problems during the creation of CNN models and aims to reduce the high computational training cost of CNNs.

The article entitled "Learning Facial Expressions with 3D Mesh Convolutional Neural Network" presents an approach to facial expression classification, which uses visualization to review and refine newly trained CNNs. Jin et al. demonstrate the use of visualization to understand the underlying network's behavior. In particular, high- and low-activation nodes can be identified and interpreted. Moreover, the authors demonstrate that low-activation nodes can be removed from the network to improve the overall classification performance.

Visual Analytics Applications

The last four articles introduce several interesting visual analytics applications, including examining resume data, finding similar peers, analyzing visual–verbal consistency, and exploring information diffusion on social media. In "ResumeVis: A Visual Analytics System to Discover Semantic Information in Semi-structured Resume Data," Zhang and Wang report that text mining and natural language processing techniques have seen enormous progress in recent years and are an integral part of numerous applications managing large numbers of text documents. Combining these methods with interactive visual analysis has a great potential in fields where automatic processing alone is insufficient for informed decision making. The ResumeVis approach shows a fruitful combination for a specific type of semi-structured texts, namely, curricula vitae. The approach indicates that powerful analyses are possible with this visual analytics system, which supports various human resource-related tasks.

In "Visual Interfaces for Recommendation Systems: Finding Similar and Dissimilar Peers," Du et al. present visual interfaces for recommendation applications through which users can readily find similar peers by retrieving, comparing, and refining peer groups interactively. Their study introduces a new hierarchical visualization, namely, LikeMeDonuts, that displays a visual summary of individual peers and allows users to compare the seed record with the displayed peers visually. This approach enables users to change its layout interactively to meet individual requirements and can also suggest a proper layout to achieve improved scalability and reduce visual clutter. Their work also introduces Ranking Glyph and History Heatmap visualizations, in addition to LikeMeDonuts, to assist users in finding and comparing similar peers.

The article entitled "CapVis: Towards Better Understanding of Visual-Verbal Saliency Consistency" introduces an interesting study on how people's attention on an image matched their descriptions of the same image. CapVis, a visual analysis system, is developed to support the study and reveals the relationship between eye fixations and image captions. A set of features has been proposed to capture the visual–verbal saliency consistency and is visualized in the system for further understanding the image content. Liang et al. also investigate different methods on how a human and computational model may look and describe an image. The investigation results are used to evaluate and calibrate the captioning model.

In "D-Map+: Interactive Visual Analysis and Exploration of Ego-centric and Event-centric Information Diffusion Patterns in Social Media," Chen et al. introduce a model to characterize the general information diffusion on social media used for ego- and event-centric analyses. A compact and occlusion-free hexagon map visualization is proposed to represent the information diffusion on social media visually. The approach first extracts communities from social media users and projects the communities into different regions on the map by using the hexagonal tessellation.

On the basis of the model and visualization, the work presents a visual analytics system called D-Map+ to support ego- and event-centric analyses and visualization. The system empowers users

to examine influential users, explore their behaviors, and relate their behaviors to the evolution of important events on social media visually.

Yingcai Wu
State Key Lab of CAD&CG
Zhejiang University
ycwu@zju.edu.cn

Nan Cao
Tongji University
nan.cao@gmail.com

Steffen Koch
University of Stuttgart
Steffen.Koch@vis.uni-stuttgart.de

David Gotz
University of North Carolina at Chapel Hill
gotz@unc.edu