Introduction to the Special Issue on Interactive Computational Visual Analytics

Remco Chang, Tufts University
David Ebert, Purdue University
Daniel Keim, University of Konstanz

This editorial introduction describes the aims and scope of the special issue on Interactive Computational Visual Analytics of the ACM Transactions on Interactive Intelligent Systems. It explains why visual analytics is crucial to the growing needs surrounding data analysis, and it shows how the four articles selected for this issue reflect this theme.

Categories and Subject Descriptors: C.2.2 [**Computer-Communication Networks**]: Network Protocols

General Terms: Design, Algorithms, Performance

Additional Key Words and Phrases: Wireless sensor networks, media access control, multi-channel, radio interference, time synchronization

ACM Reference Format:
Remco Chang, David Ebert, and Daniel Keim, 2014. Introduction to Special Issue on Interactive Computational Visual Analytics. *ACM Trans. Embedd. Comput. Syst.* 9, 4, Article 39 (March 2014), 2 pages.
DOI: http://dx.doi.org/10.1145/0000000.0000000

1. INTRODUCTION

Welcome to the Special Issue on Interactive Computational Visual Analytics.

Visual analytics is defined as "the science of analytical reasoning facilitated by visual interactive interfaces." Since its inception in 2006, the field has grown to encompass a wide array of topics relating to the theory, design, and development of interactive visual interfaces for the purposes of data exploration, data analysis, sense making, and decision making.

While the scope of visual analytics is broad, one principle that has emerged over the years is the need for visual analytics systems to leverage computational methods in statistics, data mining, knowledge discovery, and machine learning for large-scale data analysis. In these systems, the human operator works alongside the computational processes in an integrated fashion - the computer can sift through large amounts of data and identify the relevant information, while the human interactively explores the reduced data space to discover trends and patterns and make informed decisions. The two components operate in coordination, allowing for a continuous and cooperative analytical loop.

This special issue includes papers that address how computational methods can be integrated into interactive visualization systems from a variety of perspectives. In-

Author's addresses: R. Chang, Department of Computer Science, Tufts University; David Ebert, School of Electrical and Computer Engineering, Purdue University; Daniel Keim, Department of Computer and Information Science, University of Konstanz.

DOI: http://dx.doi.org/10.1145/0000000.0000000

tegrating statistics with visualizations, the papers by Chan et al. and Jean-Bernard demonstrate how interactivity allows a human operator to more effectively explore and utilize automated statistical methods. Using a similar approach, Riveiro examines the use of anomaly detection methods alongside visualizations to aide the analysis of real-time maritime traffic data. Finally, Chen et al. presents a method to capture sequences of analysis steps into a model to preserve the provenance of a visual data analysis task. Together, these four papers represent how computational methods can be augmented by interactive visualizations of data to better assist analysts in exploring and analyzing large and complex data.

1.1. Regression Cube: A Technique for Multidimensional Visual Exploration and Interactive Pattern Finding

The *Regression Cube* technique from Chan et al. demonstrates that even classic analysis techniques can be augmented with interactive features that extend the depth of analysis. In this case, the common operation of multivariate regression analysis using scatterplots is extended to include interactive sensitivity analysis and an iterative technique called *regression hierarchies*. These methods allow analysts to move beyond global regression analysis to explore local trends in the data, as well as trend sensitivity.

1.2. Interactive Statistics with Illmo

A widespread problem in the scientific community is making the correct choice of statistical model and tests in empirical research. Martens Jean-Bernard explains that despite the proliferation of statistical interfaces (e.g. SPSS, SAS) and flexible software (e.g. R), many researchers continue to use statistical tests without verifying that the underlying assumptions of these tests are met. The system presented by Jean-Bernard, called *Illmo*, is an interactive statistical interface that allows researchers to both test their model's assumptions and explore alternative models to find the appropriate choice for their dataset. By identifying the critical points-of-entry for humans to explore parameters in statistical models, Illmo embodies core ideas of both Visual Analytics and Intelligent Systems.

1.3. Evaluation of normal model visualization for anomaly detection in maritime traffic

The analysis of real-time data remains a challenging problem in many data-driven areas including power-grid maintenance, computer security, surveillance, and in this case, maritime traffic. In this paper, Maria Riveiro describes an approach to aiding real-time analysis by visualizing what data looks like in a normal case, which is hypothesized to increase the ability of analysts to accurately detect anomalous behavior. Riveiro compares the effectiveness of the normal model visualization to other common approaches to maritime traffic analysis, and adapts existing visualization operation models to focus on the unique challenges of anomaly detection.

1.4. Employing a Parametric Model for Analytic Provenance

When exploring a dataset using tools like those presented in this special issue, analysts leave behind not only a trail of interactions, but also a trail of insights, decisions, and discoveries. This trail of thinking, or "analytic provenance", is often lost since few systems have means to effectively record analyst interactions. The approach presented here by Chen et al. includes a new language and model to capture the analytic process for later exploration and re-use. Their approach is evaluated using a number of openly available datasets, and shows how a user can refer to their own analysis history to inform future analysis directions. As the need for scalable analytic provenance contin-

ues to grow, work in this direction will serve as a basis for capturing analytics in many scientific domains.