

Introduction to the Special Issue on Interactive Computational Visual Analytics

REMCO CHANG, Tufts University
DAVID S. EBERT, Purdue University
DANIEL KEIM, University of Konstanz

This editorial introduction describes the aims and scope of *ACM Transactions on Interactive Intelligent Systems*'s special issue on interactive computational visual analytics. 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: H.1.2 [**User/Machine Systems**]: Human Information Processing; H.5.2 [**User Interfaces**]: Graphical User Interfaces (GUI)

General Terms: Visual Analytics, Interaction

Additional Key Words and Phrases: Information visualization, exploratory data analysis, statistics, computational methods

1. INTRODUCTION

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.

Authors' addresses: Remco Chang, Tufts University, Department of Computer Science, 161 College Ave, Medford, MA 02155; David S. Ebert, Purdue University, School of Electrical and Computer Engineering, 465 Northwestern Ave, West Lafayette, IN 47907; Daniel Keim, University of Konstanz, Department of Computer and Information Science, Box 78, 78457 Konstanz, Germany.

This special issue includes articles that address how computational methods can be integrated into interactive visualization systems from a variety of perspectives. Integrating statistics with visualizations, the articles by Chan et al. and by Martens demonstrate how interactivity allows a human operator to explore and utilize automated statistical methods more effectively. Using a similar approach, Riveiro examines the use of anomaly detection methods alongside visualizations to aid the analysis of real-time maritime traffic data. Finally, Chen et al. present a method to capture sequences of analysis steps into a model to preserve the provenance of a visual data analysis task. Together, these four articles represent how computational methods can be augmented by interactive visualizations of data to better assist analysts in exploring and analyzing large and complex data.

2. THE FOUR ARTICLES

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

The *regression cube* technique of 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 scatter plots 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.

2.2. Interactive Statistics with Illmo

A widespread problem in the scientific community is making the correct choice of statistical model and tests in empirical research. Jean-Bernard Martens 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 Martens, called *Illmo*, is an interactive statistical interface that allows researchers both to test their model's assumptions and to 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.

2.3. 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 article, 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 detect anomalous behavior accurately. 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.

2.4. 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, because 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 reuse. Their approach is evaluated using a number of openly available datasets, and it shows how a user can refer to their own analysis history to inform future analysis directions. As the need for scalable analytic provenance continues to grow, work in this direction will serve as a basis for capturing analytics in many scientific domains.