Information Visualization WS 2013/14

11  Visual Analytics
11 Visual Analytics

11.1 Definitions and Motivation

- Lot of research and papers in this emerging field:
  - “Visual Analytics: Scope and Challenges” of Keim et al.
  - “Illuminating the path” of Thomas and Cook
  - ...

![Book Cover: Illuminating the Path](image)
11.1 Definitions and Motivation

– Nowadays we are confronted with an information overload.
– Reasons for this: Increasing computer power and storage capacities.
– Data is produced at incredible rates and our ability to collect and store the data is increasing at a faster rate than our ability to analyze it.
11.1 Definitions and Motivation

- The analysis of these massive, typically messy and inconsistent, volumes of data is crucial in many application domains.

- Why?
  - For decision makers, analysts or emergency response teams it is an essential task to rapidly extract relevant information from the flood of data.
11.1 Definitions and Motivation

- Big Data
  - A phenomenon defined by the rapid acceleration in the expanding volume of high velocity, complex and diverse types of data

- Big data is often defined along three dimensions:
  - Volume – size
  - Velocity – rate of input, output, change
  - Variety – different types, sources, variables
Information Visualization WS 2013/14

11 Visual Analytics

• 11.1 Definitions and Motivation

  – What do we need?
    • Advanced techniques and technologies to enable the capture, storage, distribution, management and analysis of data
    • Enable effective, efficient analysis, decision making, planning, and action
11.1 Definitions and Motivation

- Big Data Challenges:
  - Heterogeneity
  - Unprecedented volume and/or complexity
  - Multivariate, Multidimensional
  - Streaming
  - Qualitative, Quantitative, Categorical, Temporal, ...
  - Distributed sources: sensors, text, images, maps,...
  - Inconsistent and uncertain
  - Different scales and cross scales
  - ...
11.1 Definitions and Motivation

Today, a selected number of software tools is employed to help analysts to organize their information, generate overviews and explore the information space in order to extract potentially useful information.

Most of these data analysis systems still rely on interaction metaphors developed more than a decade ago and it is questionable whether they are able to meet the demands of the ever-increasing mass of information.
11.1 Definitions and Motivation

– In fact, huge investments in time and money are often lost, because we still lack the possibilities to properly interact with the databases.

– Visual analytics aims at bridging this gap by employing more intelligent means in the analysis process.
11 Visual Analytics

11.1 Definitions and Motivation

- The basic idea of visual analytics is to visually represent the information, allowing the human to directly interact with the information, to gain insight, to draw conclusions, and to ultimately make better decisions.

- The visual representation of the information reduces complex cognitive work needed to perform certain tasks. People may use visual analytics tools and techniques to synthesize information and derive insight from massive, dynamic, and often conflicting data by providing timely, defensible, and understandable assessments.
11.1 Definitions and Motivation

- Decision-makers should be enabled to examine this massive, multi-dimensional, multi-source, time-varying information stream to make effective decisions in time-critical situations.
11.1 Definitions and Motivation

- For informed decisions, it is indispensable to include humans in the data analysis process to combine flexibility, creativity, and background knowledge with the enormous storage capacity and the computational power of today's computers.

- The specific advantage of visual analytics is that decision makers may focus their full cognitive and perceptual capabilities on the analytical process, while allowing them to apply advanced computational capabilities to augment the discovery process.
11  Visual Analytics

• 11.1 Definitions and Motivation
  – Visual Analytics is defined as the science of analytical reasoning supported by interactive visual interfaces.

[Thomas and Cook, 2005]
11.2 Scope of Visual Analytics

- Can be described as:
  
  “...the science of analytical reasoning facilitated by interactive visual interfaces.” [Thomas and Cook, 2005]

- To be more precise, visual analytics is an iterative process that involves information gathering, data preprocessing, knowledge representation, interaction and decision making.
11.2 Scope of Visual Analytics

- The ultimate goal is to gain insight in the problem at hand which is described by vast amounts of scientific, forensic or business data from heterogeneous sources.

- To reach this goal, visual analytics combines the strengths of machines with those of humans. On the one hand, methods from knowledge discovery in databases (KDD), statistics and mathematics are the driving force on the automatic analysis side, while on the other hand human capabilities to perceive, relate and conclude turn visual analytics into a very promising field of research.
11.2 Scope of Visual Analytics

- Historically, Visual Analytics has evolved out of the fields of information visualization (InfoVis) and scientific visualization (SciVis).

InfoVis + SciVis
11.2 Scope of Visual Analytics

– Nowadays fast computers and sophisticated output devices create meaningful visualizations and allow us not only to mentally visualize data and concepts, but also to see and explore an exact representation of the data under consideration on a computer screen.

– However, the transformation of data into meaningful visualizations is not a trivial task that will automatically improve through steadily growing computational resources.
11.2 Scope of Visual Analytics

- Very often, there are many different ways to represent the data under consideration and it is unclear which representation is the best one.
- State-of-the-art concepts of representation, perception, interaction and decision-making need to be applied and extended to be suitable for visual data analysis.
11.2 Scope of Visual Analytics

- The fields of information and scientific visualization deal with visual representations of data.
- The main difference among the two is that scientific visualization examines potentially huge amounts of scientific data obtained from
  - sensors,
  - simulations or
  - laboratory tests.
11.2 Scope of Visual Analytics

- Typical scientific visualization applications are flow visualization, volume rendering, and slicing techniques for medical illustrations.

- In most cases, some aspects of the data can be directly mapped onto geographic coordinates or into virtual 3D environments.
11.2 Scope of Visual Analytics

- We define Information Visualization more generally as the communication of abstract data relevant in terms of action through the use of interactive interfaces.
11.2 Scope of Visual Analytics

- There are three major goals of visualization, namely
  - Presentation,
  - Confirmatory analysis, and
  - Exploratory analysis
11.2 Scope of Visual Analytics

- For presentation purposes, the facts to be presented are fixed a priori, and the choice of the appropriate presentation technique depends largely on the user. The aim is to efficiently and effectively communicate the results of an analysis.
11.2 Scope of Visual Analytics

- For confirmatory analysis, one or more hypotheses about the data serve as a starting point. The process can be described as a goal-oriented examination of these hypotheses. As a result, visualization either confirms these hypotheses or rejects them.
11.2 Scope of Visual Analytics

- Exploratory data analysis as the process of searching and analyzing databases to find implicit but potentially useful information, is a difficult task. At the beginning, the analyst has no hypothesis about the data.
11.2 Scope of Visual Analytics

- Visual Analytics is more than only visualization. It can rather be seen as an integral approach combining visualization, human factors and data analysis.

- Figure 1 illustrates the detailed scope of Visual Analytics. Concerning the field of visualization, visual analytics integrates methodology from
  - information analytics,
  - geospatial analytics, and
  - scientific analytics.
11.2 Scope of Visual Analytics

- Especially human factors (e.g., interaction, cognition, perception, collaboration, presentation, and dissemination) play a key role in the communication between human and computer, as well as in the decision-making process.
11.2 Scope of Visual Analytics

- In this context, production is defined as the creation of materials that summarize the results of an analytical effort, presentation as the packaging of those materials in a way that helps the audience understand the analytical results in context using terms that are meaningful to them, and dissemination as the process of sharing that information with the intended audience.
11.2 Scope of Visual Analytics

– In matters of data analysis, visual analytics furthermore profits from methodologies developed in the fields of data management & knowledge representation, knowledge discovery and statistical analytics.

– Note that visual analytics is not likely to become a separate field of study but its influence will spread over the research areas it comprises.
Information Visualization WS 2013/14

11  Visual Analytics

• 11.2 Scope of Visual Analytics

Fig. 1. The Scope of Visual Analytics
11.3 Historical Background

- Independent development of
  - Automatic analysis techniques such as statistics and data mining
  - Visualization techniques

- Key thoughts changed the rather limited scopes
  \(\Rightarrow\) Today called Visual analytics research

Most important step was to move from confirmatory analysis to exploratory analysis.
11.3 Historical Background

- Later the availability of graphical user interfaces with interaction techniques
  ➔ Information visualization

- Potential of integrating the user in the KDD process by
  - effective and efficient visualization techniques
  - interaction capabilities
  - knowledge transfer
  ➔ Visual data exploration or visual data mining
11.3 Historical Background

- Visual data mining was already a combination of visualization techniques and data analysis techniques.
- The term “Visual Analytics” was finally coined by Jim Thomas in the research and development agenda “Illuminating the Path”
  ➔ Homeland Security in the United States
11.3 Historical Background

Meanwhile, the term is used in a wider context, describing a new multidisciplinary field that combines various research areas including:
- Visualization
- Human-computer interaction
- Data analysis
- Data management
- Geo-spatial and temporal data processing
- Statistics
- ....
11.4 Scientific Fundamentals

- Visual Analytics evolved from Information Visualization and automatic data analysis.
- It combines former independent fields and strongly encourages human interaction in the analysis process, see Figure.
11.4 Scientific Fundamentals

- Visualization has three major goals: presentation, confirmatory analysis, and exploratory data analysis.
- Visual Analytics is more than pure visualization.
- Combination of visualization, human factors, and data analysis.
11.4 Scientific Fundamentals

- Integration of methodologies from information analytics, geospatial analytics, and scientific analytics.
- Human factors play a key role: Interaction, cognition, perception, collaboration, presentation, and dissemination. (between human and computer and decision-making process)
- Data analysis: Profits from methodologies developed in the fields of statistical analytics, data management, knowledge representation, and knowledge discovery.
11.4 Scientific Fundamentals

- Visual Analytics Process
11 Visual Analytics

11.5 Key Applications

- Physics and astronomy
- Astrophysics
- Monitoring climate and weather
- Emergency management
- Security and geographics
- Biology and medicine
- Business intelligence
- ...