Enhancing Visual Analytics with Guidance and Progression

Christian Tominski
Institute for Visual & Analytic Computing
University of Rostock
2020-02-20
Visual Analytics

"The purpose of computing is insight, not numbers.
— Richard Hamming, 1962

"Visualization offers a method for seeing the unseen. It enriches the process of scientific discovery and fosters profound and unexpected insights.
— McCormick et al., 1989

"Visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces.
— Thomas & Cook, 2005
Visual Analytics

**Information seeking mantra** (Shneiderman, 1996)
- Overview first,
- Zoom and filter,
- Then details-on-demand.

**Visual analytics mantra** (Keim et al., 2006)
- Analyse first,
- Show the important,
- Zoom, filter and *analyse further*,
- Details on demand
Visual Analytics

Van Wijk’s model of visualization adapted to visual analytics

Adapted from van Wijk, 2006
Challenges

• **Large and complex data** hard to analyze
  • Complex objectives: Where to look and what to do?
  • Complex methods and tools: How to use and parameterize?
  • Long-running computations: Slow response times!
  • Black-box computations:
  • No insight into result generation!

**Goal:** Keep the data analysis going on the human side and on the machine’s side
Outline

• **Guidance**: Support for the human part
  • Characterization
  • Conceptual model
  • Examples

• **Progression**: Enhancing the machine part
  • Conceptual model
  • Implementation
  • Applications
Guidance

- **Definition**
- First conceptualization (Schulz et al., 2013)
- Definition and characterization (Ceneda et al., 2017)

"**Guidance** is a computer-assisted **process** that aims to actively resolve a **knowledge gap** encountered by users during an **interactive** visual analytics session.

— Ceneda et al., 2017
Aspects of Guidance

• **Knowledge gap** – Why is guidance needed?
• **Input** – What information can be utilized for providing guidance?
• **Output** – How is guidance conveyed and how does it look like?
• **Degree** – How much help does guidance provide?
Aspects of Guidance

• **Knowledge gap** – Why is guidance needed?
  • Target
  • Path

• **Input** – What information can be utilized for providing guidance?
  • Data
  • Images, Specification
  • Domain, Knowledge, History
Aspects of Guidance

• **Output** – How is guidance conveyed and how does it look like?
  • Visual cues to indicate
  • Options to be selected
  • Automatic specification

• **Degree** – How much help does guidance provide?
  • Orienting
  • Directing
  • Prescribing
Conceptual Model of Guidance
Guided Multi-scale Exploration of Time Series

• Example
  • Exploring biological simulation data (Luboschik et al., 2013)
  • 1.7M time steps, each x-coordinate covers about 1,000 time steps

• Knowledge gap
  • Where and at what scale should the data be studied in detail?
Guided Multi-scale Exploration of Time Series

- **Guidance input**
  - Compute multi-scale differences
Guided Multi-scale Exploration of Time Series

• **Guidance output**
  • Difference bands visualize differences between scales

• **Guidance degree**
  • Difference bands provide orientation as to where potentially interesting facts about the data can be found
Guided Multi-scale Exploration of Time Series

Demo: Software courtesy of Martin Luboschik
Guided Multi-scale Exploration of Time Series

Demo: Software courtesy of Martin Luboschik

By Martin Luboschik. Licensed under CC BY.
Further Examples of Guidance

Model-driven guidance, Streit et al., 2012

Navigation recommendations, Gladisch et al., 2013

Guided parameterization, Ceneda et al., 2018
Guidance

• **Open Questions**
  • How to evaluate effectiveness of guidance?
  • How can the knowledge gap be inferred by the system?
  • How can the knowledge gap be conveyed to the system?
  • When is the right moment to provide guidance?
  • What degree of guidance is appropriate?
  • What are suitable visual encodings?
  • ...
Outline

• **Guidance**: Support for the human part
  - Characterization
  - Conceptual model
  - Examples

• **Progression**: Enhancing the machine part
  - Conceptual model
  - Implementation
  - Applications
Visual Analytics

• **Standard approach**
  - Monolithic operators process the data as a whole
  - For large data, processing time can be substantial
  - Resulting latency has adverse effects on visual data analysis

"Especially for **large datasets**, supporting **real-time interactivity** requires careful attention to **system design** and poses important research **challenges** ranging from low-latency architectures to intelligent sampling and aggregation methods.

— Heer & Shneiderman, 2012
Progressive Visual Analytics

• **Progressive approach** (Stolper et al., 2014)
  • Generate **partial results** of increasing completeness and correctness
  • Basic steps
    • Subdivide computations
    • Subdivide data
  • Key advantages
    • Responsiveness of the system
    • Transparency of the involved calculations
    • Control of the visual data analysis
Modelling Progressive Visual Analytics

- **Progressive operators** (Schulz et al., 2016)
  - Generate partial results of increasing quality
  - Number of results depends on
    - Algorithmic factors
    - Human factors

- **Progressive transitions** (Schulz et al., 2016)
  - Generate, transmit, and subsume chunks
  - Increase and reduce chunk granularity
  - Chunking strategies
    - Incremental chunking: Chunk samples
    - Semantic chunking: Chunk data aspects
    - Level-of-detail chunking: Chunk abstractions

Example of progressive pipeline
Implementing Progressive Visual Analytics

• **Multi-threading architecture** (Piringer et al., 2009)
Implementing Progressive Visual Analytics

• Standard approach

- Human
- Computer
- Request
- Result

• Progressive approach

- Human
- Control
- Processing
- Request
- Partial result
- Partial result
- Partial result
- Complete result

Interaction latency
Feedback latency
Applications of Progressive Visual Analytics

- Application scenarios
Progressive Data Processing

- 370,000 car crashes, progressive search on chunks of 5,000 crashes
Progressive Graph Layout

- Social network with 747 nodes and 60,050 edges
Progressive Network Mapping

- Climate network with 6,816 nodes and 232,940 edges
Progressive Display

- Using JPEG2000 for progressive image transmission (Rosenbaum et al., 2011)
Progressive Visual Analytics

• **Open Questions**
  • How to subdivide data and processing best?
  • How to measure progress?
  • When are intermediate results reliable?
  • How to communicate trustworthiness/uncertainty of intermediate results?
  • How to react to user interventions?
  • ...
Summary

- **Guidance**: Support the human part
- **Progression**: Enhancing the machine part
Thank you!

https://ivda-book.de