Visual Analytics
Theory and Application Examples

Harald Piringer
VRVis Research Center
Vienna, Austria
Overview

- VRVis – Overview

- Part 1: Visual Analytics – Theory
  - Definition
  - Motivation
  - Process

- Part 2: Application Examples
  - Statistical modeling of natural gas consumption
  - Process optimization
  - Anomaly detection in facility management data
  - Tunnel surveillance
VRVis – Research Center for Virtual Reality and Visualization

- Founded in 2000
- Located in TechGate, 1220 Vienna
- Staff approx. 70 researchers
- **Mission**: application-oriented research in Visual Computing for linking science to industry
- **Research focus**: Computer graphics, Scientific Visualization, Visual Analytics
- Close cooperation with national and international universities and companies
- **Students welcome!!** (Praktika, Diplomarbeiten)
Part 1: Visual Analytics
Definition: Visual Analytics

- ... is the science of **analytical reasoning** facilitated by **interactive visual interfaces** [Thomas 2005]
- ... combines **automated analysis** techniques with **interactive visualizations** for an effective understanding, reasoning and decision making on the basis of very **large and complex datasets** [Keim et al. 2010]
Why Visual Analytics?

...complex tasks require all abilities!
Tasks – Some Examples

- **Data Preparation**
  - Detection of data quality issues
  - Selective export
  - Aggregation

- **Exploratory analysis**
  - Understanding data
  - Identify relationships
  - Detect structures

- **Model building**
  - Feature selection
  - Validation

- **Analysis and optimization of complex systems**
  - Sensitivity analysis
  - Multi-objective optimization

- **Surveillance and monitoring**
Comparing Manual and Automated Approaches

- **Interactive Visualization (~ InfoVis)**
  - user-guided analysis possible
  - possible for contradicting, unclean data
  - understand results in context
  - uses power of human visual system
    - human involvement not always possible or desirable (expensive!)
    - limited dimensionality
    - often only qualitative results
    - (still) often unfamiliar

- **Automated Data Analysis (~ Statistics)**
  - precise definition of goals necessary
  - limited tolerance of data artifacts
  - result without explanation
  - computationally expensive
  - (after setup): results without human involvement → often only viable app.
    - scales better w.r.t. many dimensions
    - precise results
    - long history (mostly statistics)
Visual Analytics: Summary

- Multidisciplinary and integrated approach combining
  - Computer graphics (*InfoVis*, also *SciVis* and *Rendering*)
  - Computation-based processing of data (*statistics*, *data mining*, *image processing*, *simulation*, etc.)
  - Human factors (*interaction*, *cognition*, *perception*, *collaboration*, etc.)
  - And more (*databases*, *distributed systems*, *knowledge-representation*, *GIS*, etc.)

- Highly interdisciplinary
Part 2: Application Examples
Example 1: Energy Data

- **Background**
  - Hourly measurements of natural gas consumptions over 5 years
  - Several meteorological quantities

- **Goals**
  - Primary: Build accurate prediction model for gas consumption
  - Secondary: Identify data quality issues
Excursus: Visplore

- **Software for exploratory data analysis developed at VRVis**

- **Flexible Visualization**
  - More than 10 visualization types for multivariate data
  - Freely configurable and combinable

- **Interaction**
  - Parameters adjustable at any time (e.g., scaling, coloring, ...)
  - Various ways of data selection for linking views in real time
  - Explicit support for building and validating regression models
  - Export data, images, etc.

- **Scalability**
  - Real time response also for millions of data values
Viplore – Technicals

- Written in C++
- OpenGL used for graphics
- gtk+ used for GUI
- Designed for large data
  - highly multi-threaded
  - quick preview during interaction
  - memory management
- Open architecture
  - views / importers / etc. are plugins
Applications of Visplore

- **Distribution partners**
  - **AVL**: Simulation-based optimization in automotive industry
  - **Plasmo**: Surveillance and optimization of production processes
  - **Hakom**: Building prediction models for the energy sector
  - **AI-MS**: Optimization of aviation infrastructure planning

- **Applied by**
  - **Austrian Power Grid**: Optimization prediction models (e.g., for renewable energy)
  - **Kapsch TrafficCom**: Calibration of telematics infrastructure
  - **IC-Consulenten**: Temporal analysis of malfunction messages
  - **Baxter**: Analysis of clinical trials
  - ... and several others
Interested...?

- Internship / diploma thesis
  - Contribute to visplore!
  - **Interesting topics** oriented towards practice
  - Large-scale real-world project
  - Help and supervision
  - Payment
  - Ask me or write me an email: hp@vrvis.at

- Questions?
Example 2: Production Process

- **Background**
  - Measurements from a welding seam
  - Parameters controlling the welding process

- **Goals**
  - Primary: Detect reason for insufficient quality
  - Secondary: Analyze quality development over time
Example 3: Simulation Data

- **Background**
  - 100 simulation runs of a bearing of a car engine

- **Goals**
  - **Primary:** Assess accuracy and plausibility of simulated data
  - **Secondary:** Identify problematic parameter combinations
Example 4: Facility Management

- **Background**
  - ~ 240,000 malfunction messages of a large infrastructure facility
  - Data contains time, functional unit, building, type of malfunction

- **Goals**
  - Primary: Detect undesirable trends
  - Secondary: Detect data quality issues
Example 5: Surveillance Calibration

- **Background**
  - Alarm messages from video-based tunnel surveillance

- **Goals**
  - **Primary:** Identify parameters to minimize wrong positives
  - **Secondary:** Better understand detection
Example 6: Surveillance of Road Tunnels

- **Background**
  - Video surveillance: cameras every 80 to 100 meters
  - Live detection of incidents (e.g., smoke, pedestrians, lost cargo)
Example 6: Surveillance of Road Tunnels

- **Current challenges**
  - Current systems do not (perceptually) scale to numerous events → information overload in disaster scenarios
  - Difficult to understand temporal development
  - No easy access to historic video material
Example 6: Surveillance of Road Tunnels

- Data

Realtime video analysis
Stream of raw events
Processing of events
Example 6: Surveillance of Road Tunnels

Our solution: overview

- Present View
- History View
- Future View
Conclusion

- Many application areas for Visual Analytics
- Suitable to address challenging tasks
  - Require intelligent algorithms AND human domain knowledge
  - Tight integration of visualization, interaction, computation
- Versatile tools are needed to provide this tight integration