Interactive Visual Analysis of Multi-faceted Scientific Data

Johannes Kehrer
VRVis Research Center
Vienna, Austria

[joint work with Helwig Hauser, University of Bergen, Norway]
Multi-faceted Scientific Data

- **Spatiotemporal data**
- **Multi-variate data**
  (multiple data attributes, e.g., temp. or pressure)
- **Multi-modal data**
  (simulation, satellite imagery, weather stations, etc.)
- **Multi-run simulations**
  (simulation repeated with varied settings for model parameters)
- **Multi-model scenarios**
  (e.g., coupled climate model)
Categorization

- Literature review of about 200 papers on scientific data
- How are vis., interaction and comput. analysis combined?

**how to represent the data**

**interaction concepts** (linking & brushing, zooming, panning, view reconfiguration, etc.)

**what are main characteristics / features**

**visual mapping** ---**interactive visual analysis**--- **comput. analysis**

- visual data fusion
- relation & comparison
- navigation
- focus+context & overview+detail
- interactive feature spec.
- data abstraction & aggregation

[compare to Keim et al. 2009; Bertine & Lalanne 2009]
**Visual vs. Computational Analysis**

- **Interactive Visual Analysis**
  - user-guided analysis possible
  - detect interesting features without looking for them
  - 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**
  - needs precise definition of goals
  - limited tolerance of data artifacts
  - result without explanation
  - computationally expensive
  - hardly any interaction required (after setup)
  - scales better w.r.t. many dimensions
  - precise results
  - long history (mostly statistics)
Fusion at different stages of vis. pipeline [Fuchs & Hauser 2009]

- visual properties (e.g., glyphs, texture, color)
- layering techniques / transparency
- feature-based approaches (e.g., visual styles)

Helix glyphs [Tominski et al. 05]  
Layering [Kirby et al. 99]  
Feature-based vis. [Viola et al. 05]
Visual Fusion

Layering techniques [Wong ‘02]

- opacity modulation
- filigreed
- colormap enhancement
- 2D heightmap

colormap + square wave modulation
Preattentive Graphical Attributes

Textures and Colors [Healey & Enns 02]

- temperature $\rightarrow$ color
- wind speed $\rightarrow$ coverage
- pressure $\rightarrow$ size
- precipitation $\rightarrow$ orientation
Glyph-based Visualization [Lie et al. 09]

Glyph color: Normalized Temperature

Glyph Rotation (-45°, 45°): Flow Velocity

Pressure

Vapor
- Investigate similarities & differences
  - linking & brushing
  - repositioning

- **Taxonomy** [Gleicher et al. 2011]
  - side-by-side comparison
  - overlay in same coord. system
  - explicit encoding of computed differences / correlations
Comparison of Time-oriented Data

Theme river [Havre et al. 01]

Helix glyphs [Tominski et al. 05]

two-tone color [Saito et al. 05]
Comparison using Difference Views
[Daee Lampe et al. 2010]
Interactive search, zooming and panning

Grand tour [Asimov 85]

Quality metrics for high-dimensional data visualization [Bertini et al. 2011]
- **why**: ranking, view optimization, etc.
- **what**: clustering, correlations, outliers, image quality, etc.
- **where**: data vs. image space
Scatterplot Matrix Navigation [Elmqvist et al. 2008]

3D transition between 2 scatterplots

Scatterplot matrix
- **Focus+context visualization**
  - different graphical resources (space, opacity, color, etc.)
  - focus specification (e.g., by pointing, brushing or querying)

- **Clustering & outlier preservation**

Outlier-preserving focus+context [Novotný & Hauser 06]
- Brushing in multiple linked views
- Tight integration with supervised machine learning

Visual human+machine learning [Fuchs et al. 09]

User hypothesis
Selection attributes

Fitness

65%

71%

J. Kehrer
Select function graphs based on similarity

- pattern sketched by user
- similarity evaluated on gradients (1st derivative)

Advanced brushing [Muigg et al. 2008]
Algorithmically extract values & patterns

- dimensionality reduction (PCA, SOM, MDS)
- aggregation, summary statistics
- outliers, clustering, etc.

Glyph-based overview [Kehrer et al. 11]  
[Andrienko & Andrienko 11]
Cluster Calendar View [vanWijk & van Selow ’99]

time series clustered by similarity (K-means)
Open Issues

- **Scientific data are getting multi-faceted**

- **How to deal with data heterogeneity?**
  - majority of methods only address one facet of data
  - coordinated multiple views with linking & brushing
  - investigation of features across data sets & levels of abstraction

- **What are vis., interaction & comput. analysis good for?**
  - analytical methods can control steps in visualization pipeline (e.g., visualization mapping or quality metrics)
  - interactive feature specification + machine learning
Thank you for your attention!

Acknowledgements