Interactive Visual Analysis of Complex Data: Sets; Simulation Steering

Krešimir Matković
Vienna University of Technology, VRVis Research Center, Vienna
Towards Quantitative Visual Analytics with Structured Brushing and Linked Statistics

Sanjin Radoš
VRVis Research Center
Vienna, Austria
Motivation

Improve reproducibility of brushing results

Enable a quantitative reading of linked views to support, for example, decision making
Structured Brushing Space

Brushing operation
- Unconstrained (traditional)
- Constrained
- Automatic

Structured space
- Anchoring
- Extent
- Movement
Structured Brushing Space

Snap-to-Grid Brush

- confine brushing to reproducible shapes
- value-based or rank-based analysis (percentile grid)
- exact brush movement (value-accurate brushing)
Structured Brushing Space

Percentile Brush
- the brush always contains a predefined number of items, like 10%
- when moved, extent/radius of the brush is adapted

Mahalanobis Brush
- takes the data distribution into account (adapts its shape, too)
Quantitative Linked Views

Quantify the analysis results

- provide descriptive statistics about the brushed data
Quantitative Linked Views

Support the comprehension of data changes using

- paths of the center points in a scatterplot
- traces in the Trace View (for a selected dimension only)
Quantitative Linked Views

Support the comprehension of data changes using

- the relative difference plot
Interactive Visual Analysis of Set-Typed Data

Wolfgang Freiler, Kresimir Matkovic, and Helwig Hauser

IEEE InfoVis 2008
Sets as dimension

Conventional approach deals with n-dim. Euclidian spaces

Each item is a point in n-dim. Space (n-tuple)

Set as dimension

New view

Demo!

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DaysToDeath</th>
<th>Status</th>
<th>Drug</th>
<th>Age</th>
<th>Days</th>
<th>Sex</th>
<th>Ascites</th>
<th>Hepatoma</th>
<th>Spiders</th>
<th>Edema</th>
<th>Bilirubin</th>
<th>Cholesterol</th>
<th>Albumin</th>
<th>Copper</th>
<th>Alphos</th>
<th>SGOT</th>
<th>Tryglic</th>
<th>Platelets</th>
<th>ProthrombinTime</th>
<th>Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values</td>
<td>int</td>
<td>0 / 1</td>
<td>0 / 1</td>
<td>int</td>
<td>int</td>
<td>int</td>
<td>int</td>
<td>int</td>
<td>int</td>
<td>int</td>
<td>int</td>
<td>int</td>
<td>int</td>
<td>int</td>
<td>int</td>
<td>int</td>
<td>int</td>
<td>int</td>
<td>int</td>
<td>int</td>
</tr>
<tr>
<td>400</td>
<td>dead</td>
<td>plac</td>
<td>21464</td>
<td>f</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.0</td>
<td>15</td>
<td>261</td>
<td>2.6</td>
<td>156</td>
<td>1718</td>
<td>138</td>
<td>172</td>
<td>190</td>
<td>12</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4500</td>
<td>alive</td>
<td>plac</td>
<td>20617</td>
<td>f</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.0</td>
<td>1.1</td>
<td>302</td>
<td>4.1</td>
<td>54</td>
<td>7395</td>
<td>114</td>
<td>88</td>
<td>221</td>
<td>11</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1012</td>
<td>dead</td>
<td>plac</td>
<td>25594</td>
<td>m</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>1.4</td>
<td>176</td>
<td>3.5</td>
<td>210</td>
<td>516</td>
<td>96</td>
<td>55</td>
<td>151</td>
<td>12</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2466</td>
<td>dead</td>
<td>pen</td>
<td>19379</td>
<td>f</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0</td>
<td>0.3</td>
<td>280</td>
<td>4.0</td>
<td>52</td>
<td>4651</td>
<td>28</td>
<td>189</td>
<td>373</td>
<td>11</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4079</td>
<td>dead</td>
<td>plac</td>
<td>16261</td>
<td>m</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.0</td>
<td>2.1</td>
<td>456</td>
<td>4.0</td>
<td>124</td>
<td>5719</td>
<td>222</td>
<td>230</td>
<td>70</td>
<td>10</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Venn Diagrams introduced in 1880 by John Venn (1834–1923) in a paper entitled *On the Diagrammatic and Mechanical Representation of Propositions and Reasonings* in the "Philosophical Magazine and Journal of Science" (from Wikipedia)
Set Type Data

Use of Histogram and Scatterplot – possible

Set-o-Gram

Demo!
Set Type Data
Set Type Data – Radial Sets

B. Alsallakh, W. Aigner, S. Miksch, and H. Hauser: Radial Sets: Interactive Visual Analysis of Large Overlapping Sets
Set Type Data – Upset

Interactive Visual Steering
InfoVis Reference Model

Data
- Raw Data
- Data Tables

Data Transformations

Visual Form
- Visual Structures
- Views

Visual Mappings

View Transformations

Human Interaction (controls)
Motivation

Simulation

- widely used in science and engineering
- helps in understanding various phenomena
- recent trend:
  - multiple simulation runs
  - variation of various control parameters
  - study influence on output
  - understanding of simulation model
  - understanding of physical phenomena
Simulation and Visualization 1

Common practice:

- define a model
- run simulation
- visualize results
- change the model
- run simulation ...
Simulation and Visualization 2

Multiple runs:

- define a model
- define control parameters
- run simulation
  - for each combination of control parameters
  - for some of the combinations (steering)
- use interactive visual analysis
  - to understand (optimize) the model
  - to understand the phenomena
Interactive Visual Steering

Model, parameters, many simulation runs

How many is many?

- 11 parameters / 10 variations each
- $10^{11}$ combinations
- 10 simulations per minute (very fast) - $10^{10}$ minutes
- 19 000 years!
Basic Idea

Start with a simple model
Assume the rest to be ideal
Tune the first model
Expand the model
Tune expanded model
Expand the model ...

It is always possible to go back!
Basic Idea

- Start with a simple model
- Assume the rest to be ideal
- Tune the first model
- Expand the model
- Tune expanded model
- Expand the model …

It is always possible to go back!
Interactive Steering

3 loops
- Explore initial results set
- Change parameter values
- Change model

Tabular data
- Initial table
- Add / remove rows
- Add / remove columns
Interaction

3 loops

- A – real time interaction
- B – “live” parameter refinement
  (A loop functioning)
- C – model change - after initial runs and data management update
  same as B
First Step

Start with a simple model
Assume the rest to be ideal
Tune the first model
4 parameters, 750 combinations
12 minutes
Two targets
Interactive Visual Analysis
Interactive Visual Analysis
Interactive Visual Analysis
Second Step
Second Step

Different responses

High CVsize – unwanted behavior

Narrow curves – low injected mass

No pilot injection

Same actuator!
Second Step
Parameter refinement

Additional investigation needed

– refine parameters
Output at various steps
Final model

Everything is set now
We wanted to check all output curves
Actuator variations
Interesting peaks detected!
Final model

Everything is set now
We wanted to check all output curves
Actuator variations
Interesting peeks detected!
Final solution

One operation point
Choose one and test
If everything OK – done!

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Name</th>
<th>Final value</th>
</tr>
</thead>
<tbody>
<tr>
<td>I_1</td>
<td>d_sac</td>
<td>0.75 mm</td>
</tr>
<tr>
<td>I_2</td>
<td>Alpha_seat</td>
<td>50 degrees</td>
</tr>
<tr>
<td>I_3</td>
<td>C_turb</td>
<td>0.9</td>
</tr>
<tr>
<td>I_4</td>
<td>mju</td>
<td>0.6</td>
</tr>
<tr>
<td>II_1</td>
<td>CV_size</td>
<td>10 mm³</td>
</tr>
<tr>
<td>II_2</td>
<td>Z_inl</td>
<td>1.6</td>
</tr>
<tr>
<td>II_3</td>
<td>Z_out</td>
<td>2.6</td>
</tr>
<tr>
<td>III_1</td>
<td>Res_bypass</td>
<td>2.0</td>
</tr>
<tr>
<td>III_2</td>
<td>Res_Outlet</td>
<td>1.0</td>
</tr>
<tr>
<td>III_3</td>
<td>Area_Bypass</td>
<td>0.032 mm²</td>
</tr>
<tr>
<td>III_4</td>
<td>Area_Valve</td>
<td>0.07mm²</td>
</tr>
<tr>
<td>IV_1</td>
<td>HPP_Length</td>
<td>300mm</td>
</tr>
<tr>
<td>IV_2</td>
<td>RV_size</td>
<td>30 cm³</td>
</tr>
</tbody>
</table>
A Hybrid Approach to Visual Steering
Complex Engineering Systems

Complex System

- not a simple combination of its components
- have many parameters
- complicated to: understand, tune, and “optimize”
Simulation Ensemble

Different sets of control parameter values

Complex data with high dim. parameter space

Advanced analysis needed:

*Hybrid Visual Steering*
Simulation Model
Who needs it?

If you are dealing with a complex system
• like a car engine injection system, for example;
Who needs it?

If you are dealing with a complex system
• like a car engine injection system, for example;

and your system has a high dimensional parameter space
• tens or hundreds of parameters;
Who needs it?

If you are dealing with a complex system
• like a car engine injection system, for example;

and your system has a high dimensional parameter space
• tens or hundreds of parameters;

and you can run a simulation of your system fast
• in less then a second, or so;
Who needs it?

If you are dealing with a complex system
  • like a car engine injection system, for example;

and your system has a high dimensional parameter space
  • tens or hundreds of parameters;

and you can run a simulation of your system fast
  • in less then a second, or so;

and your simulation results are not simple scalars only
  • but curves, surfaces, etc.
Who needs it?

If you are dealing with a complex system
• like a car engine injection system, for example;

and your system has a high dimensional parameter space
• tens or hundreds of parameters;

and you can run a simulation of your system fast
• in less then a second, or so;

and your simulation results are not simple scalars only
• but curves, surfaces, etc.

You do need the Hybrid Steering!
General Simulation

Simulation tries to compute different physical properties using a simulation model and a set of control parameters.

\[ \mathbb{R}^n \rightarrow \text{Simulation} \rightarrow \text{Complex Simulation Results} \rightarrow \text{Extracted Scalar Features} \rightarrow \mathbb{R}^m \]

- Time series, Surfaces, Volume data, Vector fields, ...
Analysis Workflow - Basic Setup

1. Control parameters and scalar features can be visualized

2. Complex results can also be visualized
Analysis Workflow – Ensemble

Simulation ensemble is created and interactively explored

DOE module needed, extensive research available
Analysis Workflow – Steering

The loop is closed
We can initiate new runs interactively!
Many automatic optimization methods exist. Regression model based is a popular way.
Analysis Workflow – Hybrid Steering

Optimization only – 2 sources of uncertainty
- regression and feature extraction

Visual steering only
- tedious search for areas of interest in parameter space

Solution:

Hybrid Visual Steering
Analysis Workflow – Hybrid Steering

Automatic Optimization

Design of Experiment

Control Parameters

Simulation

Complex Simulation Results

Extracted Scalar Features

Automatic Optimization

Regression Model Building

Scalar parameters

Regression model

Interactively defined model domain

Scalar features

Model based optimum points

Interactive Visualization

Regression model

Scalar features

Time dependent results

Interactively defined design points

Standard simulation

Complex results

Ensemble steering

Ensemble simulation

Automatic optimization

Hybrid steering

Krešimir Matković

Analysis Workflow – Hybrid Steering

Automatic Optimization

- Design of Experiment
- Control Parameters
- Simulation
- Complex Simulation Results
- Extracted Scalar Features
- Regression Model Building
- Regression model
- Interactively defined model domain
- Scalar features
- Model based optimum points

Interactive Exploration, Analysis and Optimization

- Interactively defined design points
- Time dependent results
- Scalar parameters

Standard simulation

Ensemble simulation

Complex results

Automatic optimization

Ensemble steering

Hybrid steering
Analysis Workflow – Hybrid Steering

Design of Experiment → Control Parameters → Simulation → Complex Simulation Results → Extracted Scalar Features → Regression Model Building → Regression model → Interactively defined model domain → Scalar features → Automatic Optimization → Model based optimum points

Automatic Optimization

Interactive Exploration, Analysis and Optimization

Scalar parameters → Interactively defined design points → Interactively defined model domain → Time dependent results → Standard simulation

Standard simulation

Complex results

Automatic optimization

Ensemble steering

Hybrid steering
Analysis Workflow – Hybrid Steering

Automatic Optimization

Design of Experiment

Control Parameters → Simulation → Complex Simulation Results → Extracted Scalar Features

Regression Model Building

Scalar parameters

Regression model

Interactively defined model domain

Scalar features

Automatic Optimization

Interactive Visualization

Interactively defined design points

Time dependent results

Model based optimum points

Standard simulation

Complex results

Ensemble steering

Ensemble simulation

Automatic optimization

Hybrid steering
Analysis Workflow – Hybrid Steering

Design of Experiment → Control Parameters → Simulation → Complex Simulation Results → Extracted Scalar Features → Regression Model Building → Regression model → Automatic Optimization

Interactive Exploration, Analysis and Optimization

Automatic Optimization

Interactive Visualization

Standard simulation

Ensemble simulation

Complex results

Automatic optimization

Ensemble steering

Hybrid steering
Analysis Workflow – Hybrid Steering

Automatic Optimization

Design of Experiment

Control Parameters → Simulation → Complex Simulation Results → Extracted Scalar Features → Automatic Optimization

Regression Model Building

Scalar parameters → Regression model → Interactively defined model domain → Scalar features

Interactive Exploration, Analysis and Optimization

Scalar parameters → Interactively defined design points

Interactive Visualization

Regression model → optimum points

Standard simulation

Complex results

Automatic optimization

Ensemble steering

Hybrid steering
View Design

New view which supports identified requirements

Fully integrated in CMV system

1. Show distribution – histogram like

```
V_rail
[4000.00, 10000.00]
max count = 1003
```
View Design

New view which supports identified requirements

Fully integrated in CMV system

2. Empty bins problem
View Design

New view which supports identified requirements

Fully integrated in CMV system

3. Show brushed items
View Design

New view which supports identified requirements

Fully integrated in CMV system

4. Set optimization constraints – click on bin to add/subtract range, fine tune the range if necessary
View Design

New view which supports identified requirements

Fully integrated in CMV system

5. Works for all parameters
View Design

```
pilot_start
[0.00, 4.00]
max count = 950

P_rail
[800.00, 1600.00]
max count = 981

V_rail
[4000.00, 10000.00]
max count = 1003

V_inlet
[10.00, 30.00]
max count = 1391

D_line
[1.00, 3.00]
max count = 1062

L_line
[100.00, 500.00]
max count = 705
```
View Design

New view which supports identified requirements

Fully integrated in CMV system

6. Show optimum(s)
View Design

New view which supports identified requirements

Fully integrated in CMV system

6. Show optimum(s)
View Design

- **pilot_start**
  - Range: [0.00, 4.00]
  - Max count: 950

- **P_rail**
  - Range: [800.00, 1600.00]
  - Max count: 981

- **V_rail**
  - Range: [4000.00, 10000.00]
  - Max count: 1003

- **V_inlet**
  - Range: [10.00, 30.00]
  - Max count: 1391

- **D_line**
  - Range: [1.00, 3.00]
  - Max count: 1062

- **L_line**
  - Range: [100.00, 500.00]
  - Max count: 705
View Design

- **pilot_start**
  - [0.00, 4.00]
  - max count = 702

- **P_rail**
  - [800.00, 1600.00]
  - max count = 981

- **V_rail**
  - [4000.00, 10000.00]
  - max count = 838

- **V_inlet**
  - [10.00, 30.00]
  - max count = 1144

- **D_line**
  - [1.00, 3.00]
  - max count = 900

- **L_line**
  - [100.00, 500.00]
  - max count = 622
Case Study – Initial Data

Six parameters varied – 2700 simulation runs

- rail volume and pressure $V_{\text{rail}}$ and $P_{\text{rail}}$
- high pressure pipes geometry $L_{\text{line}}$ and $D_{\text{line}}$
- rail-pipe junction volume $V_{\text{inlet}}$
- pilot injection start $\text{pilot\_start}$
Case Study – Initial Data

30 time dependent state variables per run

Scalar features extracted

Regression model built
Analysis Goals

We want balanced system and balanced individual injectors

Main goals:

- small difference between injection pressures of individual pilot injections for each cylinder
- small difference between amounts of fuel injected of individual pilot injections for each cylinder
- as fast as possible needle opening and closing velocities for pilot and main injections
Hybrid Steering Workflow

1. Initial runs available

2. Interactive exploration starts
Hybrid Steering Workflow

- Six parameters varied
- 2700 simulation runs
- 30 time dependent state variables per run computed
- Scalar features extracted
- Regression model built
- Interactive visual analysis used to explore the data
Hybrid Steering Workflow

3. Set optimization constraints

4. Compute regression model

5. Define optimization goals

```plaintext
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Constraints</th>
<th>Max Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>pilot_start</td>
<td>[0.00, 4.00]</td>
<td>540</td>
</tr>
<tr>
<td>P_rail</td>
<td>[800.00, 1600.00]</td>
<td>900</td>
</tr>
<tr>
<td>v_rail</td>
<td>[4000.00, 10000.00]</td>
<td>675</td>
</tr>
<tr>
<td>v_inlet</td>
<td>[10.00, 30.00]</td>
<td>900</td>
</tr>
<tr>
<td>D_line</td>
<td>[1.00, 3.00]</td>
<td>900</td>
</tr>
<tr>
<td>L_line</td>
<td>[300.00, 500.00]</td>
<td>540</td>
</tr>
</tbody>
</table>
```
Hybrid Steering Workflow

3. Set optimization constraints

4. Compute regression model

5. Define optimization goals

6. Compute and show optimum
Hybrid Steering Workflow

7. Compute points in the neighborhood of the optimum using the simulation tool - curves

8. Check if better optimum exists and is it feasible
Evaluation and User Feedback

Five domain experts in several analysis sessions, one coauthors the paper

"I could never set up optimization so fast. I also see all results together with the initial runs."

"The suggestions of where to refine the parameter space based on optimization speeds up the steering..."

“Seeing all runs all the time is simply unmatched in a conventional workflow. I can also see all curves...”

“I would estimate a speedup of at least an order of magnitude when designing complex systems.”
Conclusion

Complex systems and complex data require new analysis approaches

Close coupling of simulation, optimization, and visualization

Hybrid Visual Steering - an integrated system and workflow presented

Supports experts in exploration, analysis and optimization of complex engineering systems

Very positive feed-back from engineers
Thank you!

Special thanks for used materials to H. Hauser, and colleagues from VRVis!