Visual Analytics - Introduction

Eduard Gröller

Institute of Computer Graphics and Algorithms
Vienna University of Technology
Outline

■ Introduction to visual analytics
■ Definition of visual analytics
■ Technical challenges and agenda
■ Application areas

■ Some slides courtesy of
  ◆ Silvia Miksch
  ◆ Daniel Keim / Jim Thomas
Motivation: Main Problems

Data Unmanageable – Loss of Overview

Missing Integration of Various (Heterogeneous) Information Sources

Various Interdisciplinary Methods

Missing Involvement of Users and their Tasks
• 100 million FedEx transactions per day
• 150 million VISA credit card transactions per day
• 300 million long distance calls in AT&T’s network per day
• 50 billion e-mails worldwide per day
• 600 billion IP packets per day on DE-CIX backbone
New Requirements Summary

- Volume of data, orders of magnitude larger and different levels of abstraction
- Complexity of information spaces into very high dimensions, 200 the norm
- Information often out of context, incomplete, fuzzy
- Information in all media types: text, imagery, video, voice, web, sensor data
- Time and temporal dynamics fundamentally change the approach
- Spatial, yet non-spatial abstract data
- Multiple ontologies, languages, cultures

For many applications: we now turn to data-intensive visual analytics
### Visualization for Problem Solving

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>599</td>
<td>693</td>
</tr>
<tr>
<td>2</td>
<td>525</td>
<td>693</td>
</tr>
<tr>
<td>3</td>
<td>541</td>
<td>662</td>
</tr>
<tr>
<td>4</td>
<td>542</td>
<td>611</td>
</tr>
<tr>
<td>5</td>
<td>527</td>
<td>579</td>
</tr>
<tr>
<td>6</td>
<td>505</td>
<td>529</td>
</tr>
<tr>
<td>7</td>
<td>469</td>
<td>553</td>
</tr>
<tr>
<td>8</td>
<td>409</td>
<td>558</td>
</tr>
<tr>
<td>9</td>
<td>321</td>
<td>531</td>
</tr>
<tr>
<td>10</td>
<td>318</td>
<td>606</td>
</tr>
<tr>
<td>11</td>
<td>321</td>
<td>693</td>
</tr>
<tr>
<td>12</td>
<td>243</td>
<td>693</td>
</tr>
<tr>
<td>13</td>
<td>250</td>
<td>660</td>
</tr>
<tr>
<td>14</td>
<td>253</td>
<td>579</td>
</tr>
<tr>
<td>15</td>
<td>246</td>
<td>527</td>
</tr>
<tr>
<td>16</td>
<td>230</td>
<td>489</td>
</tr>
<tr>
<td>17</td>
<td>196</td>
<td>510</td>
</tr>
<tr>
<td>18</td>
<td>192</td>
<td>497</td>
</tr>
<tr>
<td>19</td>
<td>134</td>
<td>568</td>
</tr>
<tr>
<td>20</td>
<td>94</td>
<td>493</td>
</tr>
<tr>
<td>21</td>
<td>25</td>
<td>423</td>
</tr>
<tr>
<td>22</td>
<td>87</td>
<td>467</td>
</tr>
<tr>
<td>23</td>
<td>128</td>
<td>482</td>
</tr>
<tr>
<td>24</td>
<td>183</td>
<td>473</td>
</tr>
<tr>
<td>25</td>
<td>163</td>
<td>568</td>
</tr>
</tbody>
</table>

---

Danube University Krems  
Department of Information und Knowledge Engineering [Silvia Miksch et al.]
Analytical Methods

Screen Resolution: 1024 * 768 = 786,432

Measurements of Water Level in LA Every Year: 5,256,000

Number of Cellular Phones in Austria (2005): 8,160,000

Transmitted Emails Every Hours (World-Wide): 35,388,000

Whole Data often not Presentable

1. Applying Analytical Methods (Data Reduction)
2. Visualization of Most Important Data and Information

Analytical Methods

Statistics, Machine Learning & Data Mining
Interactions

Past
Only passive Observations
Representation not Changeable
„one fits all“

Today
Active Examination with Visualizations
Dynamically Adaptable and Modifiable
→ Different Users, Tasks and Aims
Visual Analytics – What is it?

James Thomas & Kristin A. Cook:
NVAC (National Visualization and Analytics Center), Seattle, USA

"Visual Analytics is the science of analytical reasoning facilitated by interactive visual interfaces"

[Thomas & Cook 2005]
Visual Analytics is the science of analytical reasoning facilitated by interactive visual interfaces.

People use visual analytics tools and techniques to

- Synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data.
- Detect the expected and discover the unexpected.
- Provide timely, defensible, and understandable assessments.
- Communicate assessment effectively for action.

“The beginning of knowledge is the discovery of something we do not understand.”
~Frank Herbert (1920 - 1986)
Why is the topic highly relevant today?

- Very Large Data Collections are available in Databases and Data Warehouses.
- On the Basis of the Data Complex Decisions have to be made in a timely fashion.
- Pure Visualization Methods (Information Visualisation) do not work for Billions of Data Records.
- Full Automatic Knowledge Discovery Approaches only work for well-defined and clearly specifiable problems.
- Especially for adversarial situations: Fraud, Viruses, SPAM, Attacks, Competition, …
What is new?

What do we have?
- Automatic Knowledge Discovery & Information Mining
- Interactive Visual Data-Exploration

What do we need?
Tight Integration of Visual and Automatic Data Analysis Methods with Database Technology for a Scalable Interactive Decision Support

Vis’07 – Scope and Challenges of Visual Analytics – Keim / Thomas
Real-time Analysis of

- very large, complex, dynamic information
- from many diverse data sources
- in diverse formats and resolutions
- in uncertain, potentially life-threatening, and time-critical situations.

“Discovery consists of seeing what everybody has seen and thinking what nobody has thought.”
~Albert von Szent-Gyorgyi (1893 - 1986)
Technical Challenge: Scalability

Scalability w.r.t.

- Amount of Data and Dimensionality
- Number of Data Sources and Heterogeneity
- Data Quality and Data Resolution
- Dynamicity and Novelty
- Data Representation and Visual Resolution
- User Interface and Interaction
- Display Devices

“All truths are easy to understand once they are discovered; the point is to discover them.”
~ Galileo Galilei (1564-1642)
Visual Analytics Agenda

The Science of Analytical Reasoning
“... enable users to obtain deep insights that directly support assessment, planning, and decision making. “

Visual Representations & Interaction Technologies
“... take advantage of human eye’s broad bandwidth pathway into the mind to allow users to see, explore, and understand large amount of information at once.“

Data Representations & Transformations
“... covert all types of conflicting and dynamic data in ways that support visualization and analysis.“

Production, Presentation, & Dissemination
“... communicate information in the appropriate context to a variety of audience.“
Application Areas

Economic & Business Data
  Business Intelligence
  Market Analysis

Medicine & Biotechnology
  Patients’ Data Management
  Epidemiology
  Genetics

Security & Risk Management
  Disaster Management
  Computer Networks
  Transportation
  Reducing Crime and Terror Rate
  Fraud Detection

Environment & Climate Research

etc.
Visual Steering to Support Decision Making in Visdom

Jürgen Waser
Flood emergency assistance

- New Orleans 2005: 17th canal levee breach

Image courtesy of USACE, US Army Corps of Engineers
Flood emergency assistance

- Testing sandbag configurations in a virtual environment
Solution: World Lines
Solution: World Lines
Video
Worldlines – Multiple Linked Views
SimVis: Interactive Visual Analysis of Large & Complex Simulation Data

Dr. Helmut Doleisch
VRVis Research Center

http://www.VRVis.at/
Motivation

- Large data sets from simulation
- Goal: support exploration and analysis of results
  - Analyze n-dim. data interactively
  - Use 3D visualization
  - Overview, zoom and filter, detail on demand (Shneiderman’s information seeking mantra)
- Challenge:
  - Occlusion
  - Interactive data handling

Helmut Doleisch
http://www.simvis.at/
SimVis: Interactive Visual Analysis of Large & Complex Simulation Data
Interactive Data Handling

- **sample data set size:**
  - 540 million data items
  - currently working to expand to billions

<table>
<thead>
<tr>
<th></th>
<th>cells</th>
<th>timesteps</th>
<th>attributes</th>
<th>cells * timesteps</th>
<th>cells * timesteps</th>
<th>attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>704.900</strong></td>
<td>20</td>
<td>16</td>
<td></td>
<td>14,098,000</td>
<td>225,568,000</td>
<td></td>
</tr>
<tr>
<td><strong>150.124</strong></td>
<td>600</td>
<td>6</td>
<td></td>
<td>90,074,400</td>
<td>540,446,400</td>
<td></td>
</tr>
<tr>
<td><strong>7.680.000</strong></td>
<td>288</td>
<td>15</td>
<td></td>
<td>2,211,840,000</td>
<td>33,177,600,000</td>
<td></td>
</tr>
</tbody>
</table>

Helmut Doleisch  
http://www.simvis.at/  
SimVis: Interactive Visual Analysis of Large & Complex Simulation Data
SimVis

- VRVis’ solution for these challenges
- Feature-based visualization framework

SimVis key features:
- Multiple, linked views
- Interactive feature specification
- Focus+Context visualization
- Smooth feature boundaries
- Explicit feature representation
- On-the-fly attribute derivation

Helmut Doleisch
http://www.simvis.at/
SimVis: Multiple Views

- Scatterplots, histogram, 3D(4D) view, etc.

Helmut Doleisch
http://www.simvis.at/

SimVis: Interactive Visual Analysis of Large & Complex Simulation Data
**Brushing**

- Move/alter/extend brush interactively
- Update linked F+C views in real-time

---

**30 SimVis: Interactive Visual Analysis of Large & Complex Simulation Data**

- **color: temp.**
- **vel.**
- **pressure**
- **TKE**

Helmut Doleisch
http://www.simvis.at/
Conclusions

- Visual Analytics is an opportunity worth considering
- Collaboration between academia, industry, national laboratories, and government (national and international) is key
- For each of you:

*The best is yet to come...*