CSE 564
Visualization & Visual Analytics

Interaction & Information Navigation

Klaus Mueller

Computer Science Department
Stony Brook University
<table>
<thead>
<tr>
<th>Lecture</th>
<th>Topic</th>
<th>Projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Intro, schedule, and logistics</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Applications of visual analytics, basic tasks, data types</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Introduction to D3, basic vis techniques for non-spatial data</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Data assimilation and preparation</td>
<td>Project #1 out</td>
</tr>
<tr>
<td>5</td>
<td>Data assimilation and preparation</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Bias in visualization</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Data reduction and dimension reduction</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Visual perception</td>
<td>Project #2(a) out</td>
</tr>
<tr>
<td>9</td>
<td>Visual cognition</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Visual design and aesthetics</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Cluster analysis: numerical data</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Cluster analysis: categorical data</td>
<td>Project #2(b) out</td>
</tr>
<tr>
<td>13</td>
<td>High-dimensional data visualization</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Dimensionality reduction and embedding methods</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Principles of interaction</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Midterm #1</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Visual analytics and the visual sense making process</td>
<td>Final project proposal call out</td>
</tr>
<tr>
<td>18</td>
<td>Evaluation and user studies</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Visualization of time-varying and time-series data</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Visualization of graph data</td>
<td>Final project proposal due</td>
</tr>
<tr>
<td>21</td>
<td>Visualization of text data</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Foundations of scientific and medical visualization</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Computer graphics and volume rendering</td>
<td>Project 3 out</td>
</tr>
<tr>
<td>24</td>
<td>Scientific and medical visualization</td>
<td>Final Project preliminary report due</td>
</tr>
<tr>
<td>25</td>
<td>Illustrative rendering</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Midterm #2</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Data journalism</td>
<td>Final Project slides and final report due</td>
</tr>
</tbody>
</table>
Too Much Data?

How can we deal with data overload

- see the forest for the trees (or the other way around)
Too Much Data?

Internet routes (1/15/2005)
(NY Museum of Modern Art)
The Key to Overcome the Data Deluge: Interaction

Allow users to control what is currently shown:

• level of detail
• extent of the data (spatial, values)
• aspects of the data (attributes)

But do not leave the user lost in the forest

• provide navigation hints

Two powerful paradigms:

• overview, and detail on demand (forest and trees)
• focus and context (trees and forest)

Interaction needs to be interactive (as in responsive)

• user needs get quick visual feedback on actions
Interaction: Key to Visual Analytics

Puts the human in the loop
  • appeals to human’s expertise and intuition

Requires a suitable human-computer interface
  • recall the lectures on color and perception

Interaction can help with:
  • making sense of it all
  • putting things in proper context
  • data overload (scalability)
  • telling stories with data (explain findings to others)

Evaluate effectiveness
  • do human users actually benefit?
  • user studies!
Stephen Few (chapter 4):

- compare
- sort
- add variables
- re-scale
- re-express
- filter
- highlight
- annotate
- bookmark
- aggregate
- re-visualize
- zoom and pan
- details on demand
Example

Assume you have been offered a car to buy

- assume you are mostly interested in horsepower, weight, acceleration
- the car you have been offered has 60 hp, 1834 kg, 8 s
See the car with other available cars

hard to see how it really ranks
See the car in the context of other available cars

it is a low-horsepower car
Is horsepower correlated to weight?
• are there trade-offs?

hard to see what is going on
Scale horsepower into the same range than weight
  • could also normalize each to (0.0, 1.0)

There seems to be a positive correlation
  • cars with higher horsepower are also heavier
Another Variable

How does it relate to acceleration?

non-intuitive that acceleration is less for high horsepower cars
Acceleration should really be 1/acceleration

- should be measured in 1/sec (and not sec)

• now higher horsepower cars also seem to have higher acceleration (but the influence is quite minor) → is there a higher-D relationship?
Example: Graph of concepts

- related concepts group closer

from: http://www.mkbergman.com/date/2008/02/
Example: Graph of concepts

- only keep top 750 connected nodes
Example: Graph of concepts

- only keep top 350 connected nodes
Zooming

Example: Graph of concepts
• only keep Saab neighborhood
Zooming

Example: Graph of concepts

- only keep Saab neighborhood, zoom in more
As discussed, good ways to aggregate all data into a single display are:

- biplots (project all data into a PCA vector basis)
- multidimensional Scaling (MDS)
- parallel coordinates
The Visual Information-Seeking Mantra

- devised 1996 by Ben Shneiderman (U Maryland, College Park)
- summarizes many visual design guidelines
- in some ways inspired by human vision/behavior
- provides an excellent framework for designing Information visualization applications.

Overview, zoom and filter, then details-on-demand

....
Overview and Detail

Information space overview plus some detail

• maintains (some) context with the detail currently focused on

Leica Microsystems

WikiViz
Focus + Context

Overview and detail
- disjoint views, maybe connected by a fan
- but: they simultaneously show both overview and details
- require the viewer to consciously shift his/her focus of attention

Focus + context
- one single view which shows information in direct context
- maintains continuity across the display
- do not require viewer to shift back and forth
- a good F+C paradigm is the lens
- but: there will be distortion
Fisheye Lenses

Fisheye lenses
- physically correct and therefore familiar
- keep target item in focus
- less relevant peripheral items are dropped or reduces in size
- distortion
Bifocal Lens

Complete Mapped Information Space

Principle of the Bifocal Display

Bifocal Display Seen by the User
Avoid aliasing in transition regions by low-pass filtering
Generalized Lenses

Computers can go beyond (stretch) the laws of physics

• example: multi-perspective lens rendering, gaze-directed, …

Rademacher/Bishop

MC Escher

Loeffelmann/Groeller
Multi Camera Rigs

Shoot a space-time video
Generalized Lenses

no lens

Transition region

Wang et al., 2005
Hyperbolic Tree fisheye lens

- Xerox PARC/Inxight
- selectively and smoothly reduce complexity as user focus changes
Lenses in Information Visualization

Table Lens (Rao and Card, 1994)

- uses a DOI (degree of interest) lens
- fuses symbolic and graphical detail driven by the DOI lens
Perspective Wall

- Xerox PARC/Inxight
- details on the center panel are at least three times larger than the details on a flat wall that fits the field of view
Illustrating the concept of a magic lens. (a) shows a conventional map of an area, (b) shows the location of services (gas, water and electricity pipes) in the same area, and (c) a (movable) magic lens shows services in an area of interest, in context.

[Video]
Panning
  • smooth movement of a viewing frame over a 2D image of greater size

• Zooming
  • increasing magnification of a decreasing fraction (or vice-versa) of a 2D image under the constraint of a viewing frame of constant size

Transfer of the focus of attention:
  • zoom out → pan → zoom in
Efficient transfer of the focus of attention:

- zoom out → pan → zoom in

Furnas, Bederson, 1995
Intelligent Zooming

Depending on scrolling speed, zoom more or less

- allows efficient navigation of large documents
- employs semantic zooming

Igarashi, Hinckley, 2000
Semantic Zoom

Standard zoom:

- shows a down/up scaled version of the object/image

Semantic zoom:

- shows a different representation determined by the space available
Semantic Zooms: Maps
Could show different levels/aspects of information

- on a map, show either parking lots, bars, or restaurants
- zoom in by price range (cheap first, then more expensive…)
- zoom in by preference (favorite food first, then less favorite…)
- may combine these criteria into a preference function
Smooth Semantic Zooms with Texture Synthesis

Zooming strategy:

- magnify low resolution level image
- synthesize new detail at transition point
- minify synthesized next high level image
- weighted blending

MRI level

histology level

cell level

Transition point

Present image

weight = 1

level $i$

weight = 0

Newly synthesized image

Level $i+1$
Smooth Semantic Zooms with Texture Synthesis: Results

Plain zoom

Semantic Zoom

3D
Semantic Zooms: Information Visualization

Could show different levels/aspects of information

- on a map, show either parking lots, bars, or restaurants
- zoom in by price range (cheap first, then more expensive…)
- zoom in by preference (favorite food first, then less favorite…)
- may combine these criteria into a preference function

Zoom levels may require access rights

- members only
- big wallets only
- classified information

Human being → Male → Police man → British policeman → Inspector Blanding

information access rights
Exploded Views

Video
An Exploded View Paradigm to Disambiguate Scatterplots
Interactive technique

- Highlighting
- Brushing and Linking

At least two things must be linked together to allow for brushing

- select a subset of points
- see the role played by this subset of points in one or more other views

Example systems

- Graham Will’s EDV system
- Ahlberg & Sheiderman’s IVEE (Spotfire)
Linking Types of Assist Behavior to Position Played
Baseball Data: Scatterplots and Histograms and Bars

- Select high salaries
- Avg. career home runs (y) vs. avg. career hits (x) (batting ability)
- Avg. assists (x) vs. avg. putouts (y) (fielding ability)
- How long in majors
- Distribution of positions played
- Anything interesting?
What was Learned from Interaction w/ the Baseball Data?

- Seems impossible to earn a high salary in the first three years
- High salaried players have a bimodal distribution (peaking around 7 & 13 yrs)
- Hits/Year a better indicator of salary than HR/Year
- High paid outlier with low HR and medium hits/year. Reason: person is player-coach
- There seem to be two differentiated groups in the put-outs/assists category (but not correlated with salary) Why?
Brushing: Highlighting

Use mouse interaction to highlight points and lines in
• parallel coordinates
• scatterplots
Interaction in Parallel Coordinate
Accelerate Understanding

Dashboards should pass the 5-second test
Important rules:

- most important view goes on top or top left
- legends go near their views
- avoid using multiple color schemes
- use 5 views or fewer
- provide interactivity