Visualizing Multi-dimensional Data

Seth Horrigan

Computer Visualization
Fall 2008
Motivation

- Multi-dimensional datasets are common
  - Digital cameras
  - Wall-street stocks
  - Motor vehicles
  - Cellular telephones

- A mixture of interval, ordinal, and nominal data can be visualized well using a table
## Motivation

<table>
<thead>
<tr>
<th>Features</th>
<th>See What's the Same or What's different</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lens aperture</strong></td>
<td>F/3.0-5.8</td>
</tr>
<tr>
<td><strong>Equivalent 35mm focal length</strong></td>
<td>38 - 132 mm</td>
</tr>
<tr>
<td><strong>Optical zoom</strong></td>
<td>3.4 x</td>
</tr>
<tr>
<td><strong>Display (projector) technology</strong></td>
<td>TFT active matrix</td>
</tr>
<tr>
<td><strong>Memory storage capacity</strong></td>
<td>32 MB</td>
</tr>
<tr>
<td><strong>Video input type</strong></td>
<td>Digital camera</td>
</tr>
<tr>
<td><strong>Flash memory form factor</strong></td>
<td>SD Memory Card</td>
</tr>
<tr>
<td><strong>Digital zoom</strong></td>
<td>4 x</td>
</tr>
<tr>
<td><strong>Battery technology</strong></td>
<td>Alkaline</td>
</tr>
<tr>
<td><strong>Weight</strong></td>
<td>5.8 oz</td>
</tr>
<tr>
<td><strong>Effective sensor</strong></td>
<td>7,100,000 pixels</td>
</tr>
</tbody>
</table>
Motivation

• Questionnaire surveys produce special dataset
  ○ Interval
  ○ Ordinal

• Usually compared only within-variable
Motivation

5. In the past 2 weeks, how many times have you viewed a previous day's lab material - this includes both while in the lab and from elsewhere.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4</td>
<td>4%</td>
</tr>
<tr>
<td>1-2</td>
<td>33</td>
<td>33%</td>
</tr>
<tr>
<td>3-4</td>
<td>35</td>
<td>35%</td>
</tr>
<tr>
<td>4+</td>
<td>27</td>
<td>27%</td>
</tr>
<tr>
<td>Total</td>
<td>99</td>
<td>100%</td>
</tr>
</tbody>
</table>

6. Have you ever experienced any of the following problems while submitting an assignment/homework? (Check all that apply)

- Forgetting the process for submission: 27 (27%)
- Forgetting to submit an assignment: 27 (27%)
- Being unsure whether you submitted an assignment: 51 (52%)
- Being concerned whether you submitted the correct file(s): 50 (59%)
- I had none of these difficulties: 20 (20%)
- Other, please specify: 4 (4%)

7. In lab, would you prefer to discuss questions with your classmates in person or in an electronic forum?

- In Person: 68 (69%)
- Electronically: 31 (31%)

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>In Person</td>
<td>68</td>
<td>69%</td>
</tr>
<tr>
<td>Electronically</td>
<td>31</td>
<td>31%</td>
</tr>
<tr>
<td>Total</td>
<td>99</td>
<td>100%</td>
</tr>
</tbody>
</table>

8. If you could choose to add only one of the following features, which would it be?

- Indicator for progress through a day's material: 15 (15%)
- Search through material from all labs: 87 (88%)
Motivation

• Need a good way to relate variables to each other

• Need a good way to visualize multiple ordinal variables
Shortcomings

- Ordinal variables are usually graphed against interval variables
Shortcomings

- Graphing ordinal against ordinal does not work well
Shortcomings

- Regression lines only help a little
Shortcomings

- Summing helps, but really encodes different data
Initial Concept

- Introduce random jitter
Initial Concept

- Multi-dimensional matrix
- Allow continuous rotation from viewpoint to viewpoint
Initial Concept

- ScatterDice

Previous Work

- **Geometrically transformed displays**

- **Iconic displays**

- **Dense pixel displays**

- **Dimensional stacked displays**
Previous Work

- **Overview of methods**
  

- **Encoding variables**
  
Current Concept

- Color + Position + Size + Small multiples
Current Concept

- Much more difficult to interpret how the individual data points aggregate to the whole
- Allow many dimensions of data to be visualized using position

- Also considering how to specifically enhance a scatterplot to convey the data
Technical Challenges

- Fitting many variables into small space
  - Design through prototyping
- Determining data encoding (colors? texture?)
  - Reference previous research, experimentation
- Maintaining part-to-whole relationships
  - With each design, record what information is conveyed or lost
- Building prototype
  - Use existing knowledge and work in Prefuse and Flare
Milestones

- 10/31 - 5 other solution concepts
- 11/5 - Determine how, if at all to include interval data
- 11/5 - Create storyboards
- 11/10 - Determine color or other encoding
  - Create legend
- 11/20 - Build automatic optimal layout
- 11/25 - Design and build interaction
- 12/1 - Build Attribute-explorer style filters
- 12/10 - Create final presentation and paper