Multivariate Visual Representations

CS 4460/7450 - Information Visualization
Jan. 28, 2010
John Stasko

Agenda

• General representation techniques for multivariate (>3) variables per data case
  – But not lots of variables yet...
Quick Quiz

How Many Variables?

- Data sets of dimensions 1, 2, 3 are common
- Number of variables per class
  - 1 - Univariate data
  - 2 - Bivariate data
  - 3 - Trivariate data
  - >3 - Hypervariate data  Focus Today
Earlier

- We examined a number of tried-and-true techniques/visualizations for presenting multivariate (typically $\leq 3$) data sets
  - Hinted at how to go above 3 dimensions

Representations

Some standard ways for low-d data
Hypervariate Data

• How about 4 to 20 or so variables (for instance)?
  – Lower-dimensional hypervariate data
  – (Much higher dimensions next week)
  – Many data sets fall into this category

More Dimensions

• Fundamentally, we have 2 geometric (position) display dimensions
• For data sets with >2 variables, we must project data down to 2D
• Come up with visual mapping that locates each dimension into 2D plane

• Computer graphics: 3D->2D projections
Wait a Second

- A spreadsheet already does that
  - Each variable is positioned into a column
  - Data cases in rows
  - This is a projection (mapping)

- What about some other techniques?
  - Already seen a couple

Multiple Views

Give each variable its own display

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>1</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>
Scatterplot Matrix

Represent each possible pair of variables in their own 2-D scatterplot

Chernoff Faces

Encode different variables’ values in characteristics of human face
Examples

Cute applets:  
http://www.cs.uchicago.edu/~wiseman/chernoff/  

Table Lens

- Spreadsheet is certainly one hypervariate data presentation
- Idea: Make the text more visual and symbolic
- Just leverage basic bar chart idea

Rao & Card
CHI ’94
Visual Mapping

Change quantitative values to bars

Tricky Part

What do you do for nominal data?
Instantiation

Focus on item(s) while showing the context
See It


Video

FOCUS

- Feature-Oriented Catalog User Interface
- Leverages spreadsheet metaphor again
- Items in columns, attributes in rows
- Uses bars and other representations for attribute values

Spenke, Beilken, & Berlage
UIST '96
Charactersitics

- Can sort on any attribute (row)
- Focus on an attribute value (show only cases having that value) by double-clicking on it
- Can type in queries on different attributes to limit what is presented too
Limit by Query

Figure 4: A disjunction.

Manifestation

InfoZoom

Commercial product to be demo'ed coming up
Categorical data?

• How about multivariate categorical data?

• Students
  – Gender: Female, male
  – Eye color: Brown, blue, green, hazel
  – Hair color: Black, red, brown, blonde, gray
  – Home country: USA, China, Italy, India, ...
Mosaic Plot

Spring 2010

Mosaic Plot

Women

Men

Spring 2010

Brown
Hazel
Green
Blue

Women

Men
Mosaic Plot

Intermission

- Readings reactions
- HW 1 recap
- Grade percentages
Attr Explorer

• Discuss
• What is at the heart of the technique?

Characteristics

• Multiple histogram views, one per attribute (like trellis)
• Each data case represented by a square
• Square is positioned relative to that case’s value on that attribute
• Selecting case in one view lights it up in others
• Query sliders for narrowing
• Use shading to indicate level of query match (darkest for full match)
Features

- Attribute histogram
- All objects on all attribute scales
- Interaction with attributes limits

Features

- Inter-relations between attributes – brushing
Features

• Color-encoded sensitivity

Attribute Explorer

Video

Summary

• Summary
  – Attribute histogram
  – Attribute relationship
  – Sensitivity information
  – Especially useful in "zero-hits" situations or when you are not familiar with the data at all

• Limitations
  – Limits on the number of attributes

MultiNav

• Each different attribute is placed in a different row
• Sort the values of each row
  – Thus, a particular item is not just in one column
• Want to support browsing

Lanning et al
AVI '00
Interface

Alternate UI

- Can slide the values in a row horizontally
- A particular data case then can be lined up in one column, but the rows are pushed unequally left and right
Attributes as Sliding Rods

Information-Seeking Dialog
Instantiation

Limitations

- Number of cases (horizontal space)
- Nominal & textual attributes don’t work quite as well
Dust & Magnet

- Altogether different metaphor
- Data cases represented as small bits of iron dust
- Different attributes given physical manifestation as magnets
- Interact with objects to explore data

Interface
Interaction

- Iron bits (data) are drawn toward magnets (attributes) proportional to that data element’s value in that attribute
  - Higher values attracted more strongly
- All magnets present on display affect position of all dust
- Individual power of magnets can be changed
- Dust’s color and size can connected to attributes as well

Interaction

- Moving a magnet makes all the dust move
  - Also command for shaking dust
- Different strategies for how to position magnets in order to explore the data
Parallel Coordinates

- Perhaps the best known technique for this kind of data

- Our focus next time...
Administratia

- HW 2 in next Thursday

Upcoming

- Parallel coordinates & more
  - Reading:
    Inselberg paper

- Big Data
  - Reading:
    J. Yang et al paper