Information Visualization
Multivariate Data Visualization

Krešimir Matković
Vienna University of Technology,
VRVis Research Center, Vienna
Multidimensional Multivariate Data

Use of Color & Shape to increase dimensionality

Table Lens

Scatter plot matrix

Chernoff faces

Parallel Coordinates

Star plot

Star coordinates

Parallel Sets

Mosaic Plot

Interaction and Coordinated Multiple Views, Example Application
Multidimensional Multivariate Data

Conventional approach deals with n-dim. Euclidian spaces

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

Each dimension can be categorical (nominal, ordinal) or numerical

Data example, meteorological stations in California

\[ P_i = [x_1, \ldots, x_i, \ldots, x_n] \]

<table>
<thead>
<tr>
<th>STATION</th>
<th>AVERAGE TEMP</th>
<th>PRESSURE</th>
<th>ELEVATION</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>1015</td>
<td>200</td>
<td>...</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>1020</td>
<td>300</td>
<td>...</td>
</tr>
</tbody>
</table>
Visualizing single dimensions

Each dimension (column in the table) can be visualized as 1D data

Histogram, Bar Chart, Box-Plot
Visualizing pairs of dimensions

Pairs of dimensions can be visualized as 2D data

Scatter plot

- Widely used
- Shows correlation – positive, negative, no correlation
- Shows clusters
- Multiple points – scatter, increase size
- Can be used to show up to 5 dimensions (color, size, shape)
Visualizing pairs of dimensions

Krešimir Matković
InfoVis, TU Wien, 20. 03. 2014
Visualizing many dimensions

There are 2 geometric dimensions on screen

For data sets with >2 dimensions, we must project data down to 2D

Come up with visual mapping that locates each dimension into 2D plane

A spreadsheet already does that

- Column -> dimension
- Row -> record, case, item
Table Lens

Textual table – too large

Rao and Card 1994
Table Lens

<table>
<thead>
<tr>
<th>Cereal</th>
<th>Color</th>
<th>Type</th>
<th>Calories</th>
<th>Protein</th>
<th>Fat</th>
<th>Sodium</th>
<th>Fiber</th>
<th>Carbohydrates</th>
<th>Sugars</th>
<th>Shelf</th>
<th>Potassium</th>
<th>Vitamins</th>
<th>Weight</th>
<th>Cups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fruit &amp; Berry</td>
<td>K</td>
<td>C</td>
<td>120</td>
<td>3</td>
<td>6.5</td>
<td>200</td>
<td>5</td>
<td>14</td>
<td>12</td>
<td>3</td>
<td>140</td>
<td>25</td>
<td>1.3</td>
<td>0.35</td>
</tr>
<tr>
<td>Just Right Cereal</td>
<td>K</td>
<td>C</td>
<td>110</td>
<td>2</td>
<td>4.5</td>
<td>170</td>
<td>1</td>
<td>17</td>
<td>6</td>
<td>2</td>
<td>60</td>
<td>100</td>
<td>1.1</td>
<td>0.75</td>
</tr>
<tr>
<td>Just Right Cereal</td>
<td>K</td>
<td>O</td>
<td>140</td>
<td>3</td>
<td>7.5</td>
<td>170</td>
<td>2</td>
<td>20</td>
<td>9</td>
<td>3</td>
<td>100</td>
<td>13</td>
<td>0.75</td>
<td></td>
</tr>
</tbody>
</table>
(1/6) Distortion-oriented

Table lens [Rao/Card 1994]

Initial view, no details

Krešimir Matković

InfoVis, TU Wien, 20. 03. 2014
(1/6) Distortion

Table lens [Rao/Card 1994]

Sort by "1YR"

Mutual Fund Performance - February 1999

- Click anywhere in the window to open a focus and see the details

Krešimir Matković
InfoVis, TU Wien, 20. 03. 2014
Distortion - oriented F+C vis.

Table lens [Rao/Card 1994]

Mark good „1YR“
Sort by „3YR“

Krešimir Matković

InfoVis, TU Wien, 20. 03. 2014
(1/6) Distortion

Table lens [Rao/Carlson 1994]

Mark good "3YR"
Sort by "5YR"

Mutual Fund Performance - February 1999

- Click anywhere in the window to open a focus and see the details

Krešimir Matković
InfoVis, TU Wien, 20.03.2014
## Table Lens [Rao/Card 1994]

Click anywhere in the window to open a focus and see the details.

### Mutual Fund Performance - February 1999

<table>
<thead>
<tr>
<th>Name</th>
<th>1YR</th>
<th>3YR</th>
<th>5YR</th>
<th>10YR</th>
<th>Yld</th>
<th>Ex</th>
<th>Mg</th>
<th>Ne</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jesse Barfield</td>
<td>48.04</td>
<td>41.05</td>
<td>32.0099</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jeffrey Leonar</td>
<td>33.50939</td>
<td>31.68</td>
<td>30.0699</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dornie Hill</td>
<td>23.39</td>
<td>42.2</td>
<td>29.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Billy Sample</td>
<td>35.24999</td>
<td>33.1</td>
<td>28.08999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Howard Johnson</td>
<td>54.12</td>
<td>37.7</td>
<td>27.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andret Thomas</td>
<td>42.20999</td>
<td>33.88999</td>
<td>27.78999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Billy Hatcher</td>
<td>47.93999</td>
<td>30.14</td>
<td>27.53999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Omer Korero</td>
<td>46.65</td>
<td>32.79</td>
<td>26.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cancelle Rolles</td>
<td>388</td>
<td>88.47999</td>
<td>48.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>370</td>
<td>49.4</td>
<td>42.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>371</td>
<td>33.85</td>
<td>39.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>372</td>
<td>48.01</td>
<td>35.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>183.86</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>601</td>
<td>104.22</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Krešimir Matković  
InfoVis, TU Wien, 20. 03. 2014
(1/6) Distortion

Table lens [Rao/Card]

Open „Category“

Mutual Fund Performance - February 1999

- Click anywhere in the window to open a focus and see the details

<table>
<thead>
<tr>
<th>No.</th>
<th>Category</th>
<th>YTD</th>
<th>1YR</th>
<th>3YR</th>
<th>SYR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Large Value</td>
<td>48.04</td>
<td>41.03</td>
<td>32.00999</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Large Grow</td>
<td>39.50369</td>
<td>31.88</td>
<td>30.08999</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Large Grow</td>
<td>73.39</td>
<td>42.2</td>
<td>29.58</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Large Value</td>
<td>35.24999</td>
<td>35.1</td>
<td>28.06999</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Large Grow</td>
<td>54.12</td>
<td>87.47999</td>
<td>27.38</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Large Grow</td>
<td>42.20999</td>
<td>35.86999</td>
<td>27.78999</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Large Grow</td>
<td>47.93399</td>
<td>30.14</td>
<td>27.53999</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Large Grow</td>
<td>48.85</td>
<td>32.73</td>
<td>26.7</td>
<td></td>
</tr>
</tbody>
</table>

Krešimir Matković

InfoVis, TU Wien, 20. 03. 2014
Multiplies

Idea: use multiple single visualization

- many bar charts, e.g.

Scatterplot matrix

- Show (all possible) combinations of dimensions as a scatterplot matrix
- Good overview of corellation for all depicted dimensions
- Often used in statistics
- Too small for many dimensions
- Zoom in usually possible
Scatter plot matrix

[http://junkcharts.typepad.com/junk_charts/]

Krešimir Matković  InfoVis, TU Wien, 20. 03. 2014
Chernoff Faces 1

We can easily distinguish various faces

Encode different variables in characteristics of human face

Decreasing order of perception

- Area of face
- Shape of face
- Length of nose
- Location of mouth
- Curve of smile
- Width of mouth
- Location, separation, angle, shape and width of eyes
- Location of pupil
- Location, angle, width of eyebrows
Chernoff Faces 2
Chernoff Faces Example 1

Figure 5. Six facial variations

Figure 6. A Chernoff face for the state of Ohio

[Joseph G. Spinelli and Yu Zhou]
Chernoff Faces Example 1

Figure 7. Quality of life map with Chernoff faces

[Joseph G. Spinelli and Yu Zhou]
Chernoff Faces Example 2

Life in Los Angeles
Chernoff Faces Example 3

Comments?

2005 National League: Chernoff Faces

Arizona  Atlanta  Chicago  Cincinnati
Colorado  Florida  Houston  Los Angeles
Milwaukee  New York  Philadelphia  Pittsburgh
San Diego  San Francisco  St. Louis  Washington
Parallel Coordinates

Parallel 2D axes

Add/Remove data
- Establish Patterns
- Examine interactions

Useful for recognizing patterns between the axes

Skilled user
Parallel Coordinates 2

Encode variables along a horizontal row
Vertical line specifies single variable
Blue line specifies a case

Dataset in a Cartesian graph
Same dataset in parallel coordinates
PC – Line Point Duality
PC – Line Point Duality
PC – 5D Sphere!
Parallel Coordinates - Problems

If many records – overlapping and clutter

- Reorder dimensions
- Density
- Clustering
- Outlier preserving methods

Short demo
Parallel Coordinates – Outlier Preserving

Krešimir Matković
InfoVis, TU Wien, 20.03.2014
Extended Parallel Coordinates

- Greyscale
- Color
- Histogram information on axes
- Smooth brushing
- Angular brushing

Krešimir Matković
InfoVis, TU Wien, 20. 03. 2014
Star Plot, Radar Chart

Figure 3: Star plot of automobile data. Each star represents models (bottom two rows) are shown.
Star Coordinates

Rather than represent a point as a shape, just accumulate values along a vector parallel to the coordinate axes.

Figure 1. Calculation of data point location for an 8-dimensional dataset.
Parallel Sets

“Parallel Coordinates for categorical data”

- Axis replaced with boxes
- Layout similar to PC
Parallel Sets 2

Which class has the highest number of survivors?
Which class has the highest number of survivors?

<table>
<thead>
<tr>
<th>Crew</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survivors</td>
<td>212</td>
<td>203</td>
<td>118</td>
</tr>
<tr>
<td>Non-Survivors</td>
<td>673</td>
<td>122</td>
<td>167</td>
</tr>
</tbody>
</table>
Linewidth Illusion

Playfair’s chart from the Commercial and Political Atlas (1786) showing the balance of trade between England and the East Indies.
Linewidth Illusion

Where is the difference highest?
Common Angle Plot – a Possible Solution

Heike Hofmann and Marie Vendettuoli, TVCG 2013
Mosaic Plot

Categorical data

- Not-suitable views: scatter plot, PC, Box Plot

Example data: Titanic dataset

Mosaic Plot

- Relationship between 2 or more categorical dimensions
- Start with a square of length 1
- Divide it according to one category horizontally
- Divide each sub-block vertically according to another category....
Mosaic Plot

[Steve Simon]
Mosaic Plot

http://www.perceptualedge.com/articles/visual_business_intelligence/are_mosaic_plots_worthwhile.pdf
Dense Pixel Methods - Basic Idea

Use pixel to represent one data item

- Number of pixels (~1M) determines number of items
- Relies on use of color - color mapping
- Value ranges are mapped to a fixed color sequence of full color (hue) scale but monotonically decreasing brightness
Horizontal arrangement

8 horizontal bars correspond to 8 years

Subdivisions between the bars represent 12 months within each year

Example analysis results

- Gold price was very low in the sixth year
- IBM price fell quickly after the first 1½ month
- US-Dollar exchange rate was highest in the third year

[Keim et al. 1995]
Query Dependent Arrangement
Pixel Bar Chart

Overload typical bar chart with more information about individual elements

[Keim et al.]
Pixel Bar Chart 1

Overload typical bar chart with more information about individual elements

[Keim et al.]
Pixel Bar Chart 2

Make each pixel within a bar correspond to a data point in that group represented by the bar.

Can do millions that way.

Color the pixel to represent the value of one of the data point’s variables.

[Keim et al.]
Pixel Bar Chart 2

Product type is x-axis divider

Customers ordered by

- y-axis: dollar amount
- x-axis: number of visits

Color is (a) dollar amount spent, (b) number of visits, (c) sales quantity

[Keim et al.]
Pixel Bar Chart 3

Figure 3 Pixel bar charts

Figure 6 Dividing attribute on x-axis (e.g., $D_x = \text{Product Type}$)

Figure 7 Dividing attributes on x- and y-axis (e.g., $D_x = \text{Product Type}$, $D_y = \text{Region}$)

Figure 8 Ordering attributes on x- and y-axis (e.g., $O_x = \text{Dollar Amount}$, $O_y = \text{Quantity}$)

[Keim et al.]
Pixel Bar Chart 4

Mapping specified by 5 tuple
<\(D_x, D_y, O_x, O_y, C\)>

- \(D_x\) – Attribute partitions x axis
- \(D_y\) – Attribute partitions y axis
- \(O_x\) – Attribute specifies x ordering
- \(O_y\) – Attribute specifies y ordering
- \(C\) – Attribute specifies color mapping

[Keim et al.]
Pixel Bar Chart 5

[Keim et al.]
Pixel Bar Chart – CFD Data

Velocity: 0.466674
Temperature: 318.829
RelativePressure: 60626.9

[Martin Gasser
Hannes Kiraly.]
Color Lines View 1

Pixel based technique for families of curves

- Interaction!
Color Lines View 2

Demo!

(a) The color lines view is sorted by January temperatures. This is indicated by the blue rectangle in the bottom left corner. The colors are remarkably varied in the middle, which means there is little direct correlation between January and summer temperatures.

(b) The color lines view is sorted by elevation. Cold winters correspond to high elevations, but hot summers do not correspond to low elevation.
InfoVis Techniques

D. Keim proposes a taxonomy of techniques

Standard 2D/3D display
- Bar charts, scatterplots

Geometrically transformed display
- Parallel coordinates

Iconic display
- Needle icons, Chernoff faces

Dense pixel display
- Pixel Bar Chart, Color Lines

Stacked display
- Treemaps, dimensional stacking
Focus and Context Visualization
Basic Problem

Many data sets are too large to visualize on one screen

- Too many cases
- Too many variables
- Highlight particular cases or particular variables, but viewer’s focus may change from time to time

How to work with, navigate through, and analyze a set of data that is too large to fit in the display?

Potential solutions lie in

- Representation
- Interaction
- Both

[Stasko Course Notes]
Overview

Overview first, zoom-in, details on demand!

Overview

- Starting point
- Very important
- Helps present overall patterns
- Needed for navigation

Krešimir Matković
InfoVis, TU Wien, 20. 03. 2014
Details

Overview first, zoom-in, details on demand!

Details

- Generally provided on demand
- Individual cases and variables
- How to allow user to find and focus on details of interest?
Pan and Zoom

Pan/Scroll
- Provide a larger, virtual screen by allowing user to move to different areas
- Still a problem
- Clunky interaction
- Only get to see one piece

Zoom
- Zoom out shows an overview of data space
- Zoom in allows viewer to examine details
- Getting lost - zoom in or out way too far can’t see anything

Long pan isn’t good - zoom out, pan a little, zoom in
Focus and Context 1

Combine overview + detail

- Via space - use different portions of screen to show overview and details
- Via time - alternate between overview and details sequentially in same place

No universal solution, many possibilities

Focus+context visualization: integration of (zoomed) focus & context within one view
Various classification possible

Distorsion Oriented F+C
Views-/layers-based F+C
InPlace F+C
Focus + Context

Various classification possible

Distortion Oriented F+C
Views-/layers-based F+C
InPlace F+C
Distortion-oriented F+C vis. 2

Hyperbolic tree view
[Lamping/Rao 1994]

3D hyperbolic space
[Munzner 1995]
Distortion-oriented F+C vis. 3

Generalized detail-in-context [Keahey 1998]
Fisheye views

Distortion oriented technique

Introduced by George Furnas in 1981 report, more famous article is 1986 SIGCHI paper

Definition:

“Provide[s] detailed views (focus) and overviews (context) without obscuring anything... The focus area (or areas) is magnified to show detail, while preserving the context, all in a single display.”

(Shneiderman, DTUI, 1998)
Fisheye views
Fisheye views terminology

Focal point
Level of detail
Distance from focus
Degree of interest function
Fisheye views terminology

Focal point

Level of detail

Distance from focus

Degree of interest function
Fisheye views Focal Point

Assume that viewers focus is on some item, some coordinate, some position,...
Fisheye views Level of Detail

Some intrinsic value or quantity on each data element

How important is it to you in a general sense?

Simplest example is that all data items have same level of detail
Fisheye views Distance from Focus

Calculation of how far each data item is from the focal point

Query position

Focal Point, (0.8, 0.1)

(1.0, 1.0)

(0.0, 0.0)
Fisheye views Degree of Interest Function

“DOI” - Function that determines how items in display are rendered

- DOI = Level of Detail - Distance from Focus

Level of Detail / Distance from Focus

Focal Point, (0.8, 0.1)

(0.0, 0.0)
Fisheye views Degree of Interest Function

DoI Function

Can take on various forms

Continuous - Smooth interpolation away from focus

Filtering - Past a certain point, objects disappear

Step - Levels or regions dictating rendering $0<x<0.3$
  all same, $0.3<x<0.6$ all same

Semantic changes - Objects change rendering at different levels
Distortion-oriented F+C vis. 4

Perspective wall
[Mackinlay et al. 1991]

Document lens
[Robertson/Mackinlay 1993]
Distortion-oriented F+C vis. 5

Table lens [Rao/Card 1994]
Distortion-oriented F+C vis. 6

F+C Process Vis.  
[Matkovic et al. 2002]

Focus+Context Process Visualization

Using LODs for F+C  
3D Anchoring  
Collision Avoidance  
Focus+Context Rendering
Focus + Context

Various classification possible

Distortion Oriented F+C

Views-/layers-based F+C

InPlace F+C
Views-/layers-based F+C vis. 1

Information Mural
[Jerding/Stasko 1998]

SeeSoft [Eick et al. 1992]
Views-/layers-based F+C vis. 2

Toolglass & Magic Lens [Bier et al. 1993]

[Fuhrmann et al. 1998]
Views-/layers-based F+C vis. 3

Macroscope [Lieberman 1994]
Focus + Context

Various classification possible

Distorsion Oriented F+C
Views-/layers-based F+C
InPlace F+C
In-place F+C vis. 1

GeoSpace [Lokuge/Ishizaki 1995]
In-place F+C vis. 2

Semantic Depth of Field (SDOF)  
[Kosara et al. 2001]

With this once lang lalce, now wildly bowwed, 
fifty yare a age did Nathan Swain kill seven, ahove the sunse a sunrise and a sunset.
In-place F+C vis. 3

F+C Screen [Baudisch et al. 2001]
### Generalized F+C Visualization

Focus+context visualization: **uneven use of graphics resources for visualization**

<table>
<thead>
<tr>
<th>resource space</th>
<th>approach</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>frequ.</td>
<td>focus enlargement</td>
<td>fisheye, ProcVis</td>
</tr>
<tr>
<td>opacity style</td>
<td>smooth context</td>
<td>SDOF</td>
</tr>
<tr>
<td>color sat.</td>
<td>translucent context NPR context</td>
<td>DVR, RTVR, 2ÍVR</td>
</tr>
<tr>
<td>col.+α</td>
<td>focus highlighting</td>
<td>GeoSpace, Voxelstube</td>
</tr>
<tr>
<td></td>
<td>glyph rendering</td>
<td>SimVis</td>
</tr>
</tbody>
</table>

Krešimir Matković  InfoVis, TU Wien, 20. 03. 2014
F+C visualization – comments

Very useful, often used (esp. in InfoVis)

Usually distortion-oriented, but more general concept (cf. generalization)

Requires interaction (focussing, navigation), smooth changes due to interaction

Aligns well with information drill-down, data analysis (visual information seeking mantra)

Also useable for presentation (e.g., SDOF)

Krešimir Matković  InfoVis, TU Wien, 20. 03. 2014
Thank you!

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

Krešimir Matković  
InfoVis, TU Wien, 20. 03. 2014