Using Space Effectively: 2D

Maneesh Agrawala
Jessica Hullman

CS 294-10: Visualization
Fall 2014

Last Time: Color
## Color Models

**Physical World**
- Light Energy
- Spectral distribution functions $F(\lambda)$

**Visual System**
- Cone Response
- Encode as three values $(L,M,S)$
- CIE $(X,Y,Z)$

**Mental Models**
- Opponent Encoding
  - Separate lightness, chroma $(A,R-G,Y-B)$
- Perceptual Models
  - Color “Space”
  - Hue lightness saturation
  - CIELAB
  - Munsell $(HVC)$

### Computing Cone Response

**Input Stimulus** $\Phi(\lambda)$

**Cone Response Curves**

**Product**

**Response**

\[
L = \int \Phi(\lambda)L(\lambda)\,d\lambda \\
M = \int \Phi(\lambda)M(\lambda)\,d\lambda \\
S = \int \Phi(\lambda)S(\lambda)\,d\lambda
\]
CIE LAB and LUV color spaces

Standardized in 1976 to mathematically represent opponent processing theory
Non-linear transformation of CIE XYZ

Psuedo-Perceptual Models

HLS, HSV, HSB
NOT perceptual models
Simple renotation of RGB
- View along gray axis
- See a hue hexagon
- L or V is grayscale pixel value
Cannot predict perceived lightness
Evolution of Basic Color Terms

Proposed universal evolution across languages

Using Color in Visualization
Superficial dissection of the right side of the neck, showing the carotid and subclavian arteries
http://www.bartleby.com/107/illus520.html
Palette Design + Color Names

Minimize overlap and ambiguity of color names

<table>
<thead>
<tr>
<th>Color Name Distance</th>
<th>Salience</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.47</td>
<td>blue 50.5%</td>
</tr>
<tr>
<td>1.00</td>
<td>0.30</td>
<td>orange 93.9%</td>
</tr>
<tr>
<td>1.00</td>
<td>0.39</td>
<td>green 79.8%</td>
</tr>
<tr>
<td>0.98</td>
<td>0.66</td>
<td>red 90.4%</td>
</tr>
<tr>
<td>1.00</td>
<td>0.47</td>
<td>purple 51.4%</td>
</tr>
<tr>
<td>1.00</td>
<td>0.37</td>
<td>brown 54.0%</td>
</tr>
<tr>
<td>0.98</td>
<td>0.58</td>
<td>pink 71.7%</td>
</tr>
<tr>
<td>1.00</td>
<td>0.97</td>
<td>grey 79.4%</td>
</tr>
<tr>
<td>1.00</td>
<td>0.18</td>
<td>yellow 31.2%</td>
</tr>
<tr>
<td>0.98</td>
<td>0.25</td>
<td>blue 25.4%</td>
</tr>
</tbody>
</table>

Tableau-10

Average 0.97

http://vis.stanford.edu/color-names

Palette Design + Color Names

Minimize overlap and ambiguity of color names

<table>
<thead>
<tr>
<th>Color Name Distance</th>
<th>Salience</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.30</td>
<td>blue 50.5%</td>
</tr>
<tr>
<td>1.00</td>
<td>0.21</td>
<td>red 27.3%</td>
</tr>
<tr>
<td>1.00</td>
<td>0.34</td>
<td>green 36.8%</td>
</tr>
<tr>
<td>1.00</td>
<td>0.55</td>
<td>purple 67.3%</td>
</tr>
<tr>
<td>0.07</td>
<td>0.20</td>
<td>blue 36.6%</td>
</tr>
<tr>
<td>0.00</td>
<td>0.39</td>
<td>orange 51.0%</td>
</tr>
<tr>
<td>0.35</td>
<td>0.13</td>
<td>blue 15.7%</td>
</tr>
<tr>
<td>0.35</td>
<td>0.12</td>
<td>green 21.7%</td>
</tr>
</tbody>
</table>

Excel-10

Average 0.87

http://vis.stanford.edu/color-names
Mapping Data to Color
Scale

Default rainbow maps
Avoid rainbow color maps!

1. People segment colors into classes
2. Hues are not naturally ordered
3. Different lightness emphasizes certain scalar values
4. Low luminance colors (blue) hide high frequencies

Phase Diagrams (hue scale)

Singularities occur where all colors meet

The optical singularities of bianisotropic crystals, by M. V. Berry
Phases of the Tides

Figure 1.9. Cotidal chart. Tide phases relative to Greenwich are plotted for all the world’s oceans. Phase progresses from red to orange to yellow to green to blue to purple. The lines converge on anphidromic points, singularities on the earth’s surface where there is no defined tide. [Winfree, 1987 #1195, p. 17].
Classing quantitative data

<table>
<thead>
<tr>
<th>Age-adjusted</th>
<th>Rate per 100,000 population</th>
<th>Comparative mortality ratio (HSA to U.S.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>253.8 - 328.6</td>
<td>1.24 - 1.60</td>
<td>1.16 - 1.24</td>
</tr>
<tr>
<td>236.5 - 253.7</td>
<td>1.06 - 1.56</td>
<td>1.06 - 1.56</td>
</tr>
<tr>
<td>216.2 - 236.7</td>
<td>0.99 - 1.05</td>
<td>0.86 - 0.98</td>
</tr>
<tr>
<td>199.0 - 215.3</td>
<td>0.81 - 0.98</td>
<td>0.50 - 0.81</td>
</tr>
<tr>
<td>179.5 - 199.6</td>
<td>0.66 - 0.86</td>
<td></td>
</tr>
<tr>
<td>166.7 - 179.4</td>
<td>0.61 - 0.81</td>
<td></td>
</tr>
<tr>
<td>112.4 - 166.6</td>
<td>0.50 - 0.81</td>
<td></td>
</tr>
</tbody>
</table>

Age-adjusted mortality rates for the United States

Quantitative color encoding

**Sequential color scale**
- Constrain hue, vary luminance/saturation
- Map higher values to darker colors

**Diverging color scale**
- Useful when data has a meaningful “midpoint”
- Use neutral color (e.g., grey) for midpoint
- Use saturated colors for endpoints

Limit number of steps in color to 3-9
Color Brewer

www.colorbrewer.org

Sequential color scheme

Sequential Scheme One Hue

Percent of labor force employed in SERV ICSS 1980

- 60 to 63
- 50 to 59
- 40 to 48
- 33 to 30
- 26
- 18
Sequential color scheme

**Design of sequential color scales**

**Hue-Lightness** *(Recommended)*
- Higher values mapped to darker colors
- ColorBrewer schemes have 3-9 steps

**Hue Transition**
- Two hues
- Neighboring hues interpolate better
- Couple with change in lightness
Diverging color scheme

Hue Transition
Carefully handle midpoint
  ■ Critical class
    - Low, Average, High
    - ‘Average’ should be gray
  ■ Critical breakpoint
    - Defining value e.g. 0
    - Positive & negative should use different hues

Extremes saturated, middle desaturated

Hints for the colorist

Use only a few colors (~6 ideal)
Colors should be distinctive and named
Strive for color harmony (natural colors?)
Use cultural conventions; appreciate symbolism
Beware of bad interactions (red/blue etc.)
Get it right in black and white
Respect the color blind
Assignment 3: Visualization Software

Create a small interactive visualization application – you choose data domain and visualization technique.

1. Describe data and storyboard interface
2. Implement interface and produce final writeup
3. Submit the application and a final writeup on the wiki

Can work alone or in pairs
Final write up due before class on Oct 15, 2014
Using Space Effectively: 2D

Topics
Displaying data in graphs
Banking to 45 degrees
Fitting data and depicting residuals
Graphical calculations
Zooming and distortion
Graphs and Lines

Effective use of space

Which graph is better?

Government payrolls in 1937 [Huff 93]
Aspect ratio

Fill space with data
Don’t worry about showing zero

Yearly CO2 concentrations [Cleveland 85]

Clearly mark scale breaks

Poor scale break [Cleveland 85]
Well marked scale break [Cleveland 85]
Scale break vs. Log scale

Both increase visual resolution
- Log scale - easy comparisons of all data
- Scale break – more difficult to compare across break

[Cleveland 85]
Linear scale vs. Log scale

Linear scale
- Absolute change

Log scale
- Small fluctuations
- Percent change
\[ d(10,20) = d(30,60) \]
Exponential functions ($y = ka^{mx}$) transform into lines
\[
\log(y) = \log(k) + \log(a)mx
\]
Intercept: $\log(k)$
Slope: $\log(a)m$

Semilog graph: Exponential growth
\[
y = 6^{0.5x}, \text{ slope in semilog space: } \log(6)^{0.5} = 0.3891
\]

Semilog graph: Exponential decay
\[
y = 0.5^{2x}, \text{ slope in semilog space: } \log(0.5)^2 = -0.602
\]
Log-Log graph

Power functions \((y = kx^a)\) transform into lines

Example - Steven’s power laws:

\[ S = kI^p \rightarrow \log S = \log k + p \log I \]
Aspect ratio

Fill space with data
Don’t worry about showing zero

Yearly CO2 concentrations  [Cleveland 85]

William S. Cleveland
The Elements of Graphing Data
Banking to 45° [Cleveland]

To facilitate perception of trends, maximize the discriminability of line segment orientations.

Two line segments are maximally discriminable when avg. absolute angle between them is 45°

Optimize the aspect ratio to bank to 45°

Aspect-ratio banking techniques

<table>
<thead>
<tr>
<th>Median-Absolute-Slope</th>
<th>Average-Absolute-Slope</th>
</tr>
</thead>
</table>
| \[ \alpha = \text{median} \left( \frac{R_x}{R_y} \right) \] | \[ \alpha = \text{mean} \left( \frac{R_x}{R_y} \right) \]

Has Closed Form Solution

<table>
<thead>
<tr>
<th>Average-Absolute-Orientation</th>
<th>Max-Orientation-Resolution</th>
</tr>
</thead>
</table>
| Unweighted \[ \sum_{i} \left| \theta_i(\alpha) \right| = 45° \] | Global (over all i, j s.t. i ≠ j) \[ \sum_{i} \sum_{j} \left| \theta_i(\alpha) - \theta_j(\alpha) \right|^2 \]
| Weighted \[ \frac{\sum_{i} \left| \theta_i(\alpha) \right| \cdot l_i(\alpha)}{\sum_{i} l_i(\alpha)} = 45° \] | Local (over adjacent segments) \[ \sum_{i} \left| \theta_i(\alpha) - \theta_{i+1}(\alpha) \right|^2 \]

Requires Iterative Optimization
**Slopeless line culling**

- **Standard, Aspect Ratio** = 1.97
- **Culled, Aspect Ratio** = 4.00
- Exclude line segments with zero or infinite slope

**Comparison data sets**

- **CO₂ Measurements (co2)**
- **PRMTX Mutual Fund (prmtx)**
- **Prefuse Downloads**
- **Sunspot Cycles (sunspot)**
Comparison (Results)

Average-slope (as) and Average-weighted-orientation (awo) provide similar ratios

Comparison (Results)

Average-orientation (ao) and Global-orientation-resolution (gor) provide similar ratios
Arc-length based aspect ratio

\[
\min_{\alpha \in \{0, \ldots, N-1\}} \sum_{i=1}^{N} \left| \frac{\Delta x_i}{\sqrt{\Delta y_i}} \right|
\]

Arc-length parameterization independent of spacing of points
Result are better than GOR and LOR and close to Cleveland’s AWO

[Talbot 11]

Perceptual model based aspect ratio

Ask people to estimate slope ratios for different conditions
Use data to fit a model derived from perceptual theory

\[
\beta_{ij} = \begin{cases} 
\frac{\mu[1 + (\beta_{ij})]}{\sin(\phi)} \times 100 + \gamma & \text{if HERT} \\
\frac{\mu}{\beta_{ij}} \times 100 + (\mu + \beta_{ij}) + \epsilon_{ij} & \text{if ANGLE}
\end{cases}
\]

[Talbot 13]
Perceptual model

Perceptual model (black diamond) produces flatter aspect ratios than other techniques.

Aspect Ratio = 1.17

CO$_2$ Measurements
William S. Cleveland
Visualizing Data
**Multi-Scale Banking to 45°**

**Goal**
- Optimized aspect ratios for varying scales

**Approach**
- Identify Scales of Interest
- Generate Scale-Specific Trend Lines
- Bank Trend Lines to 45°
- Filter Resulting Aspect Ratios

---

**Multi-Scale Banking to 45°**

**Use Spectral Analysis to identify trends**
- Find strong frequency components
- Lowpass filter to create trend lines
Compute Power Spectrum

- Take Discrete Fourier Transform
- Compute squared magnitudes

Power Spectrum

Smooth the Spectrum

- Convolve with Gaussian filter
  - window size = 3, \( \sigma = 1 \)

Smoothed Power Spectrum
Threshold the Spectrum

- Threshold at mean of power spectrum
- Retain last values of contiguous runs

Generate Trend Lines

- Generate trend lines with lowpass filter
Bank Trend Lines to 45°

Filter Aspect Ratios

Filter similar aspect ratios
Keep if $\alpha_{i+1} > c\alpha_i$ (c=1.25 by default)
Sunspot Cycles

Yearly values 1700-1987

Aspect Ratio = 4.23
Aspect Ratio = 14.55

Power Spectrum
Aspect Ratios

CO₂

Monthly concentrations from the Mauna Loa Observatory, 1950-1990

Aspect Ratio = 1.17
Aspect Ratio = 7.87

Power Spectrum
Aspect Ratios
### Applications

<table>
<thead>
<tr>
<th>Small Multiples</th>
<th>Trend Explorer</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Small Multiples Image" /></td>
<td><img src="image2" alt="Trend Explorer Image" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sparklines</th>
<th>Sparklines</th>
</tr>
</thead>
<tbody>
<tr>
<td>VFINX</td>
<td>119.27</td>
</tr>
<tr>
<td>GOOG</td>
<td>364.80</td>
</tr>
<tr>
<td>MSFT</td>
<td>27.14</td>
</tr>
<tr>
<td>YHOO</td>
<td>32.18</td>
</tr>
</tbody>
</table>

### Fitting the Data
[The Elements of Graphing Data. Cleveland 94]
Linear regression ...

Linear regression w/out outlier ...

[The Elements of Graphing Data. Cleveland 94]
Transforming data

How well does curve fit data?

Residual graph
- Plot vertical distance from best fit curve
- Residual graph shows accuracy of fit

[Cleveland 85]
Most powerful brain?

The Dragons of Eden [Carl Sagan]
Graphical Calculations
Nomograms

1. Compute in any direction; fix n-1 params and read nth param
2. Illustrate sensitivity to perturbation of inputs
3. Clearly show domain of validity of computation
### Theory

\[
\begin{bmatrix}
  x_1(u) & y_1(u) & w_1(u) \\
  x_2(v) & y_2(v) & w_2(v) \\
  x_3(s,t) & y_3(s,t) & w_3(s,t)
\end{bmatrix} = 0
\]

### Slide rule

**Model 1474-66**  
**Electrotechnica**  
18 Scales

Tehnolemn Timisoara Slide Rule Archive  
http://pubpages.unh.edu/~jwc/tehnolemn/
Johannes Lambert used graphs to study the rate of water evaporation as function of temperature [from Tufte 83]
Zooming

Eames’ Powers of Ten [http://www.powersof10.com/film]

Overview + details

[Hornbaek et al. 2002]
Interactive zooming

Pad++ [Bederson and Hollan 94]

Pad++
Semantic zooming
Change visual representations as zoom level changes

PAD [Perlin and Fox 93]

TableLens [Rao & Card 94]

http://www.youtube.com/watch?v=qWqTrRAC52U
Datelens

Distortion
Single view detail + context

- Focus area – local details
- De-magnified area – surrounding context
- Like a rubber sheet with borders tacked down

Nonlinear Magnification Infocenter [http://www.cs.indiana.edu/~etkeahey/research/nlm/nlm.html]

Bifocal display [Leung and Apperley 94]

2D distortion
Multifocal display  [Leung and Apperley 94]

Fisheye  [Leung and Apperley 94]
Nonlinear magnification [Leung and Apperley 94]

6 types of distortions [Carpendale & Montagnese 01]

Gaussian, Cosine, Hemisphere, Linear, Inverse Cosine and Manhattan. Top row shows transition from focus to distortion, bottom row from distortion to context.
Cartographic Distortions
Cartograms: Distort areas

Scale area by data
[From Cartography, Dent]

Election 2012 map

% voted democrat
% voted republican

http://www-personal.umich.edu/~mejn/election/
Election 2012 map

http://www-personal.umich.edu/~mejn/election/
Statistical map with shading

MURDER RATES, 1978

1.2
4.0
8.5
12.1
15.8

FIVE REPRESENTATIVE SHADINGS

RATES PER 100,000 POPULATION

Figure 8. Statistical map with shading.

[Framed and McGill 84]

Framed rectangle chart

MURDER RATES PER 100,000 POPULATION, 1978

[Framed and McGill 84]
Rectangular cartogram

American population [van Kreveld and Speckmann 04]

Rectangular cartogram

Native American population [van Kreveld and Speckmann 04]
Dorling cartogram

http://www.ncgia.ucsb.edu/projects/Cartogram_Central/types.html

States as nodes in a graph

Graphical fisheye views of graphs [Sarkar & Brown 92]
Distorting distances

Scale distance by data

[From Cartography, Dent]

London underground

http://www.thetube.com/content/history/map.asp
Comparison to geographic map

- Spatial layout is the most important visual encoding
- Geometric properties of spatial transforms support geometric reasoning
- Show data with as much resolution as possible
- Use distortions to emphasize important information