Data Visualization (DSC 530/CIS 602-01)

Color & Colormaps

Dr. David Koop
Color
Color

[Image of a pattern with green leaves and red apples]

[C. Ware]
Color and Light

- Color is a **perceptive** property: color depends on the eyes and brain.
- Visible light is a small portion of the **electromagnetic spectrum** which is composed of waves that at various frequencies (wavelengths), all traveling at the speed of light.
Electromagnetic Spectrum

Penetrates Earth's Atmosphere?

Radiation Type
- Radio
- Microwave
- Infrared
- Visible
- Ultraviolet
- X-ray
- Gamma ray

Wavelength (m)
- Radio: $10^3$
- Microwave: $10^{-2}$
- Infrared: $10^{-5}$
- Visible: $0.5 \times 10^{-6}$
- Ultraviolet: $10^{-8}$
- X-ray: $10^{-10}$
- Gamma ray: $10^{-12}$

Approximate Scale of Wavelength
- Buildings
- Humans
- Butterflies
- Needle Point
- Protozoans
- Molecules
- Atoms
- Atomic Nuclei

Frequency (Hz)
- Radio: $10^4$
- Microwave: $10^8$
- Infrared: $10^{12}$
- Visible: $10^{15}$
- Ultraviolet: $10^{16}$
- X-ray: $10^{18}$
- Gamma ray: $10^{20}$

Temperature of objects at which this radiation is the most intense wavelength emitted
- Radio: 1 K (~−272 °C)
- Microwave: 100 K (~−173 °C)
- Infrared: 10,000 K (9,727 °C)
- Visible: ~10,000,000 K (~10,000,000 °C)

[Wikimedia, NASA]

D. Koop, DSC 530, Spring 2018
Light Reflection & Absorption

How does light work?

D. Koop, DSC 530, Spring 2018

[via M. Meyer]
Color $\neq$ Wavelength

[Diagram showing relative energy density vs. wavelength with peaks at different wavelengths for yellow and brown colors]

[via M. Meyer]
Human Color Perception

- Humans are **trichromatic**—we have three different types of cones
  - S (430nm): blue
  - M (540nm): green
  - L (570nm): "red"
- Note that the response curves **overlap**
- Spectra of visible light are "covered" by these responses
- Three numbers -> color

[Vanessaezekowitz at en.wikipedia]
Human Color Perception

- Humans do not detect individual wavelengths of light
- Use rods and cones to detect light
  - rods capture intensity
  - cones capture color
Human Color Perception

Metamerism: same three responses == same color

[via M. Meyer]
Metamerism
Project Proposal

- Information
- Due Today
- Choices:
  - [CreateVis] Create a complex, interactive visualization of a dataset
  - [Research] Research project that involves visualization (new technique, evaluation, etc.)
Test 1

• **Information**
• This Wednesday in class (March 7)
• Format:
  - Multiple Choice
  - Free Response
• Questions will involve research papers discussed in class
  - List of papers on the [web site](#)
• Questions may involve interpreting code or writing pseudocode
  - Array functions
  - D3 selections
Human Color Perception

- Humans **do not** detect individual wavelengths of light
- Use **rods** and **cones** to detect light
  - rods capture intensity
  - cones capture color
Color

• Cones respond to different areas of the visible light spectrum
• Cover all wavelengths but certain wavelengths generate greater responses
• Color is determined by calculations based on the responses from the different cones
• Opponent Process Theory: three "opponent" channels
  - Light/Dark
  - Blue/Yellow
  - Red/Green
• Opposite colors are not perceived together
Opponent Process Theory

[Machado et. al, 2009]
Color Blindness

[Ishihara (Plate 9) via Wikipedia]
Color Blindness

- Sex-linked: 8% of males and 0.4% of females of N. European ancestry
- Abnormal distribution of cones (e.g. missing the S, M, or L types)
- Either dichromatic (only two types of cones) or anomalous trichromatic (one type of cones has a defect)
  - Protanopia (L missing), Protanomaly (L defect)
  - Deuteranopia (M missing), Deuteranomaly (M defect) [Most Common]
  - Tritanopia (S missing), Tritanomaly (S defect) [Rare]
- Dichromacy is rarer than anomalous trichromacy
- Opponent process model explains why colors cannot be differentiated
Color Blindness

Normal

Protopanopia

Deuteranopia

Tritanopia

[via M. Meyer]
Simulating Color Blindness

<table>
<thead>
<tr>
<th></th>
<th>Empty</th>
<th>Photop. Subst.</th>
<th>Scale Ratio</th>
<th>0.96*Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Protanopia</strong></td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td><strong>Deuteranopia</strong></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
</tbody>
</table>

[Machado et. al, 2009]
Simulating Color Blindness

(Machado et. al, 2009)

D. Koop, DSC 530, Spring 2018

[Image of brain scans showing different colorblindness simulations]
Simulating Color Blindness

[Machado et. al, 2009]
Primary Colors?

- Red, Green, and Blue
- Red, Yellow, and Blue
- Orange, Green, and Violet
- Cyan, Magenta, and Yellow
Primary Colors?

- Red, Green, and Blue
- Red, Yellow, and Blue
- Orange, Green, and Violet
- Cyan, Magenta, and Yellow
- All of the above!
Color Addition and Subtraction

![Color Addition and Subtraction Diagram](image-url)
Color Spaces and Gamuts

- All colors visible to the average human eye are contained inside the diagram.
- The colors along any line between two points can be made by mixing the colors at the end points. In this case Green + Red = Yellow.
- The edge of the diagram, called the spectral locus, represents pure, monochromatic light measured by wavelength in nanometers. These are the most saturated colors.
- The least saturated colors are at the center, emanating from white.
- Color gamut: subset of colors that can be represented by mixing the colors at its corners.
- "Line of purples": these colors are fully saturated but can only be made by mixing two colors (red and blue).

[Anatomy of a CIE Chromaticity Diagram]
Color Spaces and Gamuts

- **Color space**: the organization of all colors in space
  - Often human-specific, what we can see (e.g. CIELAB)
- **Color gamut**: a subset of colors
  - Defined by corners on in the color space
  - What can be produced on a monitor (e.g. using RGB)
  - What can be produced on a printer (e.g. using CMYK)
  - The gamut of your monitor != the gamut of someone else's != the gamut of a printer
Color Models

• A **color model** is a representation of color using some basis
• RGB uses three numbers (red, blue, green) to represent color
• Color space ~ color model, but there can be many color models used in the same color space (e.g. OGV)
• Hue-Saturation-Lightness (HSL) is more intuitive and useful
  - Hue captures pure colors
  - Saturation captures the amount of white mixed with the color
  - Lightness captures the amount of black mixed with a color
  - HSL color pickers are often circular
• Hue-Saturation-Value (HSV) is similar (swap black with gray for the final value), linearly related
Luminance

- HSL does not truly reflect the way we perceive color
- Even though colors have the same lightness, we perceive their luminance differently
- Our perception ($L^*$) is nonlinear

[Corners of the RGB color cube]

- L from HSL
  - All the same

[Luminance]

- $L^*$

[Munzner (ill. Maguire), 2014 (based on Stone, 2006)]
Perceptually Uniform Color Spaces

- $L^*a^*b^*$ allows perceptually accurate comparison and calculations of colors

[J. Rus, CC-BY-SA (changed to horizontal layout)]
Luminance Perception (Spatial Adaption)

Edward H. Adelson

[E. H. Adelson, 1995]
Luminance Perception (Spatial Adaption)

Edward H. Adelson

[E. H. Adelson, 1995]
Simultaneous Contrast
Simultaneous Contrast
Simultaneous Contrast
Simultaneous Contrast
What colors?
What colors?

Red, yellow, blue

Purple, orange do not exist!

[A. Kitaoka]
Violations of CIELAB Assumptions

- D. Szafir examined how visualization design affects color perception

- CIELAB:
  - Approximately perceptually linear
  - 1 unit of Euclidean distance = 1 Just Noticeable Difference (JND)
  - JND: people detect change at least 50% of the time

- Assumptions CIELAB makes:
  - Simple world
  - Isolation
  - Geometric

[D. Szafir, 2017]
Simple World Assumption

[D. Szafir, 2017]
Problems with Simple World Assumption

- Isolation Assumption
- Geometric Assumption

Visualizations violate three CIELAB assumptions.

Crowdsourced Sampling

Szafir, Stone, & Gleicher, 2014

Reinecke, Flatla, & Brooks, 2016

[D. Szafir, 2017]
Isolation Assumption

[D. Szafir, 2017]
Problems with Isolation Assumption

[D. Szafir, 2017]
Geometric Assumption

[D. Szafir, 2017]
Size Problem with Geometric Assumption

Visualizations violate three CIELAB assumptions

- Based Sampling
- Carter & Silverstein, 2010
- Stone, Szafir, & Setlur, 2014

[D. Szafir, 2017]
Shape Problem with Geometric Assumption

Visualizations violate three CIELAB assumptions

[D. Szafir, 2017]
Types of Geometry

- Diagonally Symmetric Marks
- Elongated Marks
- Asymmetric Marks
- Area Marks

[D. Szafir, 2017]
Run the tests!
6 (diameters, within) × 6 (color differences, within) × 3 (color axis, between)

81 participants on Mechanical Turk (5,668 trials)

25 pixels
1.0°

37 pixels
1.5°

50 pixels
2.0°

[D. Szafir, 2017]
50% JND for Scatterplot Points

Point Diameter
Visual Angle

JND in CIELAB

△b*
△a*
△L*

Stone et al.

ΔE = 1.0

Point Size: consistent with previous results

[D. Szafir, 2017]
Bar Thickness and Length: longer bars help

50% JND for Bars

Bar Length encoded as point size

JND in CIELAB

Bar Thickness

Visual Angle

$\Delta b^*$ for points

$\Delta a^*$ for points

$\Delta L^*$ for points

[D. Szafir, 2017]
Line Thickness: better than points

50% JND for Lines

JND in CIELAB

Line Thickness

Visual Angle

R² = .97
R² = .93
R² = .90

Δb* for points
Δa* for points
ΔL* for points

[D. Szafir, 2017]
Color perception in real-world visualizations is complicated
Akiyoshi Kitaoka's Illusion pages
Colormap

• A colormap specifies a mapping between colors and data values
• Colormap should follow the expressiveness principle
• Types of colormaps:

Binary

\[
\begin{array}{c}
y \\
n \\
\end{array}
\]

Diverging

\[
\begin{array}{c}
-1 \\
0 \\
+1 \\
\end{array}
\]

Categorical

\[
\begin{array}{c}
T \\
F \\
A \\
\end{array}
\]

Sequential

\[
\begin{array}{c}
3 \\
2 \\
1 \\
\end{array}
\]

[Munzner (ill. Maguire), 2014]
Categorical vs. Ordered

- Hue has no implicit ordering: use for categorical data
- Saturation and luminance do: use for ordered data

[Munzner (ill. Maguire), 2014]
Categorical Colormap Guidelines

- Don't use too many colors (~12)
- Remember your background has a color, too
- Nameable colors help
- Be aware of luminance (e.g. difference between blue and yellow)
- Think about other marks you might wish to use in the visualization
Categorical Colormaps

[Colorbrewer2.org]
Categorical Colormaps
Number of distinguishable colors?

[Sinha & Meller, 2007]
Number of distinguishable colors?

[Sinha & Meller, 2007]

[Scale (mb)]
0 40 80
Discriminability

- Often, fewer colors are better
- Don't let viewers combine colors because they can't tell the difference
- Make the combinations yourself
- Also, can use the "Other" category to reduce the number of colors
Ordered Colormaps

- Used for ordinal or quantitative attributes
- \([0, N]\): Sequential
- \([-N, 0, N]\): Diverging (has some meaningful midpoint)
- Can use hue, saturation, and luminance
- Remember hue is not a magnitude channel so be careful
- Can be **continuous** (smooth) or **segmented** (sharp boundaries)
  - Segmented matches with ordinal attributes
  - Can be used with quantitative data, too.
Continuous Colormap

US EPA Regional Oxidant Model -- Midwest Ozone (ppbv): June 26, 1987, 18:00

[Bergman et al., 1995]
Segmented Colormap

US EPA Regional Oxidant Model -- Midwest
Ozone (ppbv): June 26, 1987, 18:00

[Bergman et al., 1995]