CP SC 881.001
Data Visualization

Joshua Levine
levinej@clemson.edu
Lecture 06
Color
Sept. 11, 2014
Agenda

- Lab 01 Questions — due Sept. 16 (Tuesday!)
- Design Critique: new due date is each Wed.
- No questions on the paper readings are due next week (not much to read!)
Continuing from Lec05
Marks

• Basic graphical elements / primitives

• Classified according to number of spatial dimensions required

- Points (0-dimensional)
- Lines (1-dimensional)
- Areas (2-dimensional)
Channels

- Parameters that control the appearance of marks

- **Position**
  - Horizontal
  - Vertical
  - Both

- **Color**
  - Black
  - Red
  - Green

- **Shape**
  - Triangle
  - Star
  - Slash
  - L-shape

- **Tilt**
  - Slanted

- **Size**
  - Length
  - Area
  - Volume
(how much)

**Magnitude Channels: Ordered Attributes**
- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)
- Color luminance
- Color saturation
- Curvature
- Volume (3D size)

(what or where)

**Identity Channels: Categorical Attributes**
- Spatial region
- Color hue
- Motion
- Shape
Time
Recall: Attribute Semantics

Temporal

- What makes time a challenge (semantically)?
Recall: Attribute Semantics

Temporal

• What makes time a challenge (semantically)?

• Hierarchical:
Recall: Attribute Semantics

Temporal

• What makes time a challenge (semantically)?

• Hierarchical:
  • Many levels millennia to nanoseconds
Recall: Attribute Semantics

Temporal

• What makes time a challenge (semantically)?

• Hierarchical:
  • Many levels millennia to nanoseconds
  • Not all nested perfectly (how do weeks fit into months)?
Recall: Attribute Semantics

Temporal

• What makes time a challenge (semantically)?

• Hierarchical:
  • Many levels millennia to nanoseconds
  • Not all nested perfectly (how do weeks fit into months)?

• Periodicity
Recall: Attribute Semantics

Temporal

- What makes time a challenge (semantically)?

- Hierarchical:
  - Many levels millennia to nanoseconds
  - Not all nested perfectly (how do weeks fit into months)?

- Periodicity

- Can have either key OR value semantics
Recall: Attribute Semantics

Temporal

- What makes time a challenge (semantically)?

- Hierarchical:
  - Many levels millennia to nanoseconds
  - Not all nested perfectly (how do weeks fit into months)?

- Periodicity

- Can have either key OR value semantics
  - Affects how we interpret attribute type
Recall: Attribute Semantics

Temporal

• What makes time a challenge (semantically)?

• Hierarchical:
  • Many levels millennia to nanoseconds
  • Not all nested perfectly (how do weeks fit into months)?

• Periodicity

• Can have either key OR value semantics
  • Affects how we interpret attribute type

• Use of time-series vs. dynamics.
Using Time as an Encoding Channel

- Fundamental Question: External vs. Internal Memory
  - Easy to compare views by moving eyes
  - Hard to compare view with memory of what you saw

- Recall: Visualization’s purpose is to serve as an external aid to **augment** working memory
Can you spot 12 differences between these pictures?
Can you spot 12 differences between these pictures?
ComParrot

Can you spot 12 differences between these pictures?

Solution:

1. Top tree ear removed.
2. Nose line on left giraffe removed.
3. Shadow on lower left coconut removed.
4. Leaf on left.
5. Eye on left giraffe removed.
6. Bottom shadow on right giraffe colored.
7. Small leaf on right of tree colored.
8. Horn on right.
9. Giraffe moved.
10. Shadow on left side shorter.
11. Giraffe tail longer.
Animation: can it facilitate?

BARBARA Tversky and JULIE BAUER MORRISON†
Department of Psychology, Jordan Hall; Bldg. #420 Stanford, CA 94305-2130 USA.

Mireille Betrancourt
I.N.R.I.A. Rhone-Alpes, 655, Av. de l’Europe, 38330 Montbonnot St-Martin, France.
email: mireille.betrancourt@inria.fr.

(Received 1 March 2001 and accepted in revised form 4 April 2002)

Graphics have been used since ancient times to portray things that are inherently spatiovisual, like maps and building plans. More recently, graphics have been used to portray things that are metaphorically spatiovisual, like graphs and organizational charts. The assumption is that graphics can facilitate comprehension, learning, memory, and inference. Assumptions aside, research on static graphics has shown that carefully designed and appropriate graphics prove to be beneficial for complex systems. Effective graphics conform to the Congruence Principle, which states that the content and format of the graphic should correspond to the format of the concepts to be conveyed. From this, it follows that animated graphics should be effective in portraying change over time. Yet the research on the efficacy of animated over static graphics is not encouraging. In cases where animated graphics were used to portray static or active systems, no clear advantage was shown.
Good Use: Animating Transitions

Animated Transitions in Statistical Data Graphics

Jeffrey Heer
George G. Robertson

Microsoft Research
Good Use: Animating Transitions

Animated Transitions in Statistical Data Graphics

Jeffrey Heer
George G. Robertson

Microsoft Research
Bad Use: Comparing Complex State Changes Over Time
Bad Use: Comparing Complex State Changes Over Time
Color
WOW, HONEY, YOU'RE MISSING A BEAUTIFUL SUNSET OUT HERE!

I'LL COUNT TO 10, AND THEN... PON!

DAD, HOW COME OLD PHOTOGRAPHS ARE ALWAYS BLACK AND WHITE? DIDN'T THEY HAVE COLOR FILM BACK THEN?

SURE THEY DID. IN FACT, THOSE OLD PHOTOGRAPHS ARE IN COLOR. IT'S JUST THE WORLD WAS BLACK AND WHITE THEN.

YES, THE WORLD DIDN'T TURN COLOR UNTIL SOMETIME IN THE 1930S. AND IT WAS PRETTY GRAY BACK THEN.

REALLY?

THAT'S REALLY WEIRD.

WELL, TRUTH IS STRANGER THAN FICTION.

BUT THEN WHY ARE OLD PAINTINGS IN COLOR? IF THE WORLD WAS BLACK AND WHITE, WOULDN'T ARTISTS HAVE PAINTED IT THAT WAY?

NOT NECESSARILY. A LOT OF GREAT ARTISTS WERE INSANE.

BUT...BUT HOW COULD THEY HAVE PAINTED IN COLOR ANYWAY? WOULDN'T THEIR PAINTS HAVE BEEN SHADES OF GRAY BACK THEN?

OF COURSE, BUT THEY TURNED COLORS LIKE EVERYTHING ELSE DID IN THE '30S.

SO WHY DIDN'T OLD BLACK AND WHITE PHOTOS TURN COLOR TOO?

BECAUSE THEY WERE COLOR PICTURES OF BLACK AND WHITE. REMEMBER?

THE WORLD IS A COMPLICATED PLACE, HOBES.

WHENEVER IT SEEMS THAT WAY, I TAKE A NAP IN A TREE AND WAIT FOR DINNER.
What are the primary colors?

1. Red, Green, Blue
2. Red, Yellow, Blue
3. Orange, Green, Violet
4. Cyan, Magenta, Yellow
What are the primary colors?

1. Red, Green, Blue
2. Red, Yellow, Blue
3. Orange, Green, Violet
4. Cyan, Magenta, Yellow
5. All of the above
“Tying color to information is as elementary and straightforward as color technique in art, “To paint well is simply this: to put the right color in the right place,” in Paul Klee’s ironic prescription. The often scant benefits derived from coloring data indicate that even putting a good color in a good place is a complex matter. Indeed, so difficult and subtle that avoiding catastrophe becomes the first principle in bringing color to information: Above all, do no harm.”
Purposes of Color

- To label (color as a noun)
- To measure (color as a quantity)
- To represent and imitate (color as a symbol)
- To enliven and decorate (color as beauty)

Matterhorn, Landeskarte der Schweiz (Wabern, 1983)
Functions of Color

Identify, Group, Layer, Highlight

Color also tells us much that is useful about the material properties of objects. This is crucial in judging the condition of our food. Is this fruit ripe or not? Is this meat fresh or putrid? What kind of mushroom is this? It is also useful if we are making tools. What kind of stone is this? Clearly, these can be life-or-death decisions. In modern hunter–gatherer societies, men are the hunters and women are the gatherers. This may have been true for long periods of human evolution, which could explain why it is mostly men who are color blind. If they had been gatherers, they would have been more than likely to eat poison berries—a selective disadvantage.

In the modern age of supermarkets, these skills are much less valuable; this is perhaps why color deficiencies so often go unnoticed.

The role that color plays ecologically suggests ways that it can be used in information display. It is useful to think of color as an attribute of an object rather than as its primary characteristic. It is excellent for labeling and categorization, but poor for displaying shape, detail, or spatial layout. These points are elaborated in this chapter. We begin with an introduction to the basic theory of color vision to provide a foundation for the applications. The latter half of the chapter consists of a set of four visualization problems requiring the effective use of color; these have to do with color selection interfaces, color labeling, pseudocolor sequences for mapping, color reproduction, and color for multidimensional discrete data. Each has its own special set of requirements. Some readers may wish to skip directly to the applications, sampling the more technical introduction only as needed.

Trichromacy Theory

The most important fact about color vision is that we have three distinct color receptors, called cones, in our retinas that are active at normal light levels—hence trichromacy. We also have rods, sensitive at low light levels, but they are so overstimulated in all but the dimmest light that their influence on color perception can be ignored. Thus, in order to understand color vision, we need only consider the cones. The fact that there are only three receptors is the reason for the basic three-dimensionality of human color vision. The term color space means an arrangement of colors in a three-dimensional

Figure 4.1
Finding the cherries is much easier with color vision.

Colin Ware, Information Visualization: Perception for Design
Physics of Light
Light is Electromagnetic Radiation

- Visible spectrum is “tiny”
- Wavelength range: 380-740 nm
Isaac Newton, 1666

What happens when a second prism is put in one of the colors?
Color $\neq$ Wavelength

But rather, a combination of wavelengths and energy.
Human Visual System
Recall: Photoreceptors

Rods

- Approximately 100-150 million rods.
- Non-uniform distribution across the retina
- Sensitive to low-light levels (scotopic vision)

Cones

- Approximately 6-7 million cones.
- Sensitive to daytime-light levels (photopic vision)
- Detect color by the use of 3 different kinds:
  - Red (L cone): 564-580nm wavelengths (65% of all cones)
  - Green (M cone): 534-545nm (30% of all cones)
  - Blue (S cone): 420-440nm (5% of all cones)
Hunters
Gatherers
Cones (SML)
(short, medium, long)

In this chapter, a number of color spaces, designed for different purposes, are discussed. Complex transformations are sometimes required to convert from one color space to another, but they are all three dimensional, and this three-dimensionality derives ultimately from the three cone types. This is the reason why there are three different colors of liquid crystal in a television screen — red, green, and blue — and this is the reason why we learn in school that there are three primary paint colors — red, yellow, and blue. It is also the reason why printers have a minimum of three colored inks for color printing — cyan, magenta, and yellow. Engineers should be grateful that humans have only three color receptors. Some birds, such as chickens, have as many as 12 different kinds of color-sensitive cells. A television set for chickens would require 12 types of differently colored pixels!

Figure 4.2 shows the human cone sensitivity functions. The plots show how light of different wavelengths is absorbed by the three different receptor types (S, M, L). It is evident that two of the functions, L and M, which peak at 540 nanometers and 580 nanometers, respectively, overlap considerably; the third, S, is much more distinct, with peak sensitivity at 450 nanometers. The short-wavelength S receptor absorbs light in the blue part of the spectrum and is much less sensitive, which is another reason (besides chromatic aberration, discussed in Chapter 2) why we should not show detailed information such as text in pure blue on a black background.

Because only three different receptor types are involved in color vision, it is possible to match a particular patch of colored light with a mixture of just three colored lights, usually called primaries. It does not matter that the target patch may have a completely different spectral composition. The only thing that matters is that the matching primaries are balanced to produce the same response from the cone receptors as the target patch.

Figure 4.2 Cone sensitivity functions. The colors are only rough approximations to spectrum hues.

Abbreviations: S, short-wavelength cone sensitivity; M, medium-wavelength cone sensitivity; L, long-wavelength cone sensitivity.
Trichromacy

- Our 3 cones type cover the visible spectrum
  - Theoretically, all we need are 2 though (why?)
- Most birds, some fish, reptiles, and insects have 4, some as many as 12 (Mantis Shrimp)
- This is the “reason” why many of our color spaces are 3D
Mantis Shrimp

16 Photoreceptors, 12 for color sensitivity!
Opponent Process Model

- Trichromacy explains how the eye receives the signals while opponent process theory explains how the signals are processed.

- Visual system is oriented around **differences** between the responses.

---

*Figure 4.10* In the color opponent process model, cone signals are transformed into black–white (luminance), red–green, and yellow–blue channels.

Colin Ware, *Information Visualization: Perception for Design*
Color Models
Light Mixing

- **Additive** mix of colored lights
  - Add up wavelengths of light to make new colors
- Primary: RGB
- Secondary: CMY
- Neutral = R + G + B
- Commonly used by monitors, projectors, etc.
Ink Mixing

- **Subtractive** mix of transparent inks
  - Start with white and other wavelengths are selectively filtered.
  - Primary: CMY
  - Secondary: RGB
  - ~Black: C + M + Y
  - Actually use CMYK to get true black
Paint Mixing

- Physical mix of opaque paints
- Primary: RYB
- Secondary: OGV
- Neutral: R + Y + B
- Additive or Subtractive?
Paint Mixing

- Physical mix of opaque paints
- Primary: RYB
- Secondary: OGV
- Neutral: R + Y + B
- Additive or Subtractive? (Mostly Subtractive! The mixing isn’t perfect and includes reflection off the paint.)
Color Spaces
RGB Color Space

- Additive, useful for computer monitors (but that’s about it!)

- A poor match for the mechanics of how we see
  - Not perceptually uniform (e.g. more “greens” than “yellows”)
  - Also, the channels are not visually separable!

Figure 4.4 illustrates the concept. Three projectors are set up with overlapping beams. In the figure, the beams only partially overlap so that the mixing effect can be illustrated, but in a color-matching experiment they would overlap perfectly. To match the lilac-colored sample, the projectors are adjusted so that a large amount of light comes from the red and blue projectors and a smaller amount of light comes from the green projector.

Figure 4.5 The three-dimensional space formed by three primary lights. Any internal color can be created by varying the amount of light produced by each of the primaries.
\((H,S/C,L/B/V)\)

Color Space

- **Hue** - what people think of as color

- **Saturation** - purity, distance from grey (amount of white mixed in)
  - Also called Chroma

- **Lightness** - from dark to light (amount of black mixed in)
  - Also Brightness or Value

- More intuitive for designers and artists
Figure 10.3. Comparing HSL lightness, true luminance, and perceptually linear luminance $L^*$ for six colors. The computed HSL lightness $L$ is the same for all of these colors, showing the limitations of that color system. The true luminance values of these same six colors, as could be measured with an instrument. The computed perceptually linear luminance $L^*$ of these colors is the best match with what we see. After [Stone 06].
Tristimulus Experiment

- Color Matching Experiment in 1931
- CIE = International Commission on Illumination (Commission internationale de l'éclairage)
- Since some weighting factors for R,G,B lights are negative, they computed a new set of weights for a new of components X,Y,Z

**RGB Weights**

**XYZ Weights**
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
Note: Colors outside the triangles cannot be accurately displayed by that space or display.
CIELab/Luv

- Designed to approximate human vision
  - Perceptual uniform, but uses positive and negative values
- L approximates luminance
- (a,b) approximate chromaticity or M-to-G and Y-to-B channels
Colormaps
What is a Colormap?

• Specifies a mapping between color and values
  • Sometimes called a transfer function

• Colormaps can be:
  • categorical vs. ordered
  • sequential vs. diverging
  • segmented vs. continuous
  • univariate vs. bivariate

• Recall: expressiveness in visual encoding — Match colormap to attribute type characteristics!
What is a Colormap?

• Specifies a mapping between color and values
  • Sometimes called a **transfer function**

• Colormaps can be:
  • categorical vs. ordered
  • sequential vs. diverging
  • segmented vs. continuous
  • univariate vs. bivariate

• Recall: expressiveness in visual encoding — Match colormap to attribute type characteristics!
What is a Colormap?

• Specifies a mapping between color and values
  • Sometimes called a **transfer function**

• Colormaps can be:
  • categorical vs. ordered
  • sequential vs. diverging
  • segmented vs. continuous
  • univariate vs. bivariate

• Recall: expressiveness in visual encoding — Match colormap to attribute type characteristics!
What is a Colormap?

• Specifies a mapping between color and values
  • Sometimes called a transfer function

• Colormaps can be:
  • categorical vs. ordered
  • sequential vs. diverging
  • segmented vs. continuous
  • univariate vs. bivariate

• Recall: expressiveness in visual encoding —
  Match colormap to attribute type characteristics!
What is a Colormap?

• Specifies a mapping between color and values
  • Sometimes called a **transfer function**

• Colormaps can be:
  • categorical vs. ordered
  • sequential vs. diverging
  • segmented vs. continuous
  • univariate vs. bivariate

• Recall: expressiveness in visual encoding —
  Match colormap to attribute type characteristics!
What is a Colormap?

- Specifies a mapping between color and values
  - Sometimes called a **transfer function**
- Colormaps can be:
  - categorical vs. ordered
  - sequential vs. diverging
  - segmented vs. continuous
  - univariate vs. bivariate
- Recall: expressiveness in visual encoding — Match colormap to attribute type characteristics!
Encoding Nominal Data
Categorical Colormaps

• Categorical colors are easier to remember if they are nameable

• Typically designed by using color as an integral identity channel to encode a single attribute

Colin Ware, Information Visualization: Perception for Design
Recall: Popout

Hue and lightness

Lightness only

Based on slide from Stone
Brightness and Saturation
Draw Attention
Distinguishability

Only good at distinguishing 6-12 simultaneous colors
Color vs. Size

- Use more saturated colors when color coding small symbols, thin lines, or other small areas. Use less saturated colors for coding large areas.

Colin Ware, Information Visualization: Perception for Design
Example: Small Areas
Luminance Contrast

- When small symbols, text, or other detailed graphical representations of information are displayed using color on a differently colored background, always ensure luminance contrast with the background.
Form

• When applying shading to define the shape of a curved surface, use adequate luminance (as opposed to chromatic) variation.

• If large areas are defined using nearly equiluminous colors, consider using thin border lines with large luminance differences (from the colors of the areas) to help define the shapes.

Colin Ware, Information Visualization: Perception for Design
Which area is bigger, red or green?


Figure 1. Stimulus From the High-Saturation Group
Cleveland & McGill, "A Color-Caused Optical Illusion on a Statistical Graph", 1983
Effects of Color on Large Areas
Effects of Color on Large Areas

Tufte, VDQI (Vol. 1), p. 76
Summary/Guidelines

• Saturation interacts strongly with size

• In small regions:
  • Saturation is more difficult to perceive, in particular saturation and hue are not separable
  • Use bright, highly saturated colors
  • For points and lines use just two saturation levels

• In large regions
  • Higher saturation makes large areas look bigger
  • Use low saturation pastel colors for large regions and backgrounds
Tools for Color
Lec07
Required Reading
Getting Started

This tutorial is for Processing 2+. If you see any errors or have comments, please let us know. This tutorial was adapted from the book, Getting Started with Processing, by Casey Reas and Ben Fry, O'Reilly / Make 2010. Copyright © 2010 Casey Reas and Ben Fry. All rights reserved.

Welcome to Processing!

Start by visiting http://processing.org/download and selecting the Mac, Windows, or Linux version, depending on what machine you have. Installation on each machine is straightforward:

- On Windows, you'll have a .zip file. Double-click it, and drag the folder inside to a location on your hard disk. It could be Program Files or simply the desktop, but the important thing is for the processing folder to be pulled out of that .zip file. Then double-click processing.exe to start.

- The Mac OS X version is also a .zip file. Double-click it and drag the Processing icon to the Applications folder. If you're using someone else's machine and can't modify the Applications folder, just drag the application to the desktop. Then double-click the Processing icon to start.

- The Linux version is a .tar.gz file, which should be familiar to most Linux users. Download the file to your home directory, then open a terminal window, and type:
  ```
tar xvzf processing-xxxx.tgz
```
(Replace xxxx with the rest of the file's name, which is the version number.) This will create a folder named processing-2.0 or something similar. Then change to that directory:
  ```
cd processing-xxxx
```
and run it:
  ```
./processing
```