Lecture 14:

Color

Computer Graphics
CMU 15-462/15-662, Spring 2017
Why do we need to be able to talk precisely about color?
Starry Night, Van Gogh
What is color?
Light is EM Radiation; Color is Frequency

- Light is oscillating electric & magnetic field
- KEY IDEA: frequency determines color of light
- Q: What is the difference between frequency and wavelength?
Q: Why does your stove turn red when it heats up?
Heat generates light

- One of many ways light is produced:
- Maxwell: motion of charged particles creates EM field
- Thermodynamics: ...particles jiggle around!
- Hence, anything moving generates light
- In other words:
  - every object around you is producing color!
  - frequency determined by temperature
Most light is not visible!

- Frequencies visible by human eyes are called “visible spectrum”
- These frequencies what we normally think of as “color”
Natural light is a mixture of frequencies

- “White” light is really a mixture of all (visible) frequencies
- E.g., the light from our sun

![Spectrum of Solar Radiation (Earth)](image)
Additive vs. Subtractive Models of Light

- **Spectrum we just saw for the sun “emission spectrum”**
  - How much light is produced (by heat, fusion, etc.)
  - Useful for, e.g., characterizing color of a lightbulb

- **Another useful description: “absorbtion spectrum”**
  - How much light is absorbed (e.g., turned into heat)
  - Useful for, e.g., characterizing color of paint
Emission Spectrum

Describes light intensity as a function of frequency

Below: spectrum of various common light sources:

Figure credit: admesy
Emission Spectrum—Example

- Why so many different kinds of lightbulbs on the market?

- “Quality” of light:

  Incandescent:
  + more sun-like
  - power-hungry

  CFL:
  - “choppy” spectrum
  + power efficient
Absorption Spectrum

- Emission spectrum is intensity as a function of frequency
- Absorption spectrum is fraction absorbed as function of frequency

Q: What color is an object with this absorption spectrum?
This is the fundamental description of color: intensity or absorption as a function of frequency.

Everything else is merely a convenient approximation!
If you remember to use spectral description as a starting point, the issues surrounding color theory/practice will make a lot more sense!
Color reproduction is hard!

- Color clearly starts to get complicated as we start combining emission and absorption (real-world challenge!)

(What color ink should we use to get the desired appearance?)
...And what about perception?

Q: What color is this dress?
How does electromagnetic radiation (with a given power distribution) end up being perceived by a human as a certain color?
The eye

Image credit: Georgia Retina (http://www.garetina.com/about-the-eye)
The eye (optics)

Image credit: Georgia Retina (http://www.garetina.com/about-the-eye)
Photosensor response (eye, camera, …)

- Photosensor input: light
  - Electromagnetic power distribution over wavelengths: $\Phi(\lambda)$

- Photosensor output: a “response” … a number $R$
  - e.g., encoded in electrical signal

- Spectral response function: $f(\lambda)$
  - Sensitivity of sensor to light of a given wavelength
  - Greater $f(\lambda)$ corresponds to more an efficient sensor (when $f(\lambda)$ is large, a small amount of light at wavelength $\lambda$ will trigger a large sensor response)

- Total response of photosensor:

$$R = \int_{\lambda} \Phi(\lambda) f(\lambda) d\lambda$$
The eye’s photoreceptor cells: rods & cones

- **Rods** are primary receptors under dark viewing conditions (scotopic conditions)
  - Approx. 120 million rods in human eye
- **Cones** are primary receptors under high-light viewing conditions (photopic conditions, e.g., daylight)
  - Approx. 6-7 million cones in the human eye
  - Each of the three types of cone feature a different spectral response. This will be critical to color vision (much more on this in the coming slides)
Density of rods and cones in the retina

- Highest density of cones is in fovea
  (best color vision at center of where human is looking)
- Note “blind spot” due to optic nerve
ACTIVITY: Rods vs. Cones

- Need a brave volunteer from the audience!
  - Will hold up colored marker in peripheral vision
  - All you have to do is tell us what color it is (easy!)
Spectral response of cones

Three types of cones: S, M, and L cones (corresponding to peak response at short, medium, and long wavelengths)

\[ S = \int_{\lambda} \Phi(\lambda) S(\lambda) d\lambda \]

\[ M = \int_{\lambda} \Phi(\lambda) M(\lambda) d\lambda \]

\[ L = \int_{\lambda} \Phi(\lambda) L(\lambda) d\lambda \]

Uneven distribution of cone types in eye

~64% of cones are L cones, ~32% M cones
Response of S,M,L cones to monochromatic light

Figure visualizes cone’s response to monochromatic light (light with energy in a single wavelength) as points in 3D space

(plots value of S, M, L response functions as a point in 3D space)
The human visual system

- Human eye does not directly measure the spectrum of incoming light - i.e., the brain does not receive “a spectrum” from the eye
- The eye measures three response values = (S, M, L). The result of integrating the incoming spectrum against response functions of S, M, L-cones

Diagram:
- Spectrum $\Phi(\lambda)$
- Eye: Focuses light on retina, cones measure light (photopic case)
- Cone responses (S, M, L) carried along optic nerve
- Brain
Q: Is it possible for two functions to integrate to the same value?
Metamers

- Metamers = two different spectra that integrate to the same (S,M,L) response!
- The fact that metameters exist is critical to color reproduction: we don’t have to reproduce the exact same spectrum that was present in a real world scene reproduce the perceived color on a monitor (or piece of paper, or paint on a wall)
- ...On the other hand, combination of light & paint could still cause trouble—different objects appearing “wrong” under different lighting conditions.
Example: Counterfeit Detection

- Many countries print currency, passports, etc., with special inks that yield different appearance under UV light:
Ok, so color can get pretty complicated!

How do we encode it in a simple(r) way?
Color Spaces and Color Models

- Many other ways to specify a color
  - storage
  - convenience
- Instead, specify a color from some color space using a color model
- Color space is like artist’s palette: full range of colors we can choose from
- Color model is the way a particular color in a color space is specified:
  - artist’s palette: “yellow ochre”
  - sRGB color space: 204, 119, 34
Additive vs. Subtractive Color Models

- Just like we had emission & absorption spectra, we have additive and subtractive* color models
  - Additive
    - Used for, e.g., combining colored lights
    - Prototypical example: RGB
  - Subtractive
    - Used for, e.g., combining paint colors
    - Prototypical example: CMYK

*A better name than subtractive might be multiplicative, since we multiply to get the final color!
Let’s shed some light on this picture...
Other Common Color Models

- **HSV**
  - hue, saturation, value
  - more intuitive than RGB/CMYK
- **SML** — physiological model
  - corresponds to stimulus of cones
  - not practical for most color work
- **XYZ** — preceptually-driven model
  - Y captures luminance (intensity)
  - X,Z capture chromaticity (color)
  - related to, but different from, SML
- **Lab** — “perceptually uniform” modification of XYZ
Practical Encoding of Color Values

- How do colors actually get encoded digitally?
- One common encoding (e.g., HTML): 8bpc hexadecimal values*/:

  #1B1F8A

- What does this string mean? Common encoding of RGB.
- Want to store 8-bits per channel (red, green, blue), corresponding to 256 possible values
- Rather than use digits 0-9, use 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F
- Single character now encodes 16 values, two characters encode 16*16 = 256 values
- Q: Roughly what color is #ff6600?

*Upper vs. lowercase letters? Makes absolutely no difference!
Other Ways of Specifying Color?

- Other color specifications not based on continuous color space
- E.g., Pantone Matching System
  - industry standard (proprietary)
  - 1,114 colors
  - Combination of 13 base pigments
- And not to forget…
Why use different color models?

- **Convenience**
  - Is it easy for a user to choose the color they want?

- **Efficiency of encoding**
  - E.g., use more of numerical range for perceptually significant colors
  - Do color images compress well?
Example: Y’CbCr color model

- Common for modern digital video
- $Y'$ = luma: perceived luminance (same as $L^*$ in CIELAB)
- $Cb$ = blue-yellow deviation from gray
- $Cr$ = red-cyan deviation from gray
Original picture of Kayvon
Contents of CbCr color channels downsampled by a factor of 20 in each dimension (400x reduction in number of samples)
Full resolution sampling of luma ($Y'$)
Reconstructed result
(looks pretty good)
Original picture of Kayvon
Why use different color models? (cont.)

- **Convenience**
  - Is it easy for a user to choose the color they want?

- **Efficiency of encoding**
  - E.g., use more of numerical range for perceptually significant colors
  - Do color images compress well?

- **Gamut**
  - Which colors can be expressed using a given model?
  - Very different for print vs. display
Which raises a very important question:

Which actual colors (i.e., spectra) do these values get mapped to?
CIE 1931* Color Space

- Standard “reference” color space
- Encompasses all colors visible by “most” human observers
  - associated color model (XYZ) captures perceptual effects
  - e.g., perception of color (“chromaticity”) changes w/ brightness (“luminosity”)
  - different from specifying direct simulation of cones (SML)
  - ...lots more to say here!

*CIE 1931 does not mean anything important: “created in 1931 by the Commission Internationale de l’Eclairage"
sRGB Color Space

- CIE 1934 captured all possible human-visible colors
- sRGB (roughly) subset of colors available on displays, printers, ...
- Nonlinear relationship between stored RGB values & intensity
  - Makes better use of limited set of numerical values
Chromaticity Diagrams

- Chromaticity is the intensity-independent component of a color
- Chromaticity diagram used to visualize extent of a color space

A display with primaries with chromacities $P_1$, $P_2$, $P_3$ can create colors that are combinations of these primaries (colors that fall within the triangle)
Color Conversion

- Given a color specified in one model/space (e.g., sRGB), try to find corresponding color in another model (e.g., CMYK)
- In a perfect world: want to match output spectrum
- Even matching perception of color would be terrific (metamers)
- In reality: may not always be possible!
  - Depends on the gamut of the output device
  - E.g., VR headset vs. inkjet printer
- Complicated task!
- Lots of standards & software
  - ICC Profiles
Gamma correction

Old CRT display:

1. Image contains value $X$
2. CRT display converts digital signal to an electron beam voltage $V(x)$ (linear relationship)
3. Electron beam voltage converted to light: (non-linear relationship)
   $$Y \propto V(\gamma)$$
   Where: $\gamma \approx 2.5$

So if pixels store $Y$, what will the display's output look like?

Fix: pixels sent to display must store:
$$Y^{1/2.5} = Y^{0.4}$$

(Note: this effect does not apply to modern LCD displays, whose luminance output is linearly proportional to input values!)

Image credit: http://creativebits.org/mac_os_x/windows_vs_mac_monitor_gamma
Non-linear correction (for adaptation)

Goal: want viewer to perceive luminance differences as if they were present in the environment where a picture is taken (reproducing the absolute values of $Y$ is often not practical)

Example: TV camera records $Y$

Outdoor Scene

$Luminance$ from scene as measured arriving at pixels $p0$ and $p1$

$Y_0$

$Y_1$

Camera

$Y_0$

$Y_1$

Stored luminance values

$LCD$ Display

$Luminance$ output from LCD display at pixels $p0$ and $p1$

$cY_0$

$cY_1$

TV viewer (at home in dark room)

$\frac{cY_0}{cY_1}$

$\frac{Y_0}{Y_1}$

$\frac{cY_0}{cY_1} < 1$

(TV not as bright as outdoor scene)

But now consider ratio of perceived brightness of these two pixels:

Viewer in outdoor scene (bright adapted):

$\frac{Y_0^{0.5}}{Y_1^{0.5}}$

Not the same ratio due to differences in adaptation

Viewer in living room at night (dark adapted):

$\frac{cY_0^{0.4}}{cY_1^{0.4}} = \frac{Y_0^{0.4}}{Y_1^{0.4}}$

[Example credit: Marc Levoy]
Non-linear correction (for adaptation)

Goal: want viewer to perceive luminance differences as if they were present in the environment where a picture is taken (reproducing the absolute values of Y is often not practical)

Solution: TV camera stores $Y^{1.25}$

Outdoor Scene

$Y_0 \rightarrow$ Camera

$Y_0^{1.25}$

$Y_1^{1.25}$

Luminance from scene as measured arriving at pixels p0 and p1

Stored luminance values

LCD Display

$cY_0^{1.25}$

$cY_1^{1.25}$

Luminance output from LCD display at pixels p0 and p1

$c < 1$

(TV not as bright as outdoor scene)

Viewer in outdoor scene (bright adapted):

$\frac{Y_0^{0.5}}{Y_1^{0.5}}$

Relative brightnesses of TV's pixels are perceived to be similar to those in real scene!

Viewer in living room at night (dark adapted):

$\frac{(cY_0^{1.25})^{0.4}}{(cY_1^{1.25})^{0.4}} = \frac{Y_0^{0.5}}{Y_1^{0.5}}$

[Example credit: Marc Levoy]
Another reason for non-linear correction: quantization error

Consider 12-bit sensor pixel:
Can represent 4096 unique luminance values in output image

Values are \( \sim \) linear in luminance since they represent the sensor’s response
Problem: quantization error

Many common image formats store 8 bits per channel (256 unique values)
Insufficient precision to represent brightness in darker regions of image

Bright regions of image: perceived difference between pixels that differ by one step in luminance is small!
(human may not even be able to perceive difference between pixels that differ by one step in luminance!)

Dark regions of image: perceived difference between pixels that differ by one step in luminance is large!
(quantization error: gradients in luminance will not appear smooth.)

Rule of thumb: human eye cannot differentiate <1% differences in luminance
Store lightness, not luminance

Idea: distribute representable pixel values evenly with respect to perceived brightness, not evenly in luminance (make better use of available bits)

Solution: pixel stores $Y^{0.45}$
Must compute $(\text{pixel\_value})^{2.2}$ prior to display on LCD

Warning: must take caution with subsequent pixel processing operations once pixels are encoded in a space that is not linear in luminance.

e.g., When adding images should you add pixel values that are encoded as lightness or as luminance?
High-dynamic range images

- Problem: ratio of brightest object to darkest object in real-world scenes can be quite large
  - Human eye can discern ratio of 100,000:1 (even more if accounting for adaptation)

- High-dynamic range (HDR) image: encodes large range of luminance (or lightness) values
  - Common format: 16-bits per channel EXR (see environment maps in Asst. 3)

- Modern camera sensors can only sense much narrower range of luminances (e.g., 12-bit pixels)

- Modern displays can only display a much narrower range of luminances
  - Luminance of white pixel / luminance of black pixel for a high-end LCD TV ~ 3000:1 *

* Ignore most marketing specs, which are now claiming over 2,000,000:1
Tone mapping

- Tone mapping: non-linear mapping of wide range of luminances into a narrower range (for storage in low-bit depth image, or for presentation on a low-dynamic range display)
  - For examples see Debevec 1997, Reinhard 2002, Fattal 2002

- How to acquire HDR images with conventional camera?
  - Take photos at multiple exposures, combine into a single HDR image

Low-dynamic range images taken at multiple exposures (to capture detail in all regions)

Low-dynamic range image that is result of tone mapping HDR image
HDR “mode” in modern cameras

Kayvon’s iPhone
Focus on the central dot for some seconds. Look at a white wall or white sheet of paper. What do you see?
Key theme: exploit characteristics of human perception to build efficient image storage and image processing systems

- Separation of luminance from chrominance in color representations (e.g., Y’CrCb) allows reduced resolution in chrominance channels (4:2:0)

- Encode pixel values linearly in lightness (perceived brightness), not in luminance (distribute representable values uniformly in perceptual space)
Web links:

- You can find the color apps from Stanford here:
  - https://graphics.stanford.edu/courses/cs178-10/applets/locus.html
  - https://graphics.stanford.edu/courses/cs178/applets/colormatching.html

- The Nature article in which cone distribution is visualized is here:
  - http://roorda.vision.berkeley.edu/Pubs/ROORDA.PDF
What you should know:

- Why is it that we can get away with just three values for color (e.g., R-G-B)?
- What are the rods and cones, and how many types of each are there in the human visual system? What can you say about how they are distributed in the retina?
- Describe some of the different color spaces that are used to express color.
- An alternative to RGB color space is the CIE color space, with X, Y, and Z primaries. What is Y in this color space? What problem with RGB color space does the CIE color space solve?
- Given a color space expressed by some three-dimensional basis, can it be converted into any other basis through a linear operation? (True or False)
- What is gamma correction? Give an example where gamma correction is useful.
- What is high dynamic range imaging? How could you capture high dynamic range information for a real scene?
- What is tone mapping? Can you think of a good algorithm for tone mapping? What would you want to preserve?