Lecture Notes: Light and Color

Subhransu Maji

February 4, 2025

Contents

1 Overview

To understand color perception in the human eye, we must explore both the spectral properties of light and the physiological mechanisms of the eye, including the nature of the underlying sensors like the cone cells. Cone cells are sensitive to different wavelengths of light, roughly corresponding to red, green, and blue, and their combined responses form the basis of our color vision. However, color perception is not solely determined by the incoming light. It is a complex process influenced by factors such as the surrounding colors, lighting conditions, and even psychological and cultural contexts. For example, the same color can appear different when viewed against different backgrounds or under varying illuminations. This interplay between physical and contextual elements highlights the intricate nature of human color perception.

2 Spectral basis of light

For the purpose of this lecture, we will consider light as an electromagnetic wave for simplicity. The electromagnetic spectrum rages from high-frequency gamma rays, with frequencies greater than 10^{24} Hz (wavelength $< 10^{-16}$m), to long range radio waves with frequencies less than 1 Hz (wavelength $> 10^{8}$m).

Any light source can be fully described by its power spectrum $\mathrm{P}(\lambda)$, which represents the amount of energy emitted per unit time at each wavelength:

$$P(\lambda) = \text{Power at wavelength } \lambda \tag{1}$$

The human eye is sensitive to wavelengths ranging from 380 nm to 700 nm, which constitutes the visible part of the spectrum. Light with wavelengths greater than 700 nm, such as radio waves, and shorter wavelengths like ultraviolet (UV), X-rays, and gamma rays are invisible to the human eye.

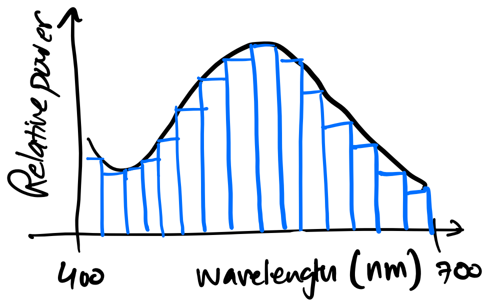

Figure 1: Spectra of a light source in the visible spectrum. The blue bars represent a discrete approximation of the continuous function which represents the relative power at each wavelength.

Therefore, to understand the color of a light source, it is sufficient to focus on the spectrum within the visible range (380–700 nm). However, these invisible wavelengths can still impact us, potentially causing damage to the eyes or skin tissue.

The solar radiation spectrum includes UV rays, visible light, and a significant amount of energy in the infrared region. Normal daylight typically has a relatively uniform power distribution across visible wavelengths. In contrast, tungsten lightbulbs emit more energy at longer wavelengths, while a ruby laser emits all its power at exactly 694.3 nm. Refer to the lecture slides to visualize the spectra of various light sources.

To compute the energy emitted per unit time between wavelengths λ_1 and λ_2, we have to integrate the power between the two values:

$$E = \int_{\lambda_1}^{\lambda_2} P(\lambda)\, d\lambda \tag{2}$$

For a discrete representation of power, wavelengths can be grouped into intervals or bins:

- Example bins: 60 nm bins with intervals 300-360 nm, 360-420 nm, etc.
- Smaller bins provide better approximations.

The energy emitted between two wavelengths can then be determined by summing the power values in the respective bins.

3 Light and Color Perception in the Human Eye

3.1 The Human Eye

The human eye contains several physical mechanisms for detecting and processing light. These include:

- **Adaptable lens:** Enables focusing on objects at varying distances.
- **Pupil:** The aperture (a hole) that allows light to enter the eye.
- **Iris:** The colored ring surrounding the pupil, containing radial muscles that control the pupil size.
- **Retina:** A light-sensitive layer that houses two types of photoreceptor cells:
  - **Rods:** Responsible for detecting brightness and intensity.
  - **Cones:** Specialized for color perception and come in three types (S, M, L) corresponding to short, medium, and long wavelengths.

Rods and cones are distributed unevenly across the retina:

- The **fovea** is a small region at the center of the visual field, about 1–2° in size. It has the highest density of cones but contains no rods.
- Each human eye contains approximately 5 million cones and 100 million rods.
- The ratio of L to M to S cones is roughly 10:5:1. Notably, there are almost no S cones in the center of the fovea.

3.2 Color Perception

Our perception of color arises from the combined responses of three types of cone cells in the retina. Each cone acts as a filter, mapping the light's power spectrum $P(\lambda)$ into three distinct values R_S, R_M, R_L:

$$R_S = \int P(\lambda) S_S(\lambda)\, d\lambda \approx \sum_{i=1}^{N} P(i) \times S_S(i) \tag{3}$$

$$R_M = \int P(\lambda) S_M(\lambda)\, d\lambda \approx \sum_{i=1}^{N} P(i) \times S_M(i) \tag{4}$$

$$R_L = \int P(\lambda) S_L(\lambda)\, d\lambda \approx \sum_{i=1}^{N} P(i) \times S_L(i) \tag{5}$$

Here $S_S(\lambda), S_M(\lambda), S_L(\lambda)$ denote the sensitivity at wavelength λ, $S_S(i), S_M(i), S_L(i)$ denotes the discrete approximation, and R_S, R_M and R_L denote the responses of the S, M and L cones respectively. This process compresses the vast amount of spectral information into just three numbers. Consequently, some spectral information is lost, and different spectra can appear identical to the human eye. Such indistinguishable spectra are called **metamers**.

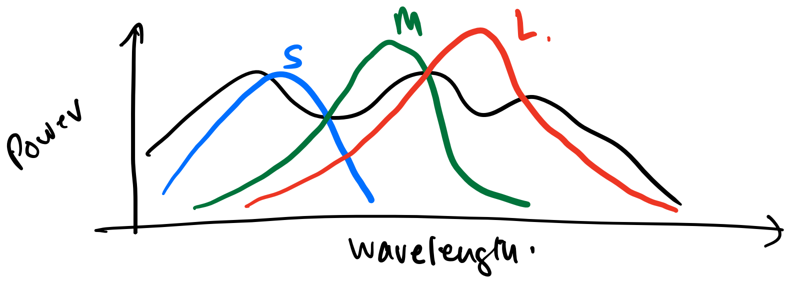

Figure 2: Caption

Example: Let F represent the power spectra of a colored flashlight, where each of the 10 values is the fraction of power produced in a 40nm range from 380nm to 780nm. (Note: some of the power can be in the non-visible spectrum so the values may sum up to less than one).

$F = [0.00, 0.00, 0.00, 0.00, 0.03, 0.08, 0.15, 0.32, 0.30, 0.12];$

For example, this flashlight produces 12 percent of its power in the range 740-780nm. Let $S_S(\lambda)$, $S_M(\lambda)$, and $S_L(\lambda)$ represent the relative absorption spectra of the cone cells in your eye. This means that, of the power absorbed by a given type of cone cell, the fraction absorbed in a given range is given by these numbers.

$S_S(\lambda) = [0.16, 0.26, 0.28, 0.15, 0.10, 0.03, 0.02, 0.00, 0.00, 0.00];$
$S_M(\lambda) = [0.00, 0.03, 0.06, 0.20, 0.31, 0.21, 0.15, 0.03, 0.01, 0.00];$
$S_L(\lambda) = [0.00, 0.00, 0.00, 0.00, 0.01, 0.04, 0.08, 0.23, 0.35, 0.29];$

When a flashlight is 5 meters from a white screen, assume that it stimulates a cone cell response (relative to the maximum possible response from that cone) given by:

$$R_c = \sum_{i=1}^{10} F(i) \times S_c(i).$$

Here F is flashlight, S_c is the absorption spectrum for cones of type c, and R_c is the response for cone cell type c. For example, from 5 meters away, the flashlight F will generate a response from the green (M) cone cells of

$$R_M = \sum_{i=1}^{10} F(i) \times S_M(i) \approx 0.0612$$

Exercise: Compute the responses for the S and L cones. (Answer:[1])

3.3 Interesting Facts about Color Vision

- The pigments for M and L cones are encoded on the X chromosome. This explains why men are more likely to be colorblind.
- Variations in the L pigment gene can lead to some women being **tetrachromatic**, meaning they may perceive an additional dimension of color.
- Color Blindness
  - **Red-green color blindness:** Caused by mutations in the L or M photoreceptors, leading to difficulty distinguishing red and green hues.
  - **Blue-yellow color blindness:** Caused by mutations in S photoreceptors, leading to challenges in distinguishing blue-green and yellow-red hues.
- Animals have varying numbers of cone types depending on their visual needs:
  - **One cone type:** Found in nocturnal animals.
  - **Two cone types:** Found in dogs.
  - **Four cone types:** Found in fish and birds.
  - **Five cone types:** Found in pigeons and some reptiles/amphibians.
  - **Twelve cone types:** Found in mantis shrimp, allowing them to perceive an extraordinary range of colors.

4 Tristimulus Theory and Color Spaces

The nature of color perception in the human eye suggests that three numbers might be sufficient for encoding color. In fact, this observation dates back to the 18th century (Thomas Young). We can also test this hypothesis through color-matching experiments, where the goal is to match the color of a light source by adjusting the power of three primary light sources. Most people can match any given light using three primaries, provided that the primaries are independent. For the same light and the same primaries, most people select the same weights. Color matching also appears to be linear, which were first described in Grassmann's laws (1853)[2].

1. If two test lights can be matched with the same set of weights, then they match each other:

$$\text{Suppose } A = u_1 P_1 + u_2 P_2 + u_3 P_3 \quad \text{and} \quad B = u_1 P_1 + u_2 P_2 + u_3 P_3. \quad \text{Then } A = B.$$

2. If we mix two test lights, then mixing their matches will also match the result:

$$\text{Suppose } A = u_1 P_1 + u_2 P_2 + u_3 P_3 \quad \text{and} \quad B = v_1 P_1 + v_2 P_2 + v_3 P_3.$$

$$\text{Then } A + B = (u_1 + v_1) P_1 + (u_2 + v_2) P_2 + (u_3 + v_3) P_3.$$

3. If we scale the test light, then the matches scale by the same amount:

$$\text{Suppose } A = u_1 P_1 + u_2 P_2 + u_3 P_3. \quad \text{Then } kA = (ku_1) P_1 + (ku_2) P_2 + (ku_3) P_3.$$

[1] $R_S \approx 0.0084$ and $R_L \approx 0.2289$

[2] https://en.wikipedia.org/wiki/Grassmann%27s_laws_(color_science)

4.1 Computing Primaries for any Light

How can we compute the weights of the primaries to match any spectral signal? One way to do this is to do a color matching experiment. However this is cumbersome. We would like to develop a mathematical way to estimating the weights given any spectral signal. Fortunately there is a way to do this by exploiting linearity of light! The idea is to think of light of a given spectrum to be a linear combination of monochromatic lights whose spectra is a delta function at a given wavelength, with linear coefficients given by the power at each wavelength. The figure below illustrates this:

Say that a monochromatic light of wavelength λ_i will be matched to $c_1(\lambda_i), c_2(\lambda_i), c_3(\lambda_i)$ of each primary p_1, p_2, p_3. We can run color matching experiments for each λ_i and store the color matching function in the rows of the matrix C given by:

$$C = \begin{pmatrix} c_1(\lambda_1) & \dots & c_1(\lambda_N) \\ c_2(\lambda_1) & \dots & c_2(\lambda_N) \\ c_3(\lambda_1) & \dots & c_3(\lambda_N) \end{pmatrix}$$

Let the given spectral signal be described by the vector t

$$t = \begin{pmatrix} t(\lambda_1) \\ \vdots \\ t(\lambda_N) \end{pmatrix}$$

Then the amounts of each primary needed to match t is given by $e = Ct$. The components e_1, e_2, e_3 describe the color of t. This follows from the linearity of light.

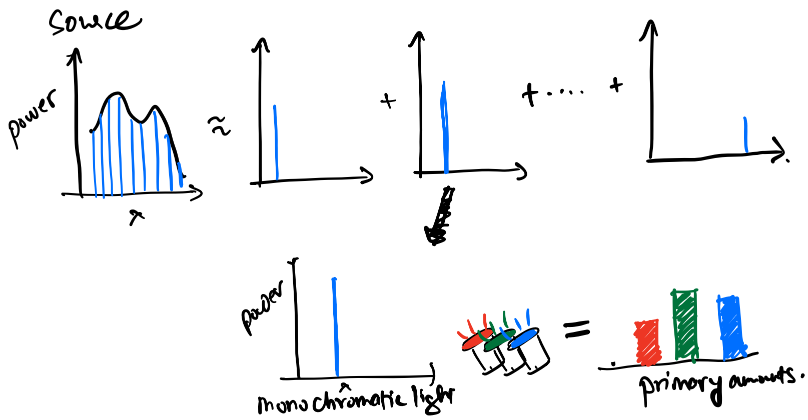

Figure 3: A light source can be thought of a linear combination of monochromatic lights with weights given by the spectral power (top). Thus the primary amounts of the light source can be computed as the weighted combination of the primary amounts of each monochromatic light (below)

4.2 Color Spaces

The RGB color space is based on primaries corresponding to red, green, and blue lights, which are typically emitted by LEDs or phosphors in monitors. The RGB color space cannot represent all visible colors, as some colors require negative weights in the matching process. These emitters are so densely packed that the human eye cannot distinguish them individually. Instead, what we perceive is the combination of several emitters within a neighborhood corresponding to a pixel. Thus, by changing the brightness of these emitters, we can alter the effective color of the pixel.

The RGB color space is not uniform in perceptual space. In other words, differences in RGB values do not necessarily correspond to perceptual color differences. The CIE uv color space is a transformation designed to better align with perceptual distance.

The HSV color space (Hue, Saturation, Value/Intensity) maps the RGB color space to dimensions that are more perceptually meaningful. Hue corresponds to the color, as represented on a color wheel. Saturation represents the vividness or purity of the color. Value corresponds to the overall brightness of the color. This transformation effectively reorients the RGB cube, aligning it along its vertex to make these dimensions easier to interpret.

There are many other color spaces designed for the specific application, device or industry needs. Some examples are:

- Lab – Similar to CIE uv. Represents lightness (L), red-green (a), and blue-yellow (b) components. Used in image processing and color difference calculations.
- CMY/CMYK (Cyan, Magenta, Yellow, Black) – Subtractive color model used for printing.
- YUV – Used in video compression; separates luminance (Y) from chrominance (UV).
- YCbCr – Derived from YUV; used in JPEG and MPEG compression standards.

The transformation from one color space to another is typically non-linear and sometimes requires calculating an intermediate color. In Python, you can apply color transformations using libraries like OpenCV or scikit-image. Conversion formulas can be looked up on Wikipedia (you don't need to memorize them!).

5 Color Phenomenon

5.1 Color Constancy

This is the ability of the human visual system to perceive color as relatively constant despite changes in illumination conditions. The perceived color results from the interaction between the light source's spectrum and the surface's reflectance. To a rough approximation, the fraction of power reflected at wavelength λ is given by: $P_{\text{reflected}}(\lambda) = R(\lambda) \cdot P(\lambda)$, where $R(\lambda)$ is the reflectance at λ.

For example, a tomato looks red in sunlight because it reflects more red light, while grapes appear purple because they reflect both blue and red light. This also means the color of a surface is influenced by the light illuminating it. For instance, under monochromatic yellow light, every surface will appear as a shade of yellow.

The human brain is adept at disentangling the effects of illumination on a surface to estimate its reflectance (or its "true" color). This means color perception depends on context and can lead to phenomena such as color constancy, the checkerboard illusion, and the controversy about the dress color.

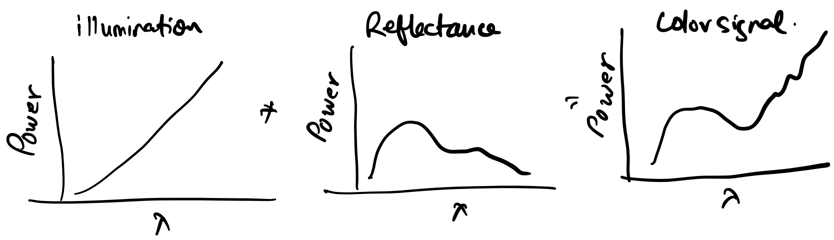

Figure 4: The spectrum of the light reflected from a surface is the product of the illumination and the surface reflectance. Thus the color of a surface depends both on the surface properties as well as the light that it is under.

5.2 Chromatic Adaption

The human visual system exhibits remarkable adaptability, adjusting its sensitivity based on the prevailing luminance in the visual field, although the exact mechanisms behind this process remain poorly understood. One way the eye adapts to different brightness levels is by changing the size of the iris opening, or aperture, which regulates the amount of light entering the eye. For example, when transitioning from bright sunlight into a dimly lit building, the iris contracts to allow less light in, helping the eye adapt to the new lighting conditions.

Similarly, the visual system adapts to different color temperatures by altering the sensitivity of the receptive cells in the retina. When exposed to an increased amount of red light, the cells sensitive to red gradually reduce their responsiveness, restoring the perception of a balanced scene where white appears as it should. Interestingly, this color adaptation is more effective in brighter environments, which explains why scenes illuminated by candlelight retain their yellowish hue, yet still appear natural to us.

5.3 White Balance

When viewing a picture on a screen or in print, our eyes adapt to the room's lighting rather than the lighting depicted in the image. If the white balance is incorrect, the image may display an unnatural color "cast."

White balancing is the process of adjusting the colors in an image to make them appear natural and accurate under varying lighting conditions. The primary goal of white balancing is to ensure that objects that are white in real life are rendered as white in the image, regardless of the light source's color temperature. For example, a photograph taken under tungsten light may appear overly orange. White balancing corrects this by reducing the orange tone, restoring a natural appearance where whites look white and other colors appear accurate.

Von Kries adaptation proposes that the visual system compensates for these changes by scaling the responses of the cone cells independently, allowing colors to appear relatively constant. This principle can be applied to correct white balance. One of the best ways is to record the color of the light using color cards—surfaces with known colors. By photographing these cards under the light source, we can infer the effect of the light by comparing the camera's measurements with the true colors of the card. For example, if a neutral object with a true color of r=g=b, is recorded as r_w, g_w, b_w then the effect of light is can be undone by multiplying the color channels with weights $1/r_w, 1/g_w, 1/b_w$.

When color cards are unavailable, alternative techniques can be used to estimate white balance, including:

1. **Gray World Assumption**: This method assumes that the average color of the image is gray. The image's average red (r_{ave}), green (g_{ave}), and blue (b_{ave}) values are calculated, and the color channels are adjusted using weights $1/r_{\text{ave}}, 1/g_{\text{ave}}, 1/b_{\text{ave}}$.

2. **Brightest Pixel Assumption**: This method assumes that the brightest pixels in the image typically represent highlights and carry the color of the light source. The weights for the color channels are set inversely proportional to the values of the brightest pixels.

3. **Gamut Mapping**: The gamut of an image (the convex hull of all pixel colors) is compared to the gamut of a "typical" image under white light. A transformation is then applied to match the image's gamut to the standard gamut, correcting the white balance.

4. **Natural Image Statistics**: Statistical models of natural images can be used to predict and correct white balance based on expected color distributions.

Modern cameras achieve white balancing through a combination of hardware sensors, machine learning algorithms, and software processing, ensuring that images and videos appear natural under various lighting conditions. For instance, iPhones are equipped with ambient light sensors that measure the color temperature and intensity of the surrounding light, while computational photography and machine learning algorithms enhance color accuracy. However, these algorithms can occasionally misfire. During wildfires in California, many people noted on Twitter that green traffic lights appeared blue in photos and videos. This happened because cameras attempted to correct the orange cast of the sky by boosting the blue and green channels, leading to this unintended color shift.

6 Optional Readings

Color and Language In their seminal work *Basic Color Terms: Their Universality and Evolution*, Berlin and Kay (1969) explored the relationship between color perception and language across different cultures. They proposed that all languages evolve basic color terms in a predictable sequence, starting with black and white, then adding red, and progressively incorporating other colors such as green, yellow, blue, brown, and more. This universality suggests that human color perception is biologically rooted, but the way colors are categorized and named is influenced by cultural and linguistic factors.

Colorization Converting a grayscale image to color, often referred to as colorization, involves adding plausible color information to an image that originally contains only intensity values. This process is used in applications such as restoring old photographs, enhancing scientific imaging, and improving visual appeal in media. Early methods relied on manual techniques, where artists applied colors to specific regions, guided by the image's content. Modern approaches leverage machine learning and deep neural networks to automate the process (and have in fact shown that colorization is a effective way to teach deep networks a good deal of visual information without requiring labels). These models are trained on large datasets of paired grayscale and color images, learning to infer contextually appropriate colors for objects, textures, and scenes. Advanced algorithms also use hints or user-provided inputs to improve accuracy and customization, ensuring the colorized output appears natural and visually consistent.