

Lecture 17: Color

1 History of Color

You already know that wavelengths of light have different colors. For example, red light has $\lambda \approx 650\text{nm}$, blue light has $\lambda \approx 480\text{nm}$ and purple light has $\lambda \approx 420\text{nm}$. But there is much more to color than pure monochromatic light. For example, why does mixing red paint and blue paint make purple paint? Surely monochromatic plane waves can't suddenly change wavelength, so what is going on when we mix colors? Also, there are colors (like cyan or brown) which do not appear in the rainbow. What wavelengths do they have? As we will see, the wavelength of monochromatic light is only the starting point for thinking about color. What we think of as a color depends on the way our brains and our visual system processes the light coming into our eyes.

The earliest studies of color were done by Newton. He understood that white light was a combination of lots of wavelengths. He performed some ingenious experiments to show this, described in his book *Optiks* (1704). For example, even though prisms turned white light into a rainbow, people thought that maybe the prism was producing the rainbow colors. Newton showed that this wasn't true. He could use lenses and mirrors to shine only red light into a prism, then only red light came out. He also separated white light into different colors, then used lenses and mirrors to put them into another prism which made white light again.

In his classic *Optiks*, Newton compared the colors going in a circle to the cycle of 5ths in music. His color wheel was

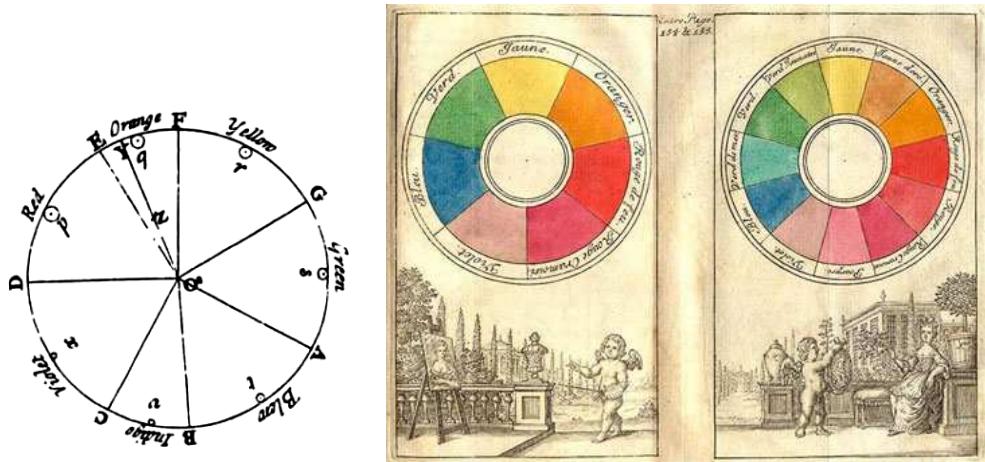


Figure 1. Newton's color circle 1704 and a contemporary rendition by Boutet (1708)

If the spectrum of visible light spans wavelengths from 350 nm to 750 nm, why should colors make a circle? Read on...

The next big advance came in 1853 by the mathematician Hermann Grassmann (Grassmann is well known in mathematical physics for his work on anti-commuting numbers which are now used to describe spinors like the electron in quantum mechanics). Grassmann was intrigued by the idea that two colors could mix to produce a different color, such as red+blue=purple. He showed that when you have an equation like this, you can add any color to either side, and it will still be a match, for example, red+blue+yellow=purple+yellow. This is known as Grassmann's law, which we will explore more quantitatively in the next section.

Another great contributor to our understanding of color was James Clerk Maxwell. He did a cool demonstration where he took black and white photos of the same colorful object using red, green and blue filters. Then when he projected those photos back again, through the same filters, the original multicolored object could be seen (this is still the way that color photography works, both film-based and digital). The amazing thing was that it wasn't just the red, blue and green colors that showed up, but the oranges and yellows and purples as well – all the colors. How is that possible? If the filters just let through pure red, blue, and green wavelengths, then none of the orange wavelengths would pass through. So how could they be reproduced? To understand this, we need to understand the perception of color.

2 Combining color

The modern theory of color was not laid down until 1931 by a set of classic experiments by W. D. Wright and John Guild in France, which built on the insights of many other people, including Newton, Grassmann and Maxwell. Their work lead to the International Commission on Illumination (CIE) and their now standard color space.

What Wright and Guild did was to systematically quantify when two colors made up of different frequencies look the same. For example, if I show you some combination of red light ($\lambda = 650\text{nm}$) and blue light ($\lambda = 480\text{nm}$), it looks the same as pure violet light ($\lambda = 420\text{nm}$):

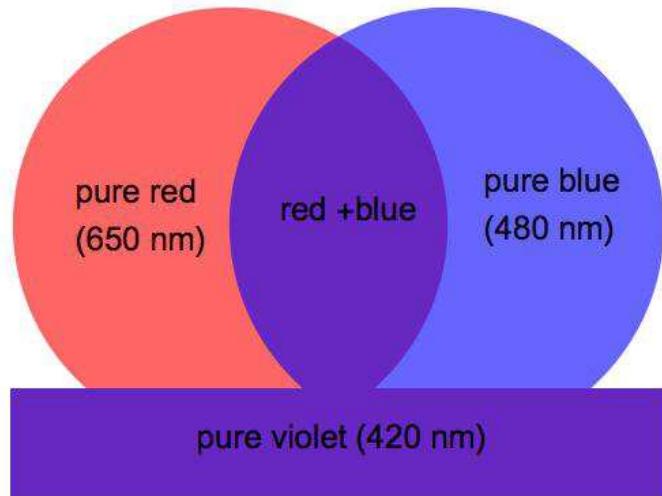


Figure 2. Mixed colors can look indistinguishable from a pure wavelength.

What Wright and Guild did was show people combinations of different wavelengths of light (the basis colors) and ask which pure wavelength of light it looked the same as. When two spectra of monochromatic light look the same, the colors are called **metamers**.

How many pure basis wavelengths do we need to combine to produce every color? Well, if we use just one color, like green, all we can do is vary the intensity – we always have green. So we need at least 2 colors. With two colors, the answer is not obvious. In principle, two colors could have worked. Two colors do work for colorblind people, and for most animals. But it's kind of obvious that 2 won't work for people, since by mixing red and blue we can only get shades of purple, never something like green. So Wright and Guild tried 3.

Wright and Guild showed people combinations of red (650 nm), blue (480 nm) and green (530 nm) and asked how much of the three colors looked the same pure light of wavelength λ . For example, they found that light of wavelength $\lambda = 420\text{ nm}$, which is a violet, can be matched with 0.97 parts blue, 0.20 red and 0.02 green. Their results are summarized in three functions $r(\lambda)$, $g(\lambda)$ and $b(\lambda)$, representing the intensities of red, green and blue light they needed to match light of a pure frequency λ . The functions look like this:

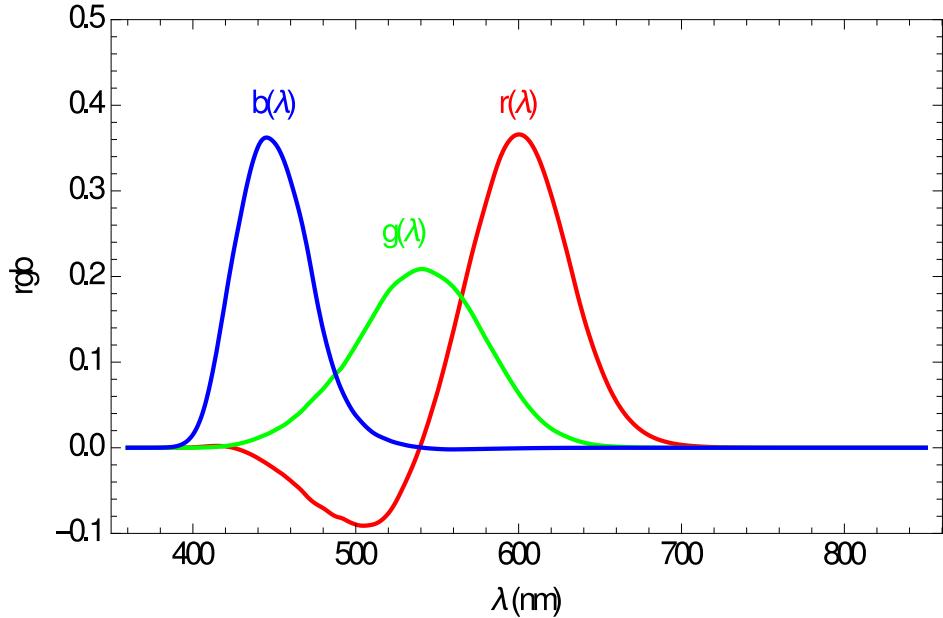


Figure 3. Original Wright and Guild RGB functions. These curves encode how we perceive color.

You should notice immediately something weird about these curves: the red one is less than zero in places. What does that mean? Well, in reality, Wright and Guild found that some colors could not be matched by mixing red, green and blue, for example, yellow with $\lambda = 580$. They got around this using Grassmann's law:

- **Grassmann's Law:** If two colors are indistinguishable (metamers), you can add any other color to them and they will still be metamers.

This essentially says that our color perception is linear. For example, Wright and Guild found that

$$I_{\text{violet}} \cong 0.97 I_{\text{blue}} + 0.20 I_{\text{red}} + 0.02 I_{\text{green}} \quad (1)$$

The “ \cong ” sign in this equation means the colors on the two sides are metamers. Violet on the left hand side here means a monochromatic violet, $\lambda = 420\text{nm}$, while the right hand side is a combination of 3 pure wavelengths. Since metamerism is a linear equivalence, Grassmann's law says that if we add more red, they will still look identical. Then

$$I_{\text{violet}} + 0.3 I_{\text{red}} \cong 0.97 I_{\text{blue}} + 0.50 I_{\text{red}} + 0.02 I_{\text{green}} \quad (2)$$

This is obvious when we write it as an equation like this, but it is very non-trivial.

One thing the linearity lets us do is take a color like yellow, which has $\lambda = 580\text{nm}$, and no matching to positive values of R, G, and B, and add to it something else, like violet to make the matching positive. For example, if

$$I_{\text{yellow}} + I_{\text{violet}} \cong 1.3 I_{\text{blue}} + 0.1 I_{\text{red}} + 0.3 I_{\text{green}} \quad (3)$$

we conclude that

$$I_{\text{yellow}} \cong 0.33 I_{\text{blue}} - 0.1 I_{\text{red}} + 0.28 I_{\text{green}} \quad (4)$$

This has a negative coefficient for red, just like in Wright and Guild's curves.

To summarize, Grassmann, Wright and Guild's observations were that our perception of light is a *linear system*. We can study perception with linear algebra. To repeat the “ \cong ” sign in the above equation indicates metmeric equivalence: “the combinations of pure wavelengths on the two sides are perceived identical by human beings”.

Negative intensities are fine mathematically, but what happens physically? Since we can see all the colors of the rainbow, are we somehow seeing negative intensities? Remarkably, the answer seems to be yes. Somehow we appreciate negative intensities as a result of higher-level cognitive processing. If this sounds nuts to you, try looking at

- Lilac Chaser: http://www.michaelbach.de/ot/col_lilacChaser

We will have more discussion of color *perception* in Sections 7 and 8 below.

3 CIE color space

Negative intensities are great mathematically, but not so practical for real world applications – you can't have an RGB monitor with negative intensities. But it would still be useful to have a representation of the total amount of color we can see, and to know which of those colors can be reproduced by a given RGB basis. A French committee called the International Commission on Illumination (CIE) got together in 1931 and came up with a very useful representation of the Wright and Guild data called the CIE color space.

The CIE commission realized that no three monochromatic colors could be mixed in positive amounts to match the full spectrum. However, they found that the could take a basis XYZ which are *sums* of monochromatic colors which would work. To make this precise, it is helpful first to write the Wright and Guild RGB values as a vector

$$\vec{R}(\lambda) = (r(\lambda), g(\lambda), b(\lambda)) \quad (5)$$

We can think of $\vec{R}(\lambda)$ as a parametrized 3-dimensional curve. The new basis will be related to \vec{R} by a linear transformation

$$\vec{X}(\lambda) = M \cdot \vec{R}(\lambda) \quad (6)$$

with M a matrix. The transformation chosen by the CIE is

$$\vec{X}(\lambda) = \begin{pmatrix} X(\lambda), Y(\lambda), Z(\lambda) \end{pmatrix} = \frac{1}{0.17} \begin{pmatrix} 0.49 & 0.31 & 0.20 \\ 0.17 & 0.81 & 0.01 \\ 0.00 & 0.01 & 0.99 \end{pmatrix} \vec{R}(\lambda) \quad (7)$$

This particular transformation is more-or-less arbitrary. It is chosen so that $X(\lambda)$, $Y(\lambda)$ and $Z(\lambda)$ are always positive. Lots of other choices would have worked just as well.

The functions $X(\lambda)$, $Y(\lambda)$ and $Z(\lambda)$ are called the **tristimulus curves** or **CIE curves**. They look like this

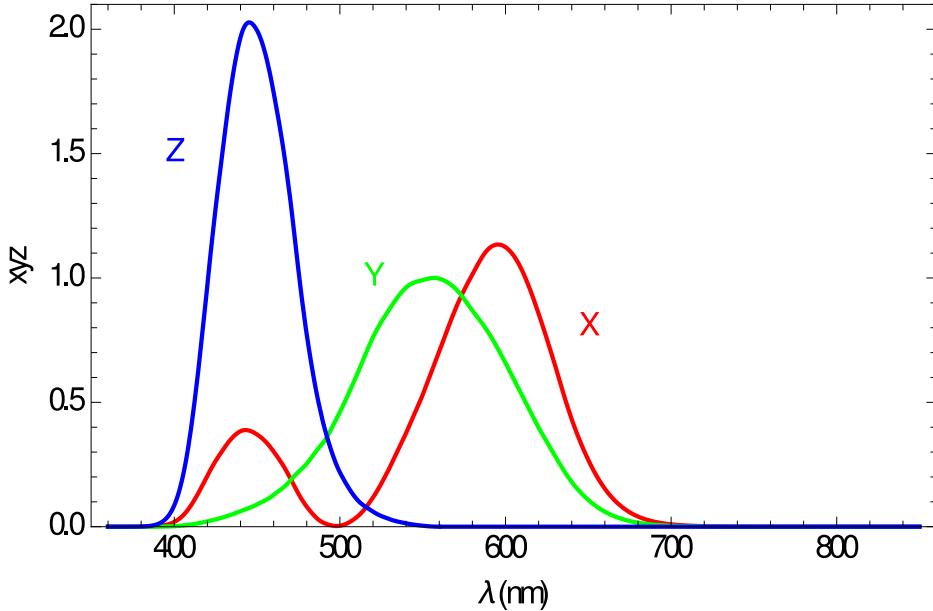


Figure 4. Standard tristimulus curves (CIE 1931).

As you can see from the figure any pure wavelength light will look identical to us as linear combination of X, Y, Z with positive coefficients. As we will see, tristimulus curves are essential to the design of film, printers, monitors, cameras, etc.

Note the small bump in the red tristimulus curve, and around 450 nm. This tells that violet looks like red + blue. It also explains the “line of purples”, which goes between red and purple and why Newton’s circle is a circle, as opposed to a line. Perhaps “explain” is too strong of a word. Of course, the CIE curves don’t actually explain why we see things that way – we could have just seen the rainbow, with red not looking anything like purple like a robot does. But somehow our mind wants to make a circle out of it. Maybe there’s an evolutionary explanation for this, I don’t know.

3.1 Normalizing the intensity

Because of the linearity, there is a well-defined notion of intensity of light: just sum the intensities in all the frequencies. In terms of the representation of a single frequency in terms of the CIE X, Y and Z values, this says that

$$I = \text{Intensity} = X + Y + Z \quad (8)$$

Then we only need two numbers to describe all the colors we can see. We call these two numbers lowercase x and y :

$$x = \frac{X}{X + Y + Z}, \quad y = \frac{Y}{X + Y + Z} \quad (9)$$

We can also define $z = \frac{Z}{X + Y + Z}$, but $z = 1 - x - y$ so it is not independent. Each frequency of pure light gives a value of $x(\lambda)$ and $y(\lambda)$. Plotting these curves gives a contour:

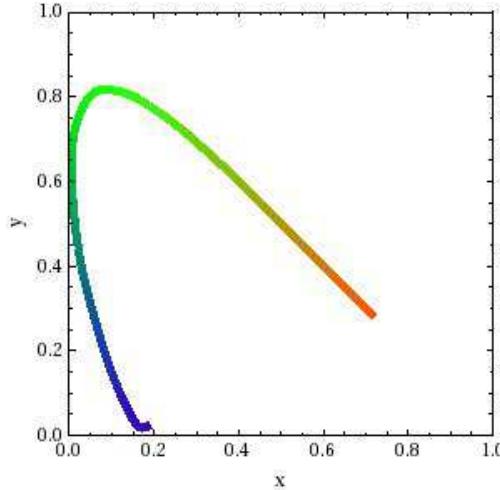


Figure 5. CIE curve: x and y values for **pure monochromatic light**, with $z = 1 - x - y$. That is, the colors of the rainbow fall along this curve.

Here’s where it gets interesting. Remember, the Wright and Guild experiments determined how much of their red, green and blue were needed to match a pure frequency λ . What happens when you mix just red and green? Or two other pure colors? Since the intensities are additive, we know that if we add two colors, the result would still match. For example, suppose we added some yellow light, with 0.3 intensity. Then we would find

$$I_{\text{violet}} + 0.3I_{\text{yellow}} \cong 0.97I_{\text{blue}} + 0.20I_{\text{red}} + 0.02I_{\text{green}} + 0.3I_{\text{yellow}} \quad (10)$$

The two sides of this equation show that two weird linear combinations of different colors are the same, even though neither side is monochromatic. What is this color? It's not monochromatic, so it's not in the rainbow.

Whatever this color is, since we can perceive it, there will be values of x , y and z which will match the perception. Which values? We can read them off of Fig. 5. Say yellow has a wavelength of 600nm and violet 420nm. Then the new x , y values will be just the linear combination of the violet and yellow above

$$x_{\text{mixed}} \cong \alpha x_{\text{violet}} + (1 - \alpha)x_{\text{yellow}}, \quad y_{\text{mixed}} = \alpha y_{\text{violet}} + (1 - \alpha)y_{\text{yellow}} \quad (11)$$

These are the equations for a line. Thus this color is somewhere on the line connecting yellow and violet on the curve in Fig. 5. Working out all the linear combinations, we get the full CIE color space, which is the gamut of colors we can perceive. It looks like this

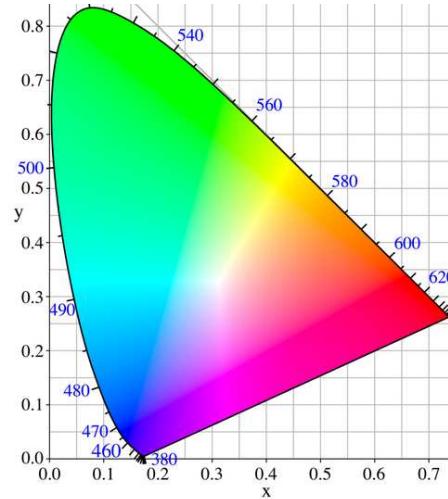


Figure 6. Approximate color space for the “CIE standard observer”. The wavelengths of monochromatic colors are marked on the boundary. The vertical line is a limitation of this RGB image.

For example, using the CIE functions, color in Eq. (10) corresponds to $x = 0.46$ and $y = 0.22$, which is a pink.

A very important point here is that if you start with any three monochromatic colors, corresponding to three points on the boundary of this curve, the **gamut** of colors you will be able to produce with that basis will be just a triangle. For example, here are two triangles using the RGB color basis of TVs and computer monitors (left) and film (right):

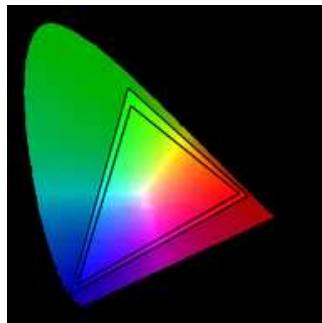


Figure 7. CIE for computer monitors.

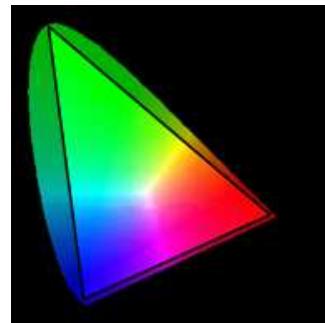


Figure 8. CIE triangle for film.

As we observed before, any monochromatic color can be made from any basis of 3 pure colors if negative intensities are allowed. If only positive intensities are allowed, then you are stuck in a triangle. What are the best colors we can choose? Those would be the colors that give the biggest triangle. This is pretty close to what they use for film. For computer monitors, the engineering requirements force them to use colors for which there are cheap LEDs, which gives a smaller triangle. In any case, we can see **the best basis will be red, green and blue**. A basis of red, yellow and blue, like Newton wanted, would get many fewer colors. In particular, RYB would have trouble making green.

Note that the “colors” X, Y, and Z correspond to the points $(x, y) = (1, 0)$, $(x, y) = (0, 1)$ and $(x, y) = (0, 0)$ respectively. Thus the triangle the CIE colors make is the bottom-left half of Fig. 6. This contains all the visible colors. It also contains lots of other linear combinations with negative intensities that don’t correspond to anything we can perceive at all. We can’t make a monitor with X, Y and Z LEDs, since X, Y and Z are not colors. The best we can do with actual colors is draw a triangle touching the color-space boundary. The point of inventing X, Y and Z is it lets us embed all perceivable colors in a 2D triangle.

4 Using the CIE values

As an example application, we can compute what color the sky is. In Lecture 16, we showed that the intensity of scattered light depends on wavelength like $I(\lambda) = c \frac{1}{\lambda^4}$ for some constant c . If you put all these scattered colors together with these relative intensities, what color do you get? By Grassmann’s linearity principle, the net color off such a signal would be equivalent to amounts of pure red, green and blue given by

$$r = \int r(\lambda) I(\lambda) d\lambda, \quad g = \int g(\lambda) I(\lambda) d\lambda, \quad b = \int b(\lambda) I(\lambda) d\lambda \quad (12)$$

Equivalently, we can use the standard x, y, z values to compute the integrals, then convert to RGB using the transfer matrix in Eq. (7) above. Performing the integrals, then normalizing the intensities so that $r + g + b = 1$, we find

$$(r, g, b) = (0.117, 0.295, 0.587) \quad (13)$$

This looks like, with various levels of saturation with white for cloud color (saturation is discussed in the next section):



Figure 9. Sky color calculated using CIE observer functions

So, the simple answer to why sky is blue is that the formula $I(\lambda) = \frac{1}{\lambda^4}$ is largest at smallest wavelengths, which is on the blue end. The more sophisticated answer is that we can figure out *what* shade of blue it is by integrating against the CIE curves. To be more careful, we should also include the fact that the incoming light is not completely flat in intensity because the sun has a spectrum.

To work out the color of the sun, we use that it is a blackbody at 5000 K. Blackbody radiation is thermal: it comes from highly excited molecules in a heat bath with their surroundings. This bath contains photons which are also thermal. The amount of each photon depends on the ratio of the photon energy $h\nu$ to the thermal energy kT . A beautiful result from quantum statistical mechanics is that the intensity of radiation at a wavelength λ from an object at temperature T is

$$I(\lambda) = \frac{2hc}{\lambda^5} \left(e^{\frac{hc}{\lambda kT}} - 1 \right)^{-1} \quad (14)$$

where h is Planck's constant, k is Boltzmann's constant and c is the speed of light. These blackbody curves look like

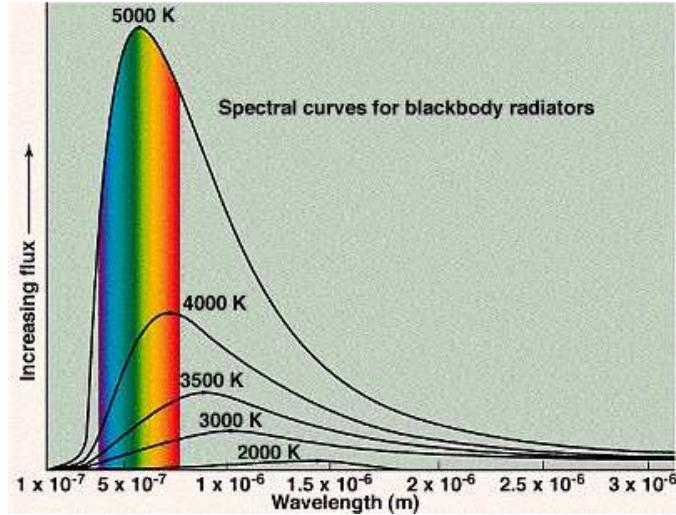


Figure 10. Blackbody radiation curves for different temperature. The sun is around 5000 K, incandescent light bulbs are around 3000 K.

Note that the sun does not have much intensity in violet, which partially explains why the sky is blue and not violet (the main reason, however, is that the λ^{-4} Rayleigh spectrum pushes the peak intensity away from violet).

A useful feature of the blackbody spectrum is that the wavelength where the spectrum is maximal scales inversely with the temperature, as given by **Wein's displacement law**:

$$\lambda_{\max} = \frac{2.898 \text{ mm K}}{T} \quad (15)$$

Plugging in room temperature $T = 300 \text{ K}$ gives $\lambda_{\max} = 9.6 \mu\text{m}$, which is in the infrared. That's why people show up in infrared cameras. Plugging in $T = 5000 \text{ K}$ gives $\lambda_{\max} = 501.4 \text{ nm}$ which is green.

To find the color of the sun, we integrate the Planck blackbody distributions against the CIE color functions, as in Eq. (12). The result is basically white. The reason the sun looks yellow is because blue light is scattered away from the sun by the sky. In fact, we can figure out the colors of the sun and sunset by subtracting the sky spectrum from the sun spectrum. Doing this, gives:

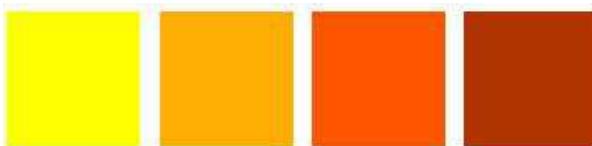


Figure 11. Sunset colors – sun color with scattering from the sky subtracted out.

More generally, each temperature translates to a particular color when the blackbody spectrum is integrated against the CIE curves. So as we vary the temperature, we trace out a curve of colors, or equivalently a path in the CIE color space. This curve has a name: the **Planckian locus**. It looks like this

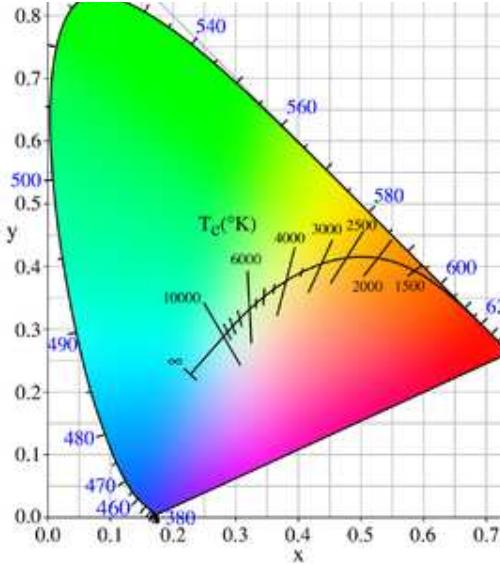


Figure 12. Color of blackbody radiation: The line is the **Planckian Locus**.

If you look at the color along this locus, you get something like

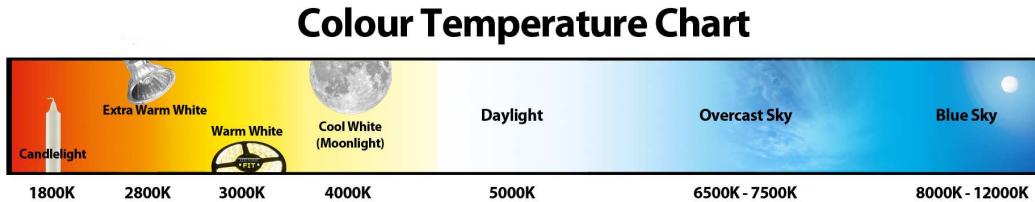


Figure 13. Colors along the Planckian locus and associated temperatures. The sun at 5700K is white.

This says that the hotter a color the more blue it is, the cooler, the more yellow or red, and in the middle it's white. Red stars, like red giants such as Betelgeuse (Orion's shoulder) are cooler, Yellow stars, like our sun are next, then white stars, such as white dwarfs, then blue stars. Sirius, the dog star, is the brightest star in the sky and is whitish-blue in color.

5 RGB and HSV

If all the colors of the rainbow are on the boundary of the CIE shape, what are all the colors in the middle? Our gamut includes a lot more than monochromatic colors.

We saw that we have a 3-dimensional basis of color perception – every color we can see can be represented by x , y , z values, representing (roughly) the intensities of red, green and blue light. One linear combination gives something equivalent to a pure monochromatic rainbow color (called **hue**). This corresponds to moving around the boundary CIE curve. One linear combination is the intensity (called **value**). This is normalized to 1 in the CIE color space. What is the remaining degree of freedom? We call it **saturation**. It is the amount of white a color has. It corresponds to moving inwards from the CIE boundary towards the white spot in the center.

Here are the three degrees of freedom in this basis

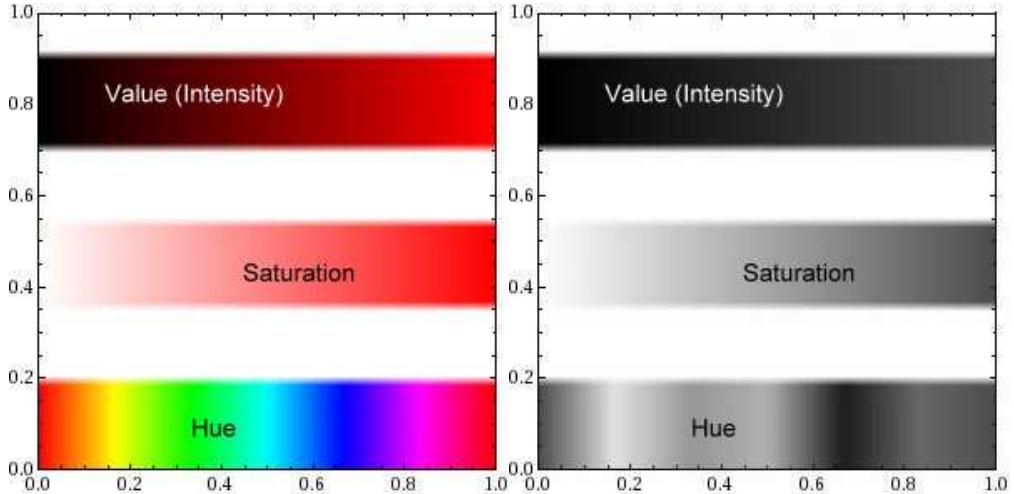


Figure 14. The HSV coordinates – using the `Hue[h,s,v]` command in mathematica.

One advantage of HSV is it lets us easily understand what colors are not in the rainbow. They are all the colors that you can get from a rainbow color (monochromatic, fixed hue), by varying the saturation (moving away from the boundary). If you vary the intensity, you just get a different value, which is normalized away in the CIE color space. Note from black-and-white part of Fig.14 that saturation, not value, characterizes how black-or-white a color is. That is, black-or-whiteness is different from brightness.

Another point you might notice is that the yellow hue seems brighter than the blue hue, despite the fact that they have the same intensity. This is because our eyes are more sensitive to yellow and green light than to blue or red light. Our overall color sensitivity curve looks like this:

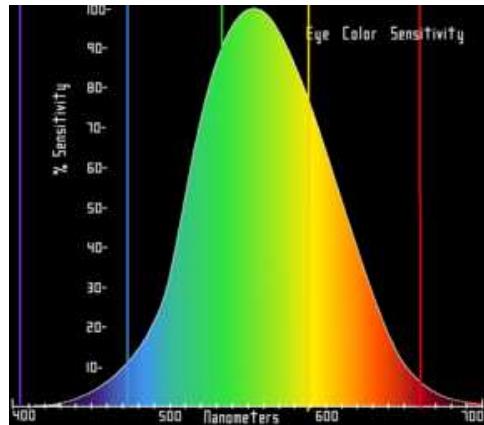


Figure 15. Color sensitivity of human eyes.

6 Additive versus Subtractive colors

The color mixing we have been talking about so far has been mixing of light. This is called **additive** mixing. For example,

$$\text{blue} + \text{red} = \text{purple} \quad (16)$$

This is in contrast to what would happen if we mixed paint: blue paint + red paint = black. This is because blue paint *reflects* blue. It absorbs red and green. red paint absorbs blue and green. Put them together and you absorb everything, giving black. This is called **subtractive mixing**. Here's an easy way to remember the difference

- Additive mixing makes things whiter.
- Subtractive mixing makes things black.

For example, here are some standard color mixing diagrams

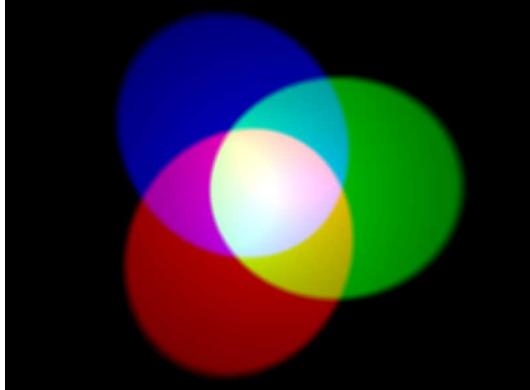


Figure 16. Additive color mixing

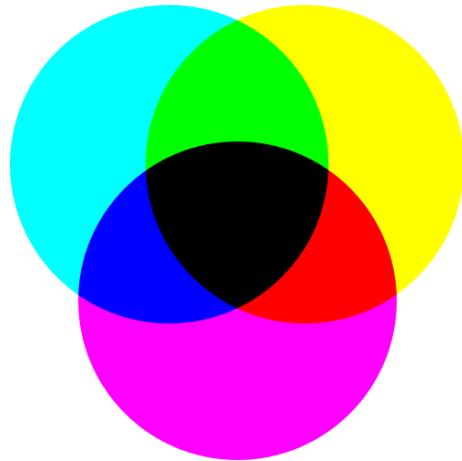


Figure 17. Subtractive color mixing

How can we figure out the subtractive color mixing rules using the CIE functions? Subtractive color mixing is multiplicative, since pigments *remove* a fraction of the incoming light. When you mix two colors, you simply multiply their intensities. For example, consider magenta and cyan. Their RGB values are

$$\text{magenta} = (1, 0, 1), \quad \text{cyan} = (0, 1, 1) \quad (17)$$

If we mix them additively we get

$$\text{magenta} + \text{cyan} = (1, 1, 2) \sim \text{lavender} \quad (18)$$

If we mix them subtractively, we need to write

$$\text{magenta} \times \text{cyan} = (1, 0, 1) \times (0, 1, 1) = (0, 0, 1) = \text{blue} \quad (19)$$

We can see the relationships between the additive and subtractive systems in the color wheels above.

The **complimentary color** \bar{C} of a color C is one which you must add to get white: $C + \bar{C} = \text{white}$. Colors on opposite sides of white (or black) are complimentary. Note that the subtractive primaries are white minus the additive primaries *e.g.* $\text{cyan}(0,1,1) = \text{white}(1,1,1) - \text{red}(1,0,0)$, which is why they are called subtractive. When you combine many colors additively, they give white and when you combine them subtractively they give black. Try also this useful additive and subtractive mixing applet

- Mixing applet http://www.michaelbach.de/ot/col_mix/index.html.

6.1 CMYK basis – printing

Why are subtractive colors important? Printing, for one thing. If you mix red, green and blue you get white light. Printers work by printing the same thing in three colors. But if you put down red blue and green ink, you will just get black, not white. This is a little subtle. You might think that if you painted a bunch of thin green lines, red lines, and blue lines next to each other, all the colors would get reflected, making the combination white. However, only 1 / 3 of the light is reflected, so it is a very dark white. Also known as black.

The translation from RGB colors, which are good on your screen, to colors which print well is not so easy. In fact, it is very complicated. First of all, what is the best basis for ink colors we can choose? For light colors, we saw it was the basis with the biggest gamut triangle, which was some kind of red, green and blue. For printing, we need the complimentary CIE space.

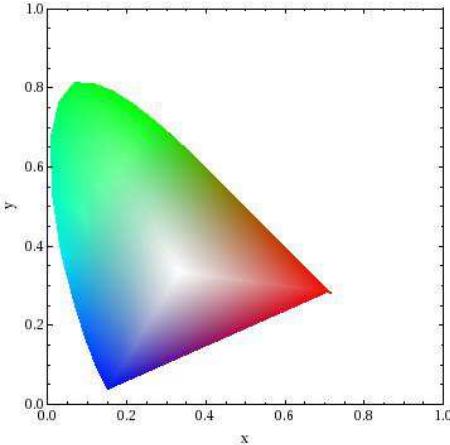


Figure 18. Additive color space

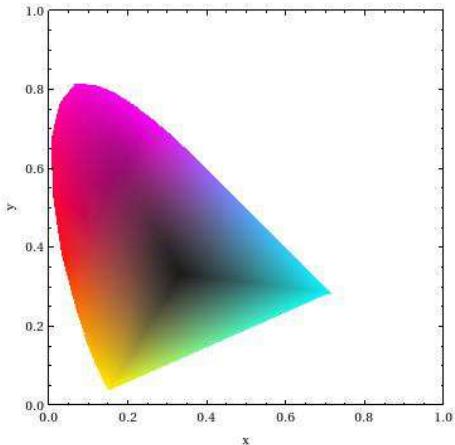


Figure 19. Subtractive color space

From this we can see that cyan, yellow and magenta make good subtractive colors. This is why printers use these three in their ink. The CMYK stands for these three colors and the K

stands for “Key”, which means the key color black. If the 3 dimensional basis is enough to produce all the colors, why also use black? Because black ink is cheaper than color ink!

This is only part of the story. To get various saturations of color, printers have to mix in white. This is done with **half-toning**, a more sophisticated version of the Ben-Day dot technique which inspired pop artist Roy Lichtenstein. In half-toning, colors are mixed, with some black, and various angles so that at large distances they look like intermediate colors

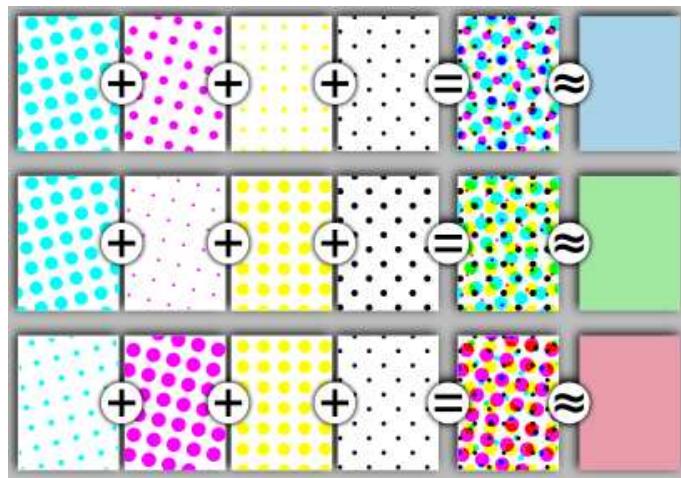
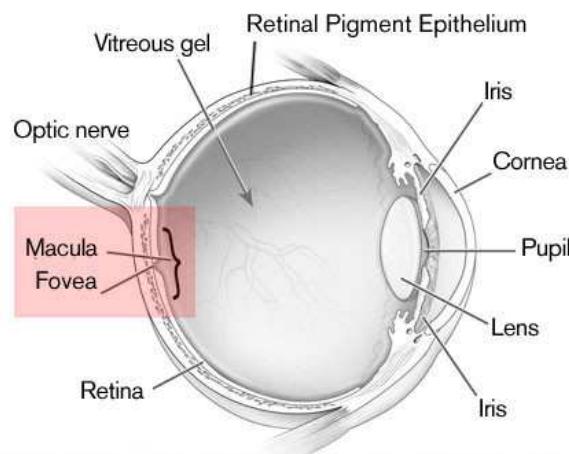


Figure 20. Half-toning

7 The Eye

Next, we'll talk about perception. Color cannot be understood without understanding more of how we humans process it with our cognitive system. To begin, let's look the way the input is processed – through our eye.

The eye looks like this



Credit: National Eye Institute, National Institutes of Health

Figure 21. The Eye. The boxed area contains most of the cones, which are how we sense color.

There are many parts of the eye which affect how we see:

- The pupil: hole where the light comes in.
- The iris: opens and closes around the pupil to let more or less light in.
- The retina: surface of back of the eye. Has the rods and cones
- The lens: focuses the light on the the retina.
- The cornea: Also helps focus light, onto the lens.
- Vitreous humor: fluid in the eye. Index of refraction ~ 1.333
- The macula: part of retina with the best resolution.
- The fovea: Small part of the macula. Consists only of cones
- The optic nerve: pipes information out of the eye. Leads to *blind spot*.

There's a lot of physics in all these parts. For now, we're going to concentrate on the parts relating to color vision. The important facts about the eye relevant to our discussion are mostly about rods and cones.

7.1 Rods and cones

Rods and cones are the photoreceptors in the eye. There are many more rods (~ 120 million) than cones (~ 6 million) in the eye. The rods are scattered all about, except in the fovea, which is the central part of the macula, which is the central part of the retina. The fovea is all cones. There are a few cones outside of the fovea, but none towards the edges of the retina. The angular distribution looks like this:

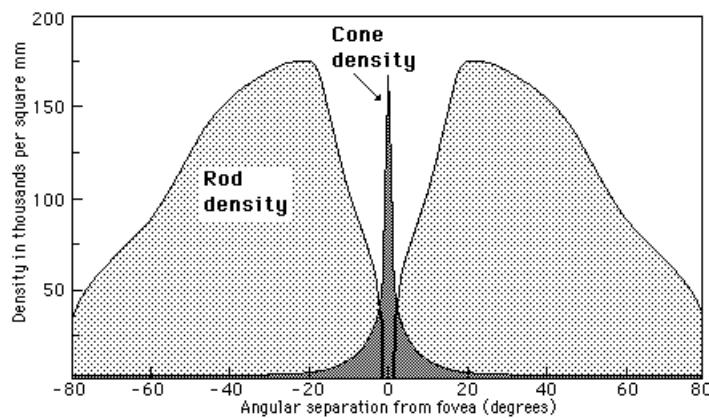


Figure 22. Rod and cone density in the eye

There are 3 kinds of **cones**, called short (blue) medium (green) and long (red), which correspond to different sensitivities in wavelength. The reason we have a 3-dimensional space of color perception (we are *trichromats*) is because we have three cones. The measured sensitivities look

like this

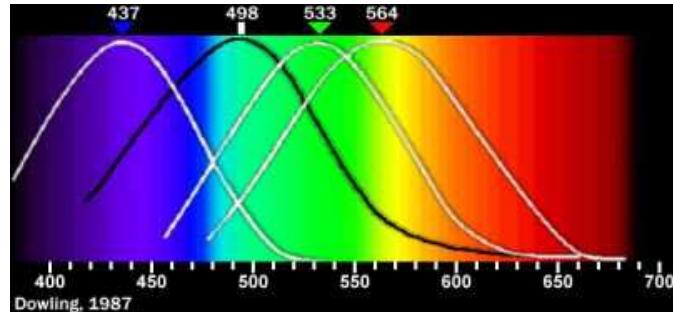


Figure 23. Sensitivity of cones (white curves) and rods (black curve)

Note that the “red” cone, is closer to having its maximal sensitivity in yellow than in red. You might have expected these to look like the X, Y, Z CIE curves, Fig. (4), but they don’t – there is no bump in the red curve at small wavelength. These are basically just the power spectra for forced simple linear systems: resonance with damping, so you wouldn’t expect a second bump. So where does the second bump in the red CIE curve come from? It must be a result of our higher level cognitive processing, which is still not well understood. Since the cones are in the central part of the eye, we have the best color perception in the center of our field of view.

Most animals, such as dogs, only have 2 cones and are *dichromats*. It’s not that dogs are color blind – they can distinguish all the colors of the rainbow, all the hues. They just can’t measure hue, saturation and value, only hue and value. So many colors that we can distinguish, dogs can’t.

Now let’s talk about **rods**. There are 20 times as many rods as cones in the eye. Rods have greater sensitivity to low light conditions and for seeing motion. Since the rods are mostly scattered around the periphery of the retina, we have better motion perception at the edge of our field of view. This makes sense evolutionarily – we want to see predators coming in from the side. When we are looking right at something, we want to know more details, like its color.

Since there are so many more rods than cones, they have different absolute light sensitivities. Overall, our eyes are sensitive to roughly 12 orders of magnitude in intensity, with the rods dominating at low intensity and the cones at high intensity.

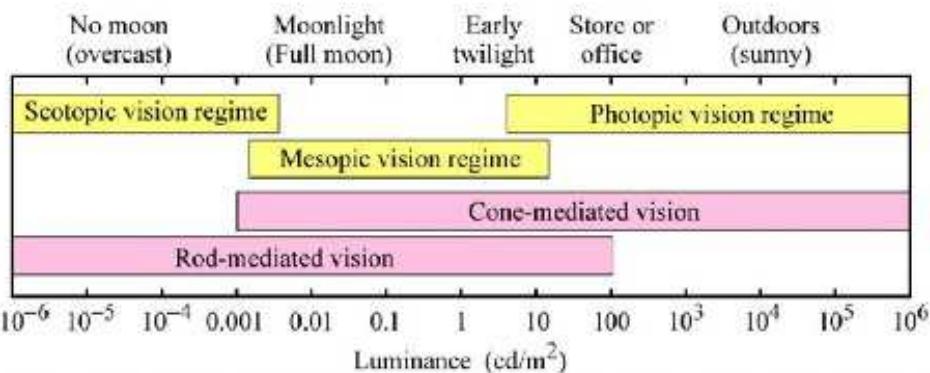


Figure 24. Sensitivity of our eyes. Cd is a candela, which is the SI unit of luminance. Scotopic, mesopic and photopic vision means vision under low-light, medium-light and well-lit conditions.

Since the rods are also in different places in the eye than cones, and since there are so many more of them, our rods do not just act like another cone with blueish sensitivity. Some input from rods is used in the interpretation of color, but we do not have a 4 dimensional space of color perception, only 3 dimensional. Basically, rods are used to interpret intensity, not color.

Rods and cones are types of neurons. The way neurons work is building up potential gradients and then discharging electrical or chemical signals. They produce an all-or-none response. When you combine lots of neurons you can get fairly smooth response curves, but the basic point is that there is some threshold sensitivity for them to work. When there is a lot of light coming in to your eye, you want these thresholds to be low, so that you can tell intensity differences between say $1000 \frac{\text{cd}}{\text{m}^2}$ and $10,000 \frac{\text{cd}}{\text{m}^2}$ (cd is a **candela**). 1 cd is about the amount of light coming off a candle). At low light, you want them to go between $10^{-4} \frac{\text{cd}}{\text{m}^2}$ and $10^{-5} \frac{\text{cd}}{\text{m}^2}$. They can do this, but it takes a little while to recalibrate. It takes longer to calibrate at lower light, simply because there's less light coming in to do the calibration with. Rods typically take 7-10 minutes to calibrate at low light. This is why it takes around 10 minutes to really see stars at night. It takes much less time to calibrate to bright light, ~ 30 seconds.

Let's go back to the rod color sensitivity curve, Fig.(23). Notice that **rods do not see red**. This is why darkrooms are red. It's not that film isn't sensitive to red, it's all about the adjustment time of your eyes. You need to have the lights all the way off when you develop the film. So if your eyes have adapted to the dark, if you put red light on, it will not spoil your rods' adaptation – the rods will still be on low sensitivity settings – low ISO. Your cones will see the red light, and they can adapt quickly. If you shined bright lights, your rods would have to readapt and it would take 10 minutes. This is also why instrument panels on boats are often illuminated with red lights, so the captains can see at night.

Another consequence of the cones not seeing red is that in dim light, red looks black. This is the origin of the Purkinje effect. Purkinje was a 19th century biologist, who observed that the red geraniums in his garden looked striking during the day, but in the early morning, before it was light, they were much darker even their leaves. The contrast is something like this:



Figure 25. Purkinje shift. Red flower during daylight (left) and at dusk (right). The right image is simulated, not a photo.

If we just took a photo of the flower at night, it would look like a dimmer version of the flower, not like the black thing on the right. That's because cameras have color sensors, which are like cones, not rods. The Purkinje effect is due to rods taking over at night. That's part of what makes color photography so tricky! More about this below.

8 Vision and human perception

Now that we know a little about the eye, we can talk about processing color. A surprising amount of color is context based.

8.1 Color adaptation

Look at this figure:

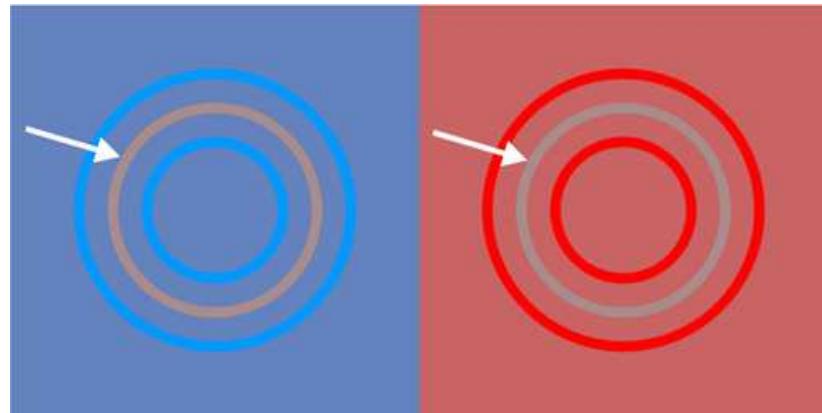


Figure 26. The two rings indicated by the arrow are the same color

The rings pointed at by the arrow look to be different colors, right? In fact, they are the same color. The issue here is that our interpretation of a color is determined in part by its surroundings. Our cones are being saturated by the backgrounds, so that the foregrounds look like they are different hues. It's also possible for our eyes to make color out of pure black and white:

- Color from nothing: http://www.michaelbach.de/ot/col_benham

The sensitivity to context can be seen also for black and white images. Perhaps the best illustration of this is the Adelson's Checkerboard:

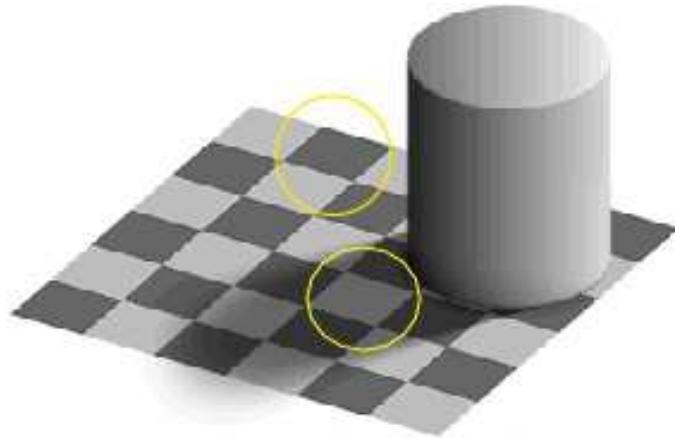


Figure 27. Do the two circled squares look the same? In fact they are exactly the same shade of gray!

Here's another example. This dress looks like a different color, depending on what's in the background:



Figure 28. What color is the dress? It depends on what's around it.

8.2 Blind Spot

Your mind can do all kinds of other things to you as well. The signals of what you are seeing have to exit the eye somehow. They do this through the optic nerve. Unfortunately, the region where the optic nerve exits your eye can't have rods or cones. So it is a blind spot. It's pretty amazing what our minds do to pretend this blind spot doesn't exist.

First, let's prove there is a blind spot. Close your left eye and stare at the cross with your right eye. Move your eye closer to the page. At some point you will see the dot disappear. This is because the dot is right in line with the blind spot



Figure 29. Blind spot

That wasn't the interesting thing (although it is interesting). Now try it again with this figure.

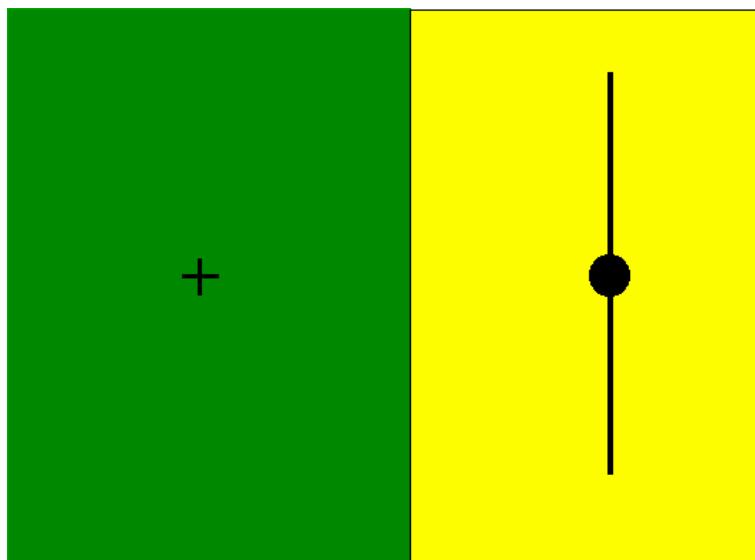


Figure 30. Blind spot #2

You should have seen a number of things: the dot went away. But it was replaced by a line! And a yellow image! Why a line? You think that's the most natural thing to be there.

Our mind tries to fill in information from context in other situations too, not just over our blind spot. For example, look at this image

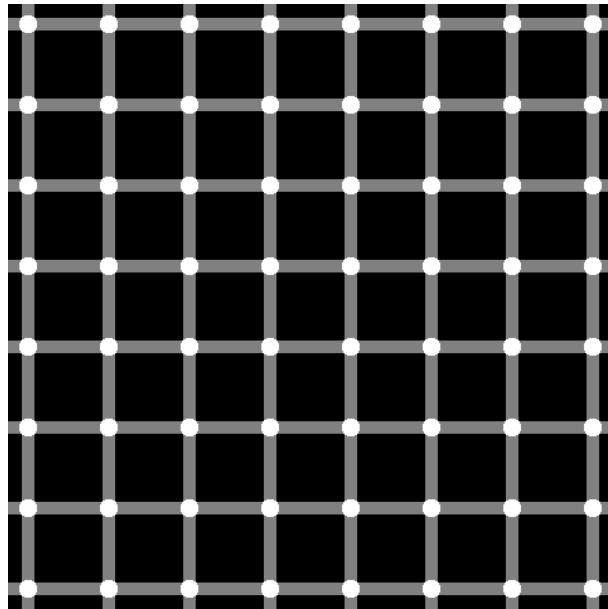


Figure 31. Count the black dots!

There are no black dots, but your mind thinks they should be there, so it puts them in. Perception is complicated!

9 Photography and imaging (optional, but interesting)

We have seen that there is a mathematical way to describe color. All the colors we can see can be written as a linear combination of intensities of a standard set of X , Y and Z “colors”. However, the color we see depends a lot on how we process that signal. This processing involves converting from a linear to a logarithmic scale, by adapting our rods and cones to the dark, and taking clues from context. This makes it very difficult to present color on text or on a computer that look like what we see in real life.

First of all, since any 3 RGB values can produce only a subset of all the colors we can see, it's impossible to get a true color image of anything. That's why photographs of rainbows never look like the real thing – they can't.

RGB values are designed to simulate the color we can see. This does great in bright light, where cones dominate, but not for dim light, where the rods are critical. Recall the Purkinje effect, Figure (25). If you took a photograph of a flower in dim light, it would look like a dimmer version of the flower, not what we see. Of course, if you look at the photo in dim light, it would start to look like the dim flower, but who wants to look at photos in the dark? Instead, we look in bright light, and see the reds perfectly well, even though they are at low intensity. So how can we get the photo to look like what we see, where the reds appear black? That is one of the main challenges in color imaging: from photography to TV to movies to computers. The main point: if you want to make the image simulate our perception, you have to do more than simply reproduce the RGB values.

9.1 Digital Cameras

It is helpful at this point to discuss some elements of how cameras work. Since no one uses film anymore, we will talk about digital cameras. We will also only concentrate on the parts having to do with color. There are a lot of interesting things we could say about the physics of a camera's lens and aperture, but this lecture is long enough already.

Digital cameras work just like you'd think they would. They have a bunch of little cavities where light enters. These are the pixels. The pixels get covered with a filter which only lets in certain light. Something like this:

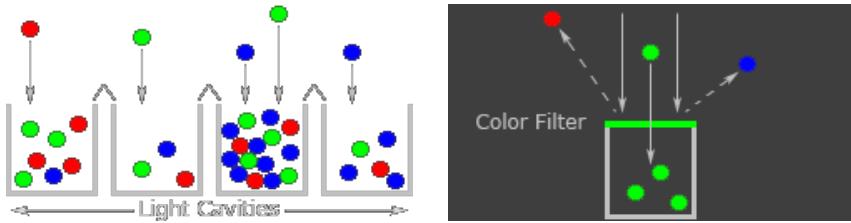


Figure 32. Components of the sensor of a digital camera

So each cavity can only take one color. Since we want a basis of 3 colors, how do we arrange the pixels? What is normally done is to use a **Bayer array**, which contains twice as many green pixels as red or blue. It looks like this:

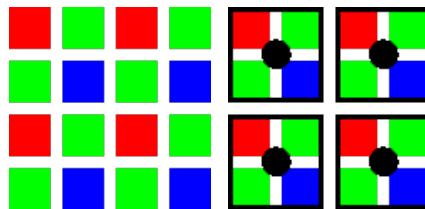


Figure 33. Bayer array. The 4 squares of pixels are turned into RGB values at each black point.

The colors are then combined into RGB values. Digital cameras generally have twice as much luminosity resolution in green than in red or blue, so the red and blue colors appear twice as noisy. This was chosen because we are maximally sensitive to green.

The output of the picture is R, G, and B intensity readings for each of the millions of pixels in your camera. In an image, as in an uncompressed .tiff file, each pixel has 8-bits (a number from 0 to 255) of information in each color channel. This is the standard **bit depth** of the image. Recall that we are sensitive to 12 orders of magnitude in intensity. So the first thing the camera has to do is to pick the range of intensities which it wants to map.

The **ISO** (International Organization for Standardization) setting of a camera tells which intensity range to use. If you are in dim light, you should use a high ISO, like 1600 ISO, which multiplies the intensities measured by roughly a factor of 1600. So going from 100 ISO to 1600 ISO, you can increase the sensitivity dramatically. Even in dim light, at 1600 ISO, there may be so few photons coming in that not all the pixels will get lit up. In any case, there will probably not be a nice smooth distribution of intensities, so the picture will often look grainy. On the other end of the range, if it is very bright out, you want to use something like ISO 100, so you will have good resolution of the bright light coming in. Here is a comparison of different ISOs

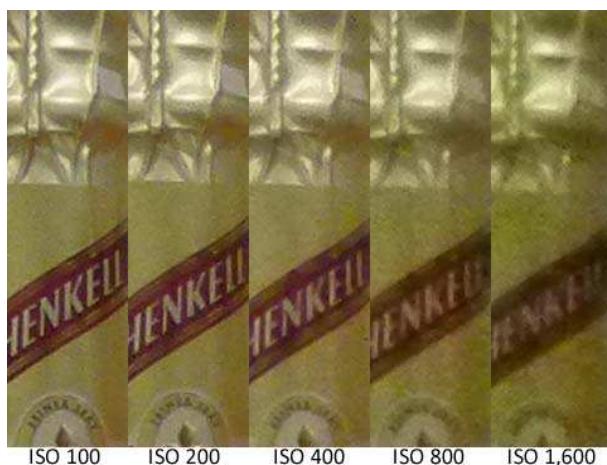


Figure 34. ISO comparison. Note how the 1600 ISO is very grainy.

Even with an ISO setting, you often still want to record more than a factor of 256 in intensity. This can be done, if we map

$$I(\lambda) \rightarrow I(\lambda)^\gamma \quad (20)$$

for some γ . The number γ is called the **gamma** factor. It is recorded into your image file when you save the photo. Because our eyes see light on essentially a logarithmic scale, in the same way we hear sound on a logarithmic scale, this gamma factor is critical to getting an image to look right.

If the γ factor is recorded in the image, and then interpreted correctly by your computer screen, or your printer, it just acts like a compression parameter for the spectrum. However cameras are very poor at figuring out the right γ . Moreover, because our minds interpret logarithmic separations in intensity on an essentially linear scale, we have a built in γ factor. So finding the right γ is not just a matter of recording the image intensity, we also need to know about our perception. A practical result is that changing the γ value can make an image look much more realistic. For example, look at this figure

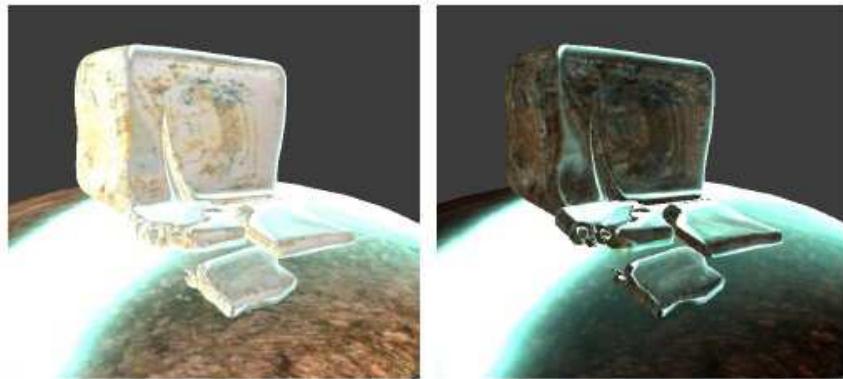


Figure 35. Original image (left) image with gamma adjusted (right)

This image has a wide range of intensities, so it is very sensitive to γ . Adjusting γ appropriately can make the image appear much more realistic.

Monitors and TVs have γ correction factors. The intensity of light is produced by the monitor can range over more than 256 bits, due to this γ factor. You can usually calibrate the γ correction by hand so that the images look proper. It is very difficult to get a monitor to look right both during the day, with lots of ambient light, and during the night, without ambient light.

9.2 Light Bulbs

Now let's talk about illumination. We can light a scene with all kinds of different light bulbs. Here is the same image lit up with three different bulbs:

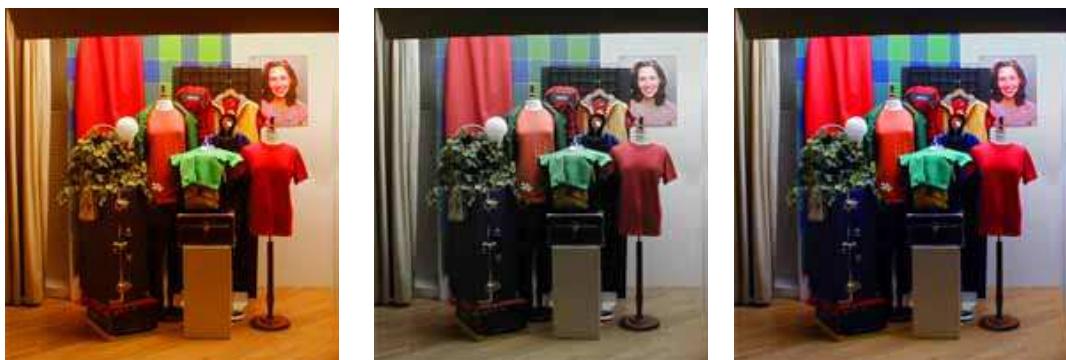


Figure 36. Same picture, with an incandescent bulb (left), GE cool white bulb (middle), and GE Chroma 50 (right).

The first picture clearly looks yellowish. But which of the other two bulbs looks more like “the real thing”. Unfortunately, there’s no exact answer for that. The colors seem brighter on the right. So we might say the colors in the middle correspond to a dimmer scene, where the reds are muted. But it’s not dimmer – it’s just different lighting. In fact, the scene *is* yellowish in incandescent light. We just don’t usually notice in a room with regular light bulbs because we compensate in our minds using facts we already know about what the scene *should* look like.

To see why light bulbs color things, we can compare the spectrum of daylight and the spectrum of an incandescent bulb

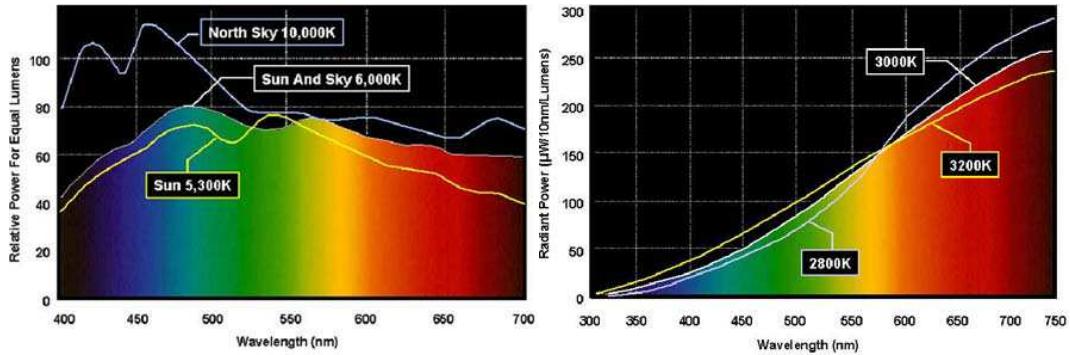


Figure 37. Spectra for daylight (left) and an Incandescent bulb (right).

Incandescent bulbs are the traditional lightbulbs. They usually have a tungsten filament which simply heats up and radiates thermally. That is, incandescent bulbs are black bodies, and therefore radiate power smoothly across the entire visible spectrum. They are generally around 3000K, which is less than the sun ($\sim 5000K$), so they usually peak in the infrared. That’s why they are so energy inefficient – most of the power goes into invisible heat radiation. Thus, the visible spectrum is just the tail of the blackbody distribution (see Fig. (10)), so they tend to tint things reddish-yellow.

The other two bulbs lighting the scene in Fig. 36 are CFLs (Compact Fluorescent Lamps). Their power spectra look like this:

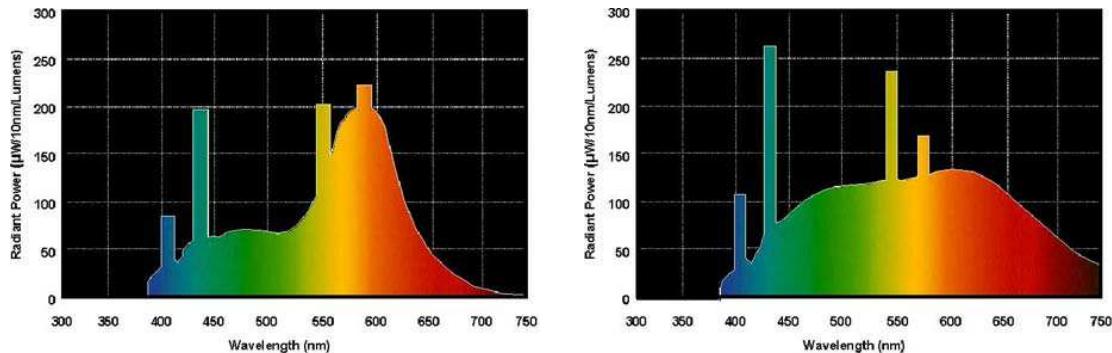


Figure 38. Spectra for a cool white (left) CFL light bulb and the GE Chroma 50 (right) CFL light bulb.

Fluorescent bulbs, including the compact kind, are usually filled with mercury. Electrical discharges excite the mercury, which then de-excites, emitting UV radiation. The UV radiation hits phosphors (complicated chemicals – not phosphorous) on the walls of the bulb which convert the UV to visible light. The result of this conversion is the smooth part of the spectrum. The peaks are the visible light emissions from mercury itself. The difference in spectra is due to different kinds of phosphor on the walls of the bulb.

One other thing about lightbulbs is that they are sometimes described as “full spectrum”. That means close-to-a-blackbody. There’s something called the **Color Rendering Index** (CRI), which is defined as the closest distance a light source is to the Planckian Locus on the CIE diagram (see Figure 12). So if you know the temperature of a lightbulb and the CRI, you can place the spectrum on the CIE the diagram. A full-spectrum bulb has CRI above 90%. An incandescent bulb is full-spectrum since it is a blackbody (CRI=100%).

9.3 White balance

Now that we understand a bit about lighting, we can ask what a picture is supposed to look like. If we compare the differently-lit photos in Fig. 36, they are all good renditions of what is seen. The incandescent photo looks yellow because it is yellow! But somehow we want it to look as if it weren’t yellow. That’s one of the hardest challenges in photography.

If we use incandescent light, things look yellow. So we want to compensate for this so things look “normal”. What is normal? The generally accepted definition is sunlight:

- **We expect things to look like they were illuminated with sunlight**

So a lot of the art of photography is trying to correct for the way color is portrayed. The best tool we have for this is **white balance**.

The white balance setting on your camera tries to **correct back to sunlight**. This is a very difficult problem, and most cameras suck at it. Here’s an example scene with incandescent lighting and various white balance settings on the camera.



Figure 39. awb is auto white balance, then tungsten and manual white balance. The bottom right is an 80a filter put directly on the lens (an old-school solution).

So you see white balance is more than just white, it corrects the whole spectrum. But white is the hardest color to correct for, since it is a careful balance of all the others. One way to correct for it is to use a pure white piece of paper and photograph it with your image. This is still commonly done by professional photographers. An alternative is to load your photo into your favorite editing program and click on a region which you know is white. Then it can try to compensate.

As a final important point, a camera actually takes the picture with more than 8 bits, usually 12. Then it applies your white balance setting and γ -correction to encode the larger range of colors and intensities it sees into a fewer number of bits. This mapping cannot be undone since information is lost in the process. It can be compensated for in post-processing, but the image will be degraded. Some point-and-shoot cameras, and all professional cameras, have a setting called RAW. Raw records all the information that the camera actually sees (the *raw* Bayer array intensities) into a file, with 12 bit depth. Then you can load up your image into photoshop and choose the white balance and γ *afterwards*. If you really want your pictures to look right, this is what you have to do. Keep in mind, if you edit a picture on a computer, the monitor has its own gamma correction, which must be calibrated, then you also have to make sure your printer is calibrated too.