Chapter 6

Color

Tuesday, Sept. 10, 2013 MIT EECS course 6.869, Bill Freeman and Antonio Torralba

Figure 6.1: Mixing of the colors of cleaning fluids: The blue-green of Windex and the yellow of Joy make green.

My first experience with color science happened when I was a child. I placed my yellow ski goggles over a light blue bedspread and saw a color different than either of those two– green! It was magical. (The scene is re-enacted in Fig. 6.1 using cleaning fluids. As we'll see from our analysis shortly, the light blue bedspread would probably better be called "cyan"). Color theory is a wonderful mixture of mathematics and aesthetics.

Why do we need color vision? People with color deficits can be unaware of their deficiencies relative to other peoples' visual systems and can function just fine in the world. Yet color makes vision much easier: it lets us isolate objects in front of backgrounds, infer physical properties of foods and surfaces (check whether fruit is ripe), and tell whether our children are sick by looking at the color of their skin.

We'll first describe the physics of color, then discuss our perception of it–the physiology and psychophysics–, and finally address how we make inferences about the world from color measurements.

6.1 Color physics

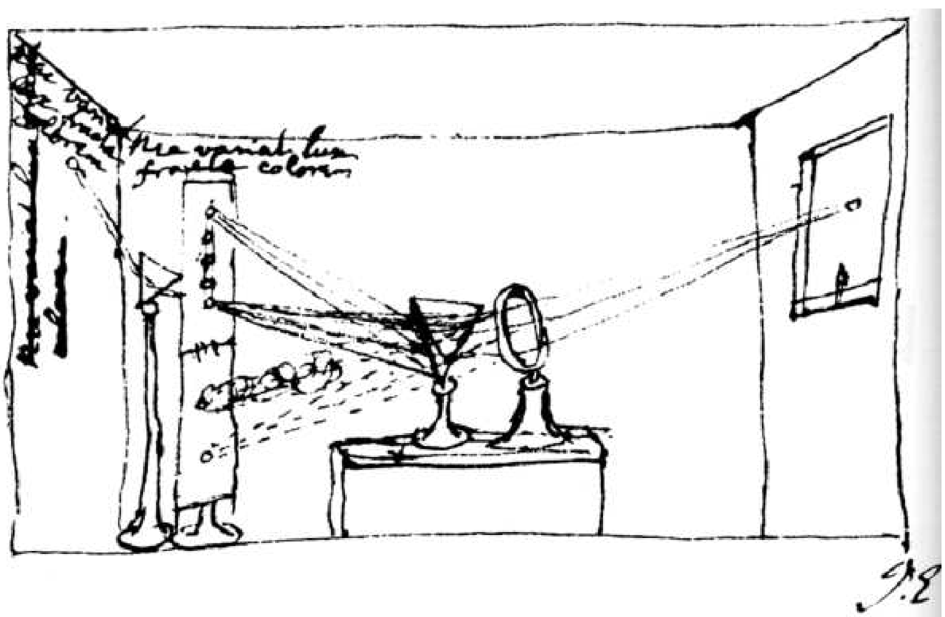
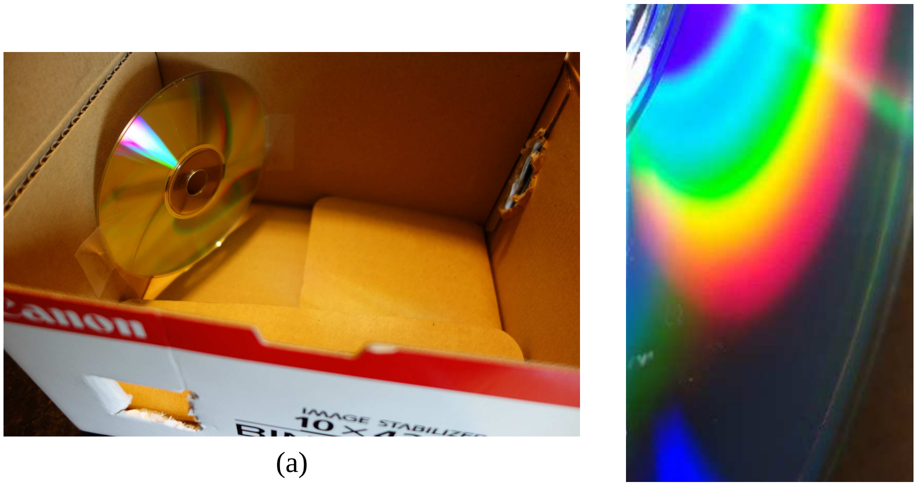


Figure 6.2: Isaac Newton's illustration of experiments with light. White light enters from a hole in the window shade at the right, where it is focused with a lens and then passes through the triangularly shaped prism. The prism bends the light rays a different amount, depending on each color. Those colors are elemental: if a color is passed through a prism again, it doesn't further break into other colors.

Electromagnetic waves surround us, at wavelengths ranging from shorter than the 0.5 nm of x-rays through the longer than 5 meters for radio waves. Our eyes are sensitve to only a narrow band of that electromagnetic spectrum, however, from approximately 400 nm for deep purple to 700 nm for deep red.

What are the properties of light? In experiments summarized by his drawing, shown in Fig 6.2, Isaac Newton revealed several intrinsic properties. Here, a pinhole of sunlight comes in through the window shade, and a lens focuses the light onto a prism. The prism then divides the white light up into many different colors. These colors seem to be elemental: if you take one of the component colors and pass it through a second prism, it doesn't split into further components; it just bends.

Such experiments led to our understanding of light and color. Sunlight has a broad distribution of light of the visible wavelengths. At an air/glass interface, light bends in a wavelength-dependent manner, so a prism disperses the different wavelength components of sunlight into different angles, and we see the different wavelengths of light as different colors, but these do not further subdivide into other colors.



(a) (b)

Figure 6.3: (a) A spectrograph constructed using a compact disk (CD). Light enters through a slit at the right, diffracting from the narrowly spaced lines of the CD. (b) Photograph of diffraction pattern from sunlight, seen thorugh hole at bottom left.

Another simple experimental set-up to reveal the spectrum of light, using very accessible parts, is the CD spectrometer depicted in Fig. 6.3 (a). Light passes through the slit at the right, and strikes a CD (with a track pitch of about 1600 nm). Constructive interference from the light waves striking the CD tracks occurs at a different angle for each wavelength of the light, yielding a separation of wavelengths by diffracting angle. The diffracted light can be viewed or photographed through the hole at the bottom left. (For construction details and more examples, see this very nice web page:

`http://www.cs.cmu.edu/˜zhuxj/astro/html/spectrometer.html`

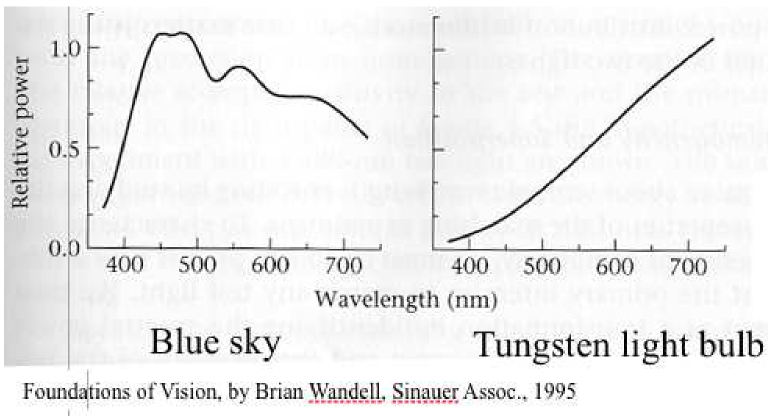
6.1.1 Radiometry and simplified reflection model

We can characterize the light by its power at each of the visible wavelengths, Fig. 6.4 shows a number of different light sources, and the spectra of the light they emit.

Light interacting with matter

We see through the interaction of light with matter. When light strikes a surface, it re-radiates with a distribution of directions and intensities at each wavelength that depends on the properties of the reflecting surface. The distribution of the outgoing light also depends on properties of the incoming light, one of the many reasons why vision is difficult.

We can summarize the changes to the light upon surface reflection with a “bi-directional reflectance distribution function”, or BRDF. The function is called “bi-directional” because it depends on both the direction of the light incident on the surface and on the direction of the reflected light being characterized. The BRDF of a surface characterizes, for each wavelength, the fractional change in the spectral power of light reflecting off the surface as a function of the angle of incidence of the light to the surface, and the viewing (reflection) angle. Shiney surfaces reflect most of the light into a single angle. Diffuse surfaces scatter the light broadly over a hemisphere of directions.



(a) (b)

(c)

(d)

Figure 6.4: (a) and (b): Plots of the power spectra of blue sky and a tungsten light bulb. Photographs show (c) a flourescent light and (d) its spectrum as viewed with the spectrograph of Fig. (6.3) (a).

In this chapter, to simplify our study of surface appearance, we will ignore many of the rich details of the BRDF. We will assume that the observed power spectrum of the reflected light does not depend on either the angles of incidence or reflection from the surface. That is commonly the case for diffuse reflections Under such conditions, the power spectrum of the reflected light, $R(\lambda)$, is simply a wavelength-by-wavelength product of the illumination power spectrum, $I(\lambda)$ and the surface reflectance spectrum, $S(\lambda)$:

$$R(\lambda) = I(\lambda)S(\lambda) \tag{6.1}$$

Wavelength-by-wavelength multiplication is also a good model for spectral changes to light caused by viewing light through an attenuating filter. The incident power spectrum is multiplied at each wavelength by the transmittance spectrum of the filter.

Even the simple model of Eq. (6.1) describes a rich visual world and allows us to make useful inferences.

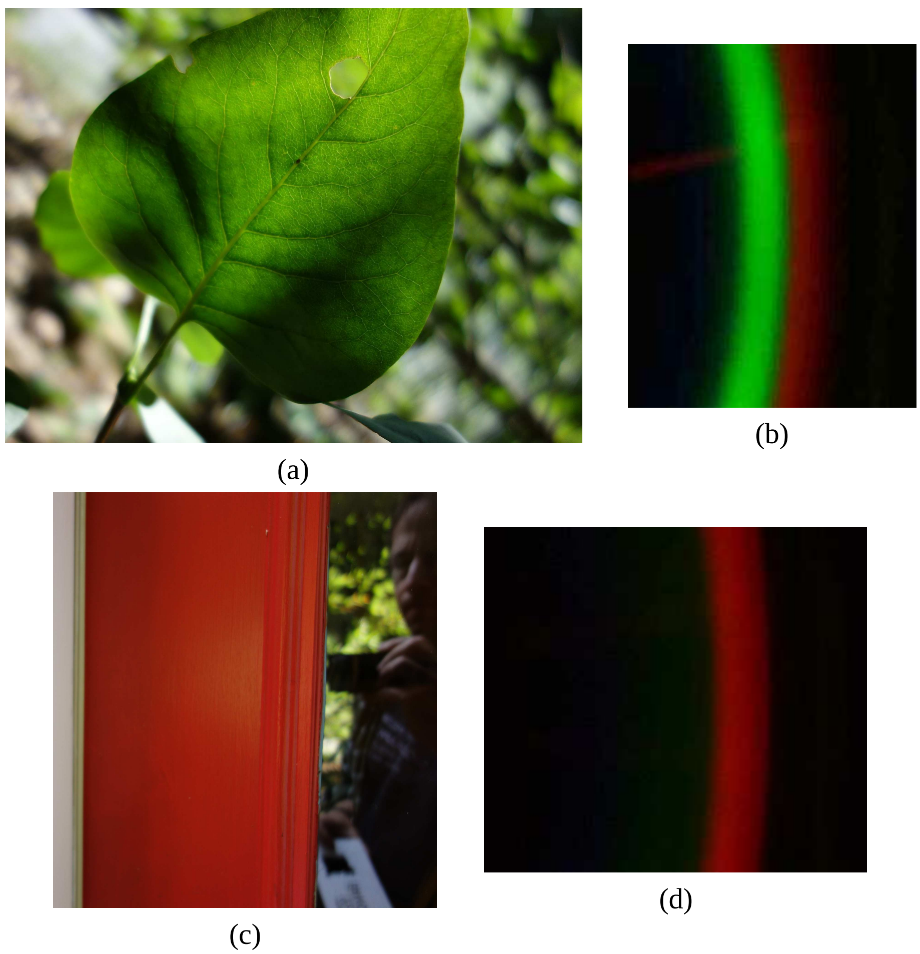
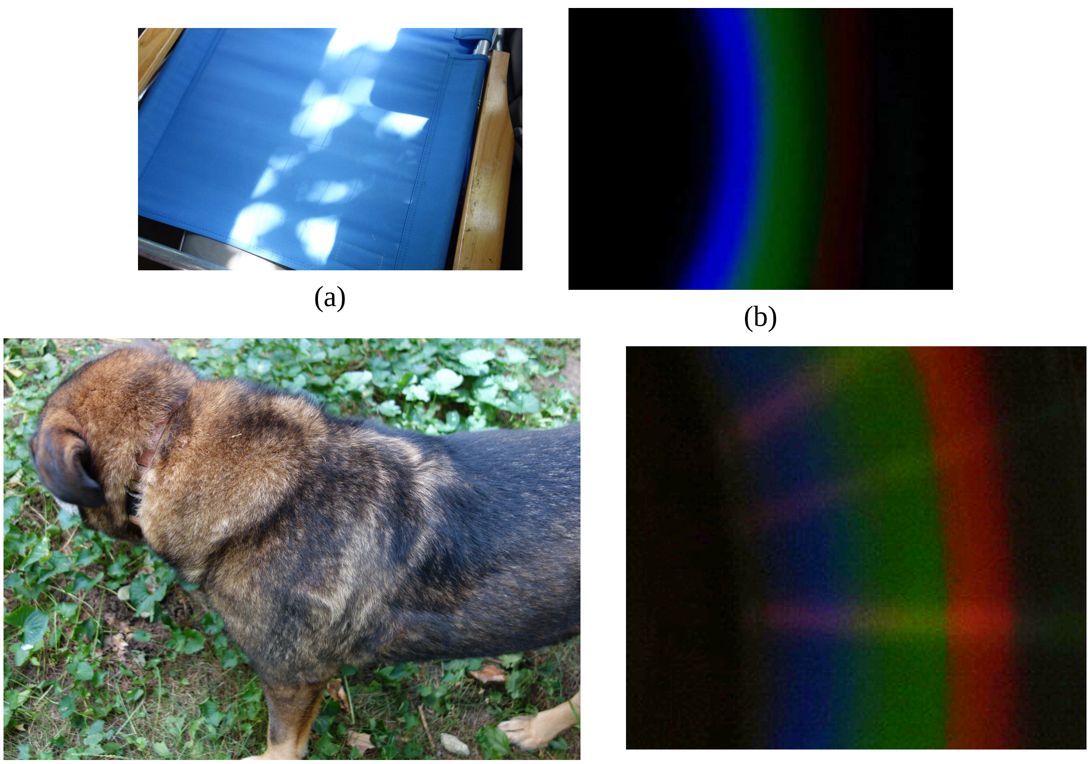


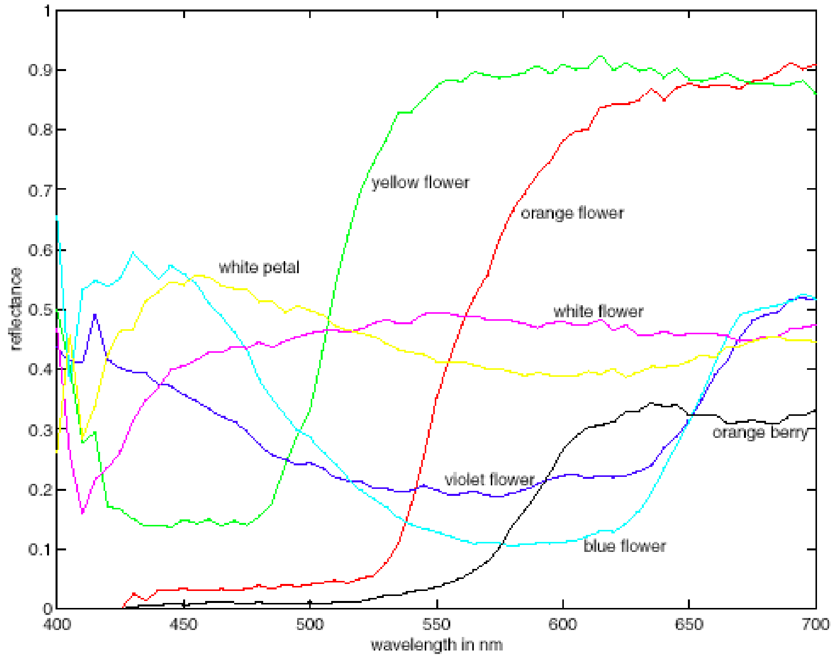
Figure 6.5: Some real-world objects and the reflected light spectra (photographed using Fig. (6.3) (a)) from outdoor viewing. (a) Leaf and (b) its reflected spectrum. (c) A red door and (d) its reflected spectrum.



(a) (b)

(c) (d)

Figure 6.6: More real-world objects and the reflected light spectra. (a) Blue-green chair and (b) its reflected light. (c) Toby the dog and (d) his reflected spectrum.



Forsyth, 2002

Figure 6.7: Observed spectra of light reflecting off the surface. Source: Forsyth and Ponce, Computer Vision, Prentice Hall.

6.1.2 Color appearance of various spectra

Figure 6.4 (a) shows some example illumination spectra. The spectrum of blue sky is on the left, and the spectrum of a tungsten light bulb (which will look orangish) is on the right. Some reflectance spectra are in Figure 6.7. A white flower reflects spectral power almost equally over all visible wavelengths. A yellow flower reflects in the green and red.

6.1.3 Cartoon Color Spectra

It's helpful to develop the skill of being able to look at a light power spectrum and to know roughly what color that spectrum would correspond to. Here is a rough description of what wavelengths correspond to what perceived colors, with a reference spectrum showing roughly what each individual wavelength, viewed by itself, looks like. (An engineer at the photographic company, Polaroid, showed this to me. I think of it as a cartoon color model, a hard-edged approximation to a much softer reality). The visible spectrum lies roughly in the range between 400 and 700 nm. We can divide into three one-hundred nm bands, which, from short to long wavelengths, corresponds to blue, green, and red (again, speaking in broad strokes). These are often called the additive primary colors, which we'll write more about.

White light is a mixture of all spectral colors. There are three other possible combinations of the three one-hundred nm bands of wavelengths, and each can be associated with a color name: Cyan is a mixture of blue and green, or roughly spectral power between 400 and 600 nm. In printing applications, this is sometimes called "minus red", since it is the full spectrum, with the red band subtracted. Blue and red, or light in the 400-500nm band, and in the 600-700nm band, is called magenta, or minus green. Red and green, with spectral power from 500-700 nm, make yellow, or minus blue.

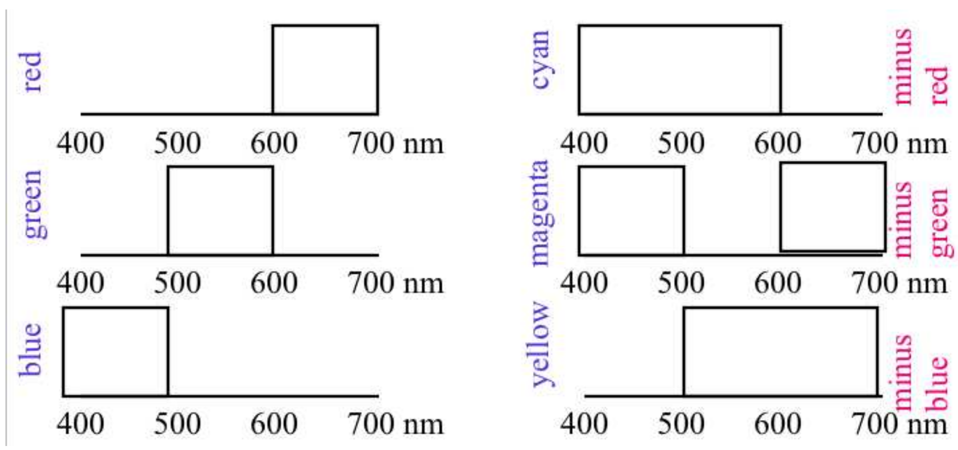


Figure 6.8: Cartoon model for the reflectance spectra of observed colors

6.1.4 Why color is useful

Here is why color is useful: it tells us something about surfaces in the world. For example, assume we have a white light source shining on a yellow egg. The light reflected from the egg, returning to the eye, will be yellow, letting us know, from a distance something about the properties of the egg's surface (that it's yellow).

Let's pose a problem that we'll address later in the chapter: given that we only observe the product of the illumination and reflection spectra, how do we know whether we are observing a white egg, viewed under yellow illumination, or a yellow egg, viewed under white illumination? See Fig. 6.23.

Before we address color appearance, we continue with two more issues with the physical properties of light spectra: color mixing, and low-dimensional models.

6.1.5 Color mixing

A light of one spectrum and color, shining on a surface of another color spectrum produces a third color whose spectrum is the wavelength-by-wavelength product of the two colors, as described by Eq. (6.1). This can be thoughtof as a form of *color mixing*, where the illumination color and the surface reflectance color mix to form the color of the reflected light.

There are two different ways that spectra combine when we mix colors together. While the precise way two spectra combine may depend on the details of the corresponding physical process, these two methods are a good model for many physical processes.

The first way is called additive color mixing. This is the way spectra combine when you project two lights simultaneously, so they are summed in our eye. CRT color televisions, DLP projectors, and colors viewed very closely in space or time all exhibit additive color mixing. The spectrum of the mixed color is a weighted sum of the spectra of the individual components. In the additive color mixing model, in our cartoon color model, red and green combine to give yellow.

The second way colors combine is called subtractive color mixing, but might make more sense to be called multipliciative color mixing. This is the mixing of light reflecting off a surface. Under this mixing model, the spectrum of the combined color is proportional to the product of the mixed components. This color mixing occurs when light reflects off a surface, or passes through a sequence of attenuating spectral filters, such as with photographic film, paint, optical filters, and crayons. An example of color mixing under the subtractive model, cyan and yellow combine to give green, since the cyan filter attenuates the red components of white light, and yellow filter would remove the remaining blue components, leaving only the green spectral region of the original white light.

Figure 6.9 shows the cartoon spectra of these color mixing examples.

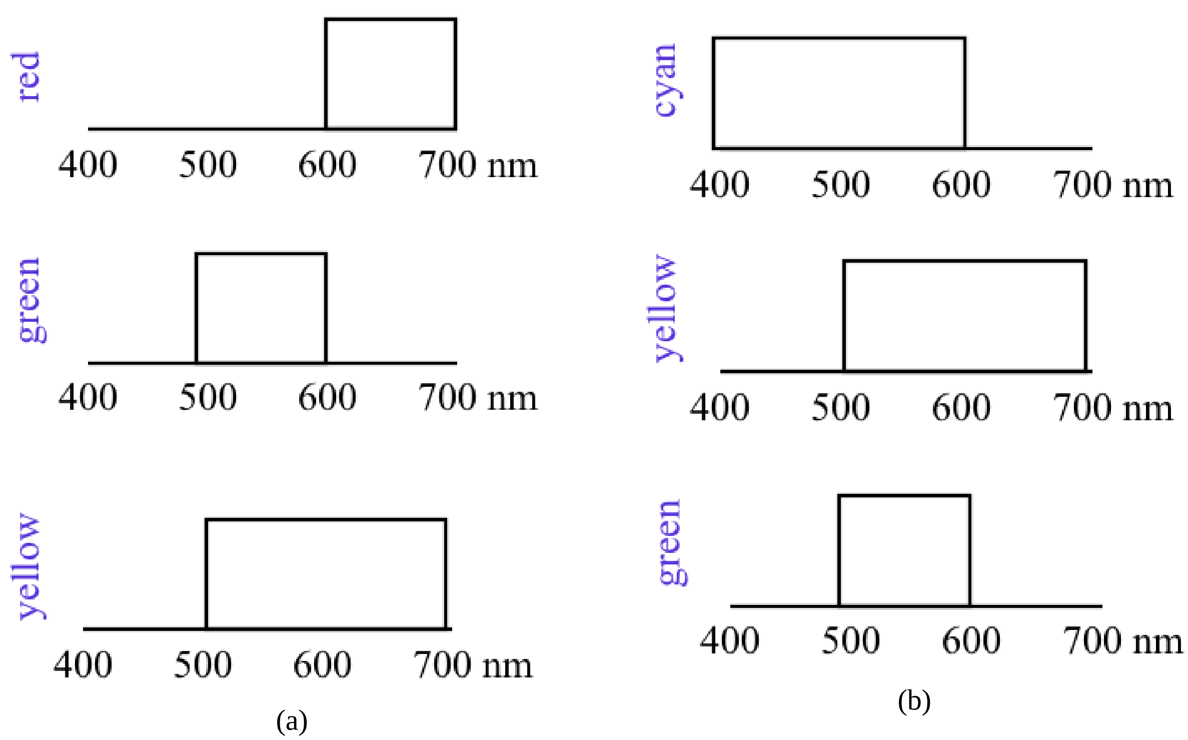


Figure 6.9: Examples of color mixing, in the world of cartoon color spectra. (a) In additive mixing, red and green combine to give yellow. (b) Under subtractive mixing, cyan and yellow mix to give green.

6.1.6 Low-dimensional models for spectra

Before we turn to color perception, let's introduce a mathematical model for light spectra that makes them much easier to work with. In general, when modeling the world, we want to keep everything as simple as possible, and that usually means working with as few degrees of freedom as possible. Color spectra seem like relatively high-dimensional objects, since we can pick any combination of numbers, from 400 to 700 nm, as we'd like. Even sampling only every 10 nm of wavelength, that gives us 31 numbers for each spectrum.

It turns out that for many real-world spectra, those 31 numbers are not independent and in practise spectra have far fewer degrees of freedom. It is common to use low-dimensional linear models to approximate real-world reflectance and illumination spectra. Any given spectrum, say $S(\lambda)$, is approximated as some linear combination of "basis spectra", $u(\lambda)$. For example, a 3-dimensional linear model of $S(\lambda)$ would be

$$\begin{pmatrix} \vdots \\ S(\lambda) \\ \vdots \end{pmatrix} \approx \begin{pmatrix} \vdots & \vdots & \vdots \\ u_1(\lambda) & u_2(\lambda) & u_3(\lambda) \\ \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \tag{6.2}$$

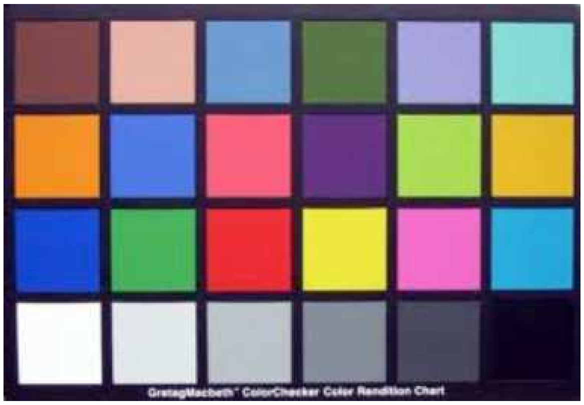
The basis spectra can be found from a collection of training spectra. If we write the training spectra as columns of a matrix, D, then performing a singular value decomposition on D yields

$$D = U * \Lambda * V' \tag{6.3}$$

where U is a set of orthonormal spectral basis vectors, Λ is a diagonal matrix of singular values, and V' is a set of coefficients. The first n columns of U are the n basis spectra that can best approximate the spectra in the training set, in a least squares sense.

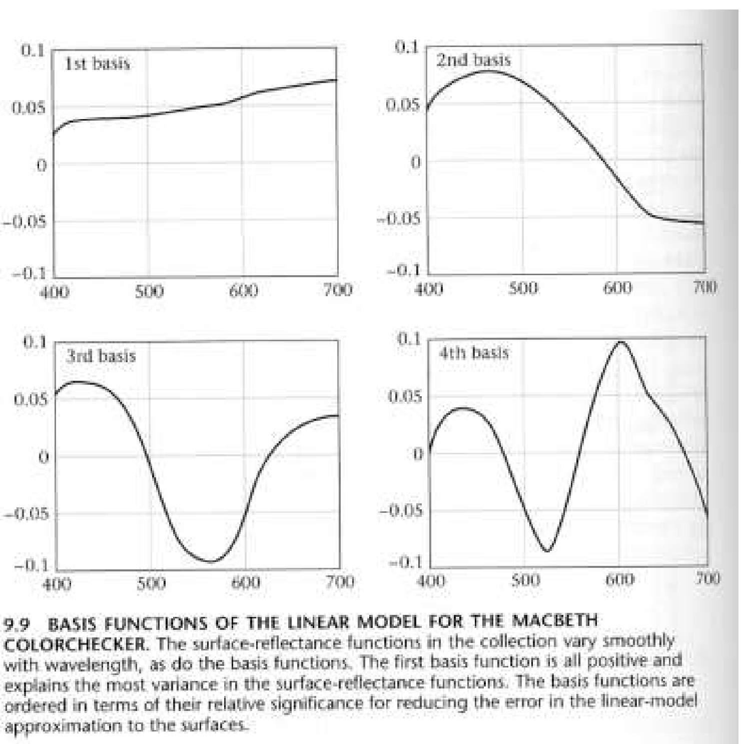
Figure 6.10 shows a demonstration, with a particular collection of surface reflectance spectra, $u_i(\lambda)$ that this works quite well. The "Macbeth Color Checker", a tool of color scientists and engineers, is a standard set of 24 color tiles, always made the same way Figure 6.10 (a). (So iconic that this woman, Figure 6.10 (b), a dedicated color scientist, I presume, has tatooed a Macbeth color checker on her arm! Alas, I'm sure the tatoo colors can only be an approximation to the real Macbeth colors).

The reflectance spectra of each Macbeth color chip has been measured. The first four basis spectra, calculated using the measured reflectance spectra and Eq. (6.3), are shown in Figure 6.10 (c). The rows of Figure 6.10 (d) show the Macbeth color checker spectra, as optimally approximated by 1, 2, and 3 basis functions, respectively. The spectra are pretty well approximated by a 3-dimensional linear model, as you can see from the plots.

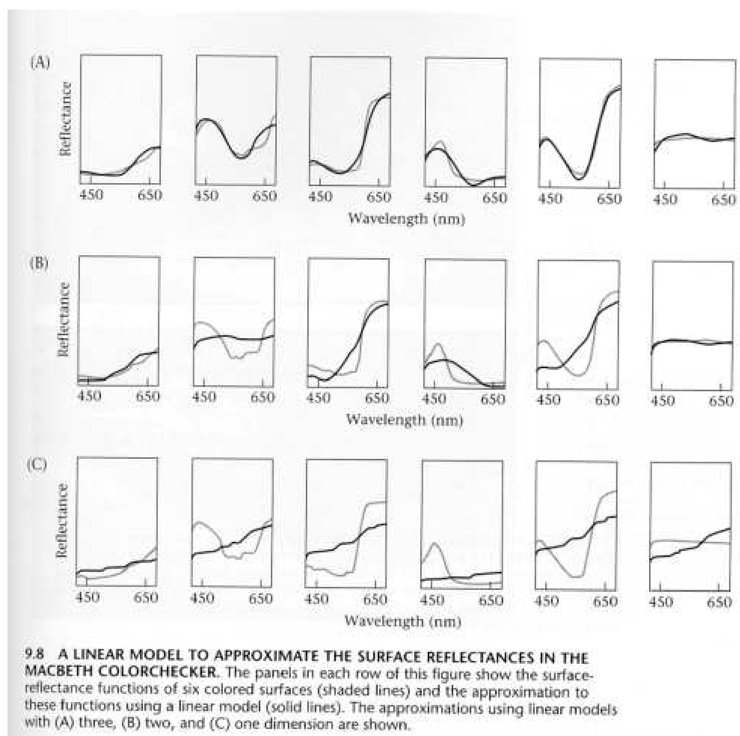


(a)

(b)



(c)



(d)

Figure 6.10: (a) The Macbeth color checker, of such iconic status that a woman (b) has tatooed it on her arm. The bottom two figures are from Foundations of Vision, by Brian Wandell, Sinauer Assoc., 1995 (c) Basis functions from which the Macbeth color checker reflectance spectra can be approximated, in (d), using 1, 2, and 3 basis functions (each row of (d), bottom to top).

6.2 Color measurement: assigning categories and numbers to color spectra

Now, we turn to the measurement of color appearance. Do all people peceive colors in the same way? Not all human languages divide up the space of all colors in the same groupings, which, in principle, could imply differences in perception among those different human groups. For example, some languages, which tend to be spoken by people at high northern latitudes, divide up the colors that English speakers call “blue” into a finer set of categories,

`http://www.nature.com/news/2007/070430/full/news070430-2.html.`

There may be cultural or physiological reasons why people living in such locations would form finer categories of particular color shades.

Figure 6.11: English speakers lump all these shades into “blue”, while Russian speakers put them into two different verbal categories.

Despite these differences across languages, it turns out that most humans match colors very consistently and one can reliably assign numbers that predict color matches.

If you can assign coordinates to a color percept, there are a wealth of applications. You can build a machine to display colors that match some desired output. You can ensure that the colors of manufactured items are consistent. Companies can trademark colors, so we need to be able to specify what is being trademarked. We have color standards for foods, for example. Figure 6.12 shows a chart showing french fry color standards, one among many standards for food colors.

To see how to quantify colors, we first need to understand the machinery of the eye.

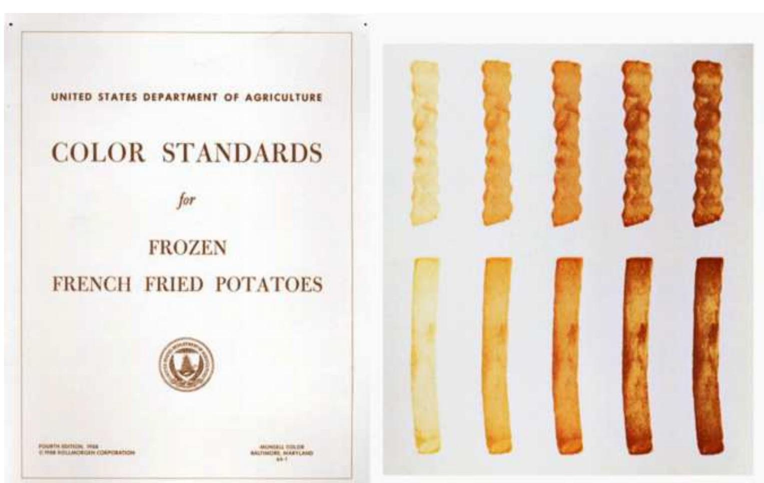


Figure 6.12: French fries color standard

6.2.1 The machinery of the eye

Given that spectra are generally low-dimensional, an organism doesn't need a large number of spectral measurements to measure natural spectra. For our color vision, our eyes have three different classes of photoreceptors, which determines the fact that there are three primary colors, three color layers in photographic film, three colors of dots on a display screen, and why two color coordinates are needed to specify any color, independent of its overall intensity (three color dimensions, minus one for the overall color normalization).

What is the machinery of human vision? Here is a drawing of the rod and cones of the eye, and some of the nerve cells connecting them. The tall ones are the rods and the short ones are the cones, both at the top layer. By the way, where does the light come in, in this drawing? Differently than how you or I might design things, the light comes in at the bottom, passes through the nerve fibers and blood vessels, then reaches the photosensitive detectors at the top of the image. Evolution may have determined that there are benefits to having the photodetectors on the inside, where they can more easily receive nourishment from blood vessels.

The retina consists of 3 classes of color receptors. Figure 6.13 (b) shows the variability of the numbers and spatial layout of color receptors (the red, green, and blue of the figure is *much* more saturated than the spectral sensitivities of the cone receptor classes are). The 3 cone classes are denoted L, M, and S, for whether they are sensitive to the long, middle, or short wavelengths of the visible spectrum. The spectral sensitivity curves are shown in Fig. 6.13 (c).

In some sense, Fig. 6.13 (c) tells the whole story of spectrum-based color perception. Three detectors, with the spectral sensitivities $R_i(\lambda)$ shown here, signal their response to an incoming light spectrum, $I(\lambda)$. The response, γ_i of a cone of color class i is

$$\gamma_i = \int_\lambda R_i(\lambda) I(\lambda) d\lambda \tag{6.4}$$

This can be thought of as projection of the incoming light spectrum onto the three basis vectors $R_i(\lambda)$, projecting the high-dimensional input spectrum onto a 3-dimensional subspace.

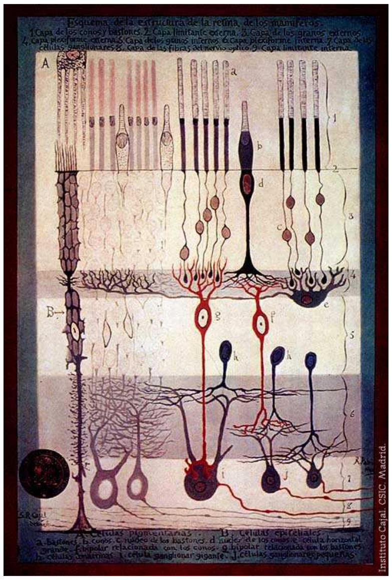
The fact that there are three different detectors means that we'll 3 numbers to describe a perceived color in the world (or 2 numbers, if we normalize for intensity). The shapes of those curves tell us which real-world spectra will look the same to us (because they'll give the same trio of photoreceptor responses) and thus will tell us how make one color look like another one.

To help understand how color is measured, and the experiments that taught us what we know about color vision, let's examine the psychophysical experiments that were done to learn what we know about color perception.

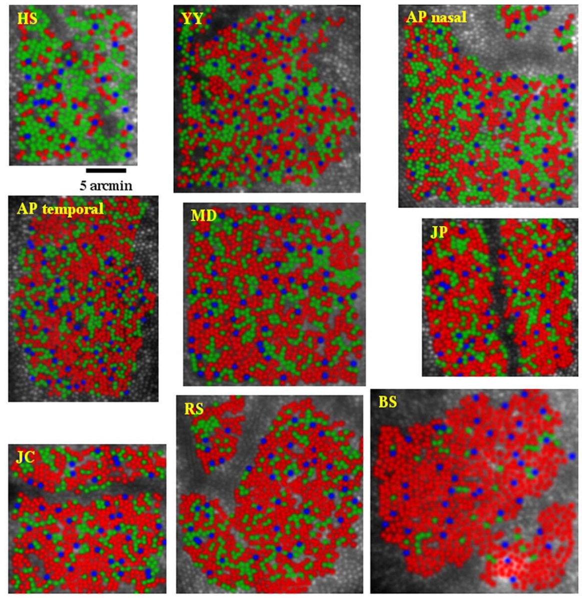
6.2.2 Color matching

Color perception measurement is mostly about color matching. We try to match a color with an additive combination of a set of reference colors, typically called primary colors. Through experimentation, it has been found that we can match the appearance of any color through a linear combination of three primary colors, stemming from the fact that we have three classes of photoreceptors in our eyes.

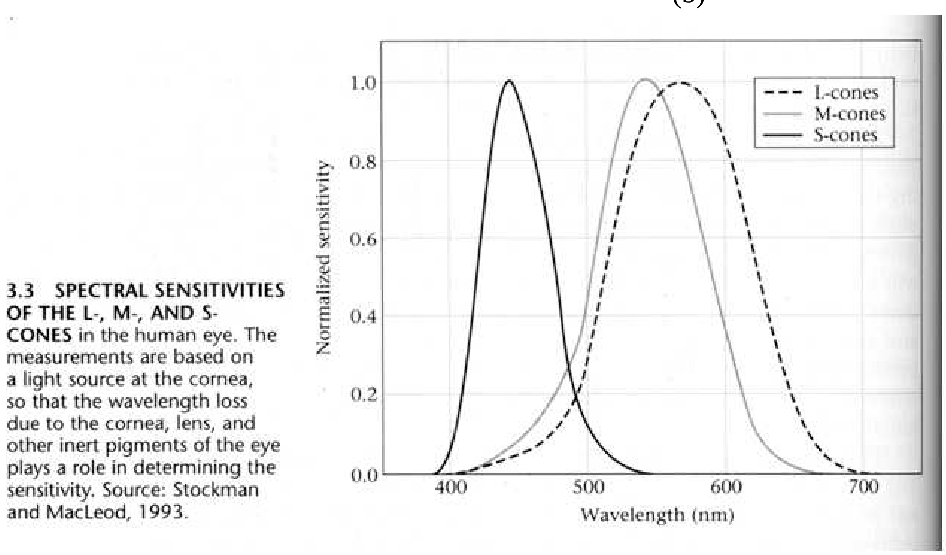
In this section, we're assuming that the color appearance is entirely determined by the spectrum of the light arriving at the eye. To ensure that this is true in the experiments, care must be taken to view the color comparisons under repeatable, controlled surrounding colors, because such details can influence the color percept. We shine a controllable combination of the primary lights on one half of a bipartite white screen, and the test light on the other half, see Fig. 6.14 (a). A grey surround field is placed around



(a)



(b)



(c)

Figure 6.13: (a) Drawing of the eye's photoreceptors by the Spanish physiologist, Ramon y Cajal. (b) View of 9 different human foveas, with the cone receptor types (L, M, or S) marked (in R, G, and B, respectively). [citation below] (c) Spectral sensitivities of the L, M, and S cones.

The Journal of Neuroscience, 19 October 2005, 25(42): 9669-9679; doi: 10.1523/JNEUROSCI.2414-05.2005, Organization of the Human Trichromatic Cone Mosaic, Heidi Hofer, Joseph Carroll, Jay Neitz, Maureen Neitz, and David R. Williams the viewing aperture, giving a view to the subject that looks something like that of the right hand side of Fig. 6.14 (a).

Finally, we arrive at how we can assign a number to any color: Pick a set of 3 lights, called primaries, and see what combination of these primaries is required to match any given color. This gives a (reproducible) representation for the color at the left: if you take these amounts of each of the selected primaries, you'll match the input color.

What if our three selected primiaries don't let us reach the test color? Fig. 6.14 (b) shows an example of that. It turns out we can always match any input test color if we "add negative light", which means to add positive light to the other side of the test comparison.

Human color matching has elegant properties that help us describe colors using linear algebra. Most every desirable linear property is satisfied with such color matching experiments. Here's one of them: if color A_1 matches color B_1, and color A_2 matches color B_2, then the sum of colors A_1 and A_2 will match the sum of colors B_1 and B_2.

That tells us that if we represent a color by the amount of the 3 primaries needed to make a match, or any number proportional to that, then we'll be able to use a nice vector space representation for color, where the observed linear combination laws will be obeyed.

That's the psychophysics. We also have in the back of our heads the mechanistic view for how colors generate signals in our brain: the light power spectrum gets projected onto the 3 photoreceptor classes spectral sensitivity curves, generating three numbers, the L, M, and S cone responses, which are the signal for that color. If we can adjust the primary color amounts, a_1, a_2, and a_3, so that their sum generates the same set of photoreceptor signals when projected onto the photoreceptor spectral response curves, we have matched the color.

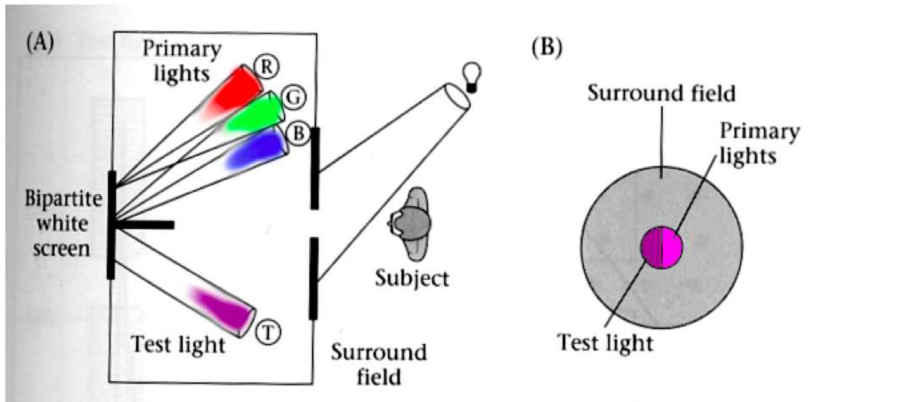
6.2.3 Linear algebraic interpretation of color perception

The psychophysics result leads to a linear algebraic interpretation of color. Let the space of all possible spectral signals be N-dimensional. In this figure, we depict that as a 3-dimensional space. The generation of cone response for a given spectral sensitivity curve can be thought of as projecting a N-dimensional signal onto a basis function and recording the resulting projection length. If we record the response of the three different spectral sensitivity curves, we are measuring the projection of our N-d vector onto each of three linearly independent basis vectors. Thus, a triplet of cone responses maps onto some coordinate in a 3-dimensional subspace (depicted as a 2-d plane here) of the original N-dimensional space.

Viewed that way, then the task of color measurement is simply the task of finding the projection of any of the possible N-d spectra into the special 3-d subspace defined by the cone spectral response curves. Any basis for that 3-d subspace will serve that task. Equivalently, we seek to predict the cone responses to any spectral signal, and projection of the spectral signal onto any 3 independent linear combinations of the cone response curves will let us do that.

So we can define a color system by simply specifying its 3-d subspace basis vectors. And we can translate between any two such color representations by simply applying a general 3x3 matrix transformation to change basis vectors. Note, the basis vectors do not need to be orthogonal, and most color system basis vectors are not.

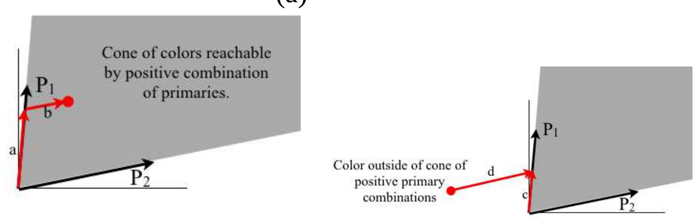
Long before scientists had measured the L, M, and S spectral sensitivity curves of the human eye, others had measured equivalent bases through psychophysical experiments. It is interesting to observe how such curves could be measured psychophysically.



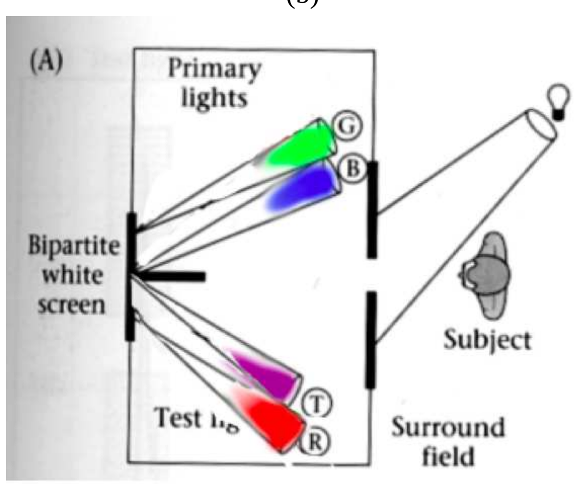
4.10 THE COLOR-MATCHING EXPERIMENT. The observer views a bipartite field and adjusts the intensities of the three primary lights to match the appearance of the test light. (A) A top view of the experimental apparatus. (B) The appearance of the stimuli to the observer. After Judd and Wyszecki, 1975.

Foundations of Vision, by Brian Wandell, Sinauer Assoc., 1995

(a)



(b)



Foundations of vision, by Brain Wandell, Sinauer Assoc. 1995

(b)

Figure 6.14: (a) Color matching experiment (From Brian Wandell, Foundations of Vision, Sinauer, 1995). (b) Vector model for color matching experiment. Primary lights can synthesize a cone of possible colors. If a desired color to match is outside that cone, we can add a primary color to the test light until it is inside the feasible cone. Adding a light to the test light side, (c), is the same as subtracting it from the primaries, but doesn't require negative light intensity.

Color matching functions

Here's what we can do to find such basis vectors, called "color matching functions", for any given set of primary lights. We exploit the linearity of color matching and find the primary light values contributing to a color match, one wavelength at a time. So for every pure spectral color as a test light, we measure the combination of these three primaries required to color match light of that wavelength. For some wavelengths and choices of primaries, the matching will involve negative light values, and remember that just means those primary lights must be added to the test light to achieve a color match.

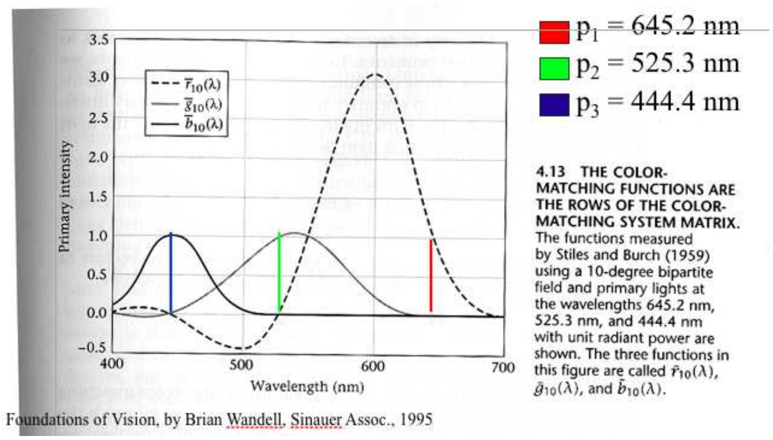


Figure 6.15: Psychophysically measured color matching functions.

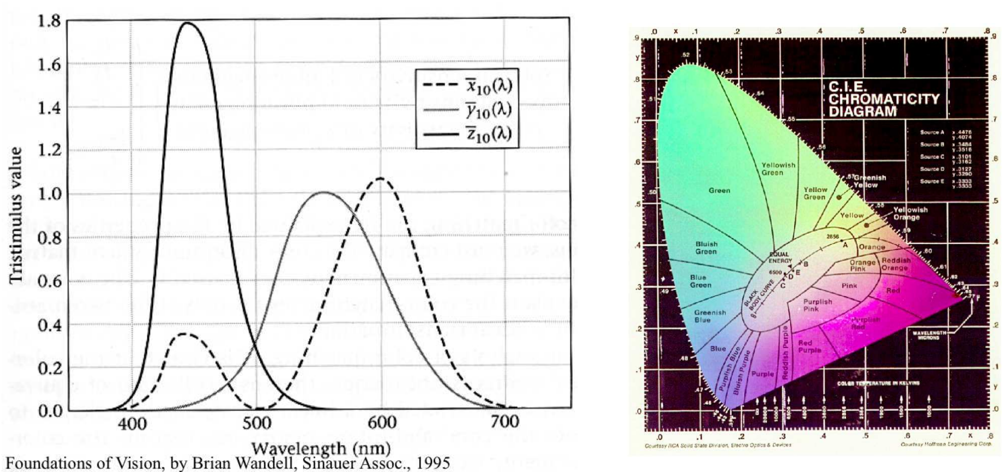
Figure 6.15 is an example of such a measured color matching function, for a particular choice of primaries, monochromatic laser lights of wavelengths 645.2, 525.3, and 444.4 nm. We can see these matches are behaving as we would expect: when the spectral test light wavelength reaches that of one of the primary lights, then the color matching function is 1 for that primary light, and 0 for the two others.

Because of the linearity properties of color matching, it's easy to derive how to control the primary lights in order to match *any* input spectral distribution, $t(\lambda)$. Let the three measured color matching functions be $c_i(\lambda)$, for $i = 1, 2, 3$. Let the matrix $\mathbf{C}$ be the color matching functions arranged in rows,

$$\mathbf{C} = \begin{pmatrix} c_1(\lambda_1) & \dots & c_1(\lambda_N) \\ c_2(\lambda_1) & \dots & c_2(\lambda_N) \\ c_3(\lambda_1) & \dots & c_3(\lambda_N) \end{pmatrix} \tag{6.5}$$

Then, by linearity, the primary controls to yield a color match for any input spectrum $\vec{t} = \begin{pmatrix} t(\lambda_1) \\ \vdots \\ t(\lambda_N) \end{pmatrix}$ will be $\sum_j C_{ij} t_j = C\vec{t}$.

So there is an infinite space of color matching basis functions to pick, so it's natural to ask whether any one choice of bases is better than another. One natural choice might be the cone spectral responses themselves, but those were only measured relatively recently, and many other systems were tried, and standardized on, earlier.



(a)

(b)

Figure 6.16: (a) CIE color matching functions. (b) The space of all colors (intensity normalized), as described in the CIE coordinate system.

6.2.4 CIE color space

One standard you should know about, because it's so common, is the CIE XYZ color space. Again, a color space is simply a table of 3 color matching functions, which must be a linear combination of all the other color matching functions, because they all span the same 3-d subspace of all possible spectra. The CIE color matching functions were designed to be all-positive at every wavelength. They're shown in Fig. 6.16 (a). What might be the benefit of having an all-positive set of color matching functions? I believe they were selected so that it would be simple to build a machine that used color filters of those spectral responses to directly measure the color coordinates of a signal.

A bug with the CIE color matching functions is that there is no all positive set of color primary lights associated with those color matching functions. But if the goal is to simply specify a color from an input spectrum, then any basis can work, regardless of whether there is a physically realizable set of primaries associated with the color matching functions.

To find the CIE color coordinates, one projects the input spectrum onto the 3 color matching functions, to find coordinates, called tristimulus values, labeled X, Y, and Z. Often, these values are normalized to remove overall intensity variations, and one then calculates $x = \frac{X}{X+Y+Z}$ and $y = \frac{Y}{X+Y+Z}$. Fig. 6.16 (b) shows the visible colors (intensity normalized) plotted in the CIE coordinates x and y.

6.2.5 Color metamerism

One final topic for the model where power spectral density determines color is metamerism, when two different spectra necessarily look the same to our eye. There is a huge space of metamers: any two vectors describing light power spectra which give the same projection onto a set of color matching functions will look the same to our eyes.

There's a sense that our eyes are missing much of the possible visual action. There's a high-dimensional space of colors out there, and we're only viewing projections onto a 3-d subspace of that.

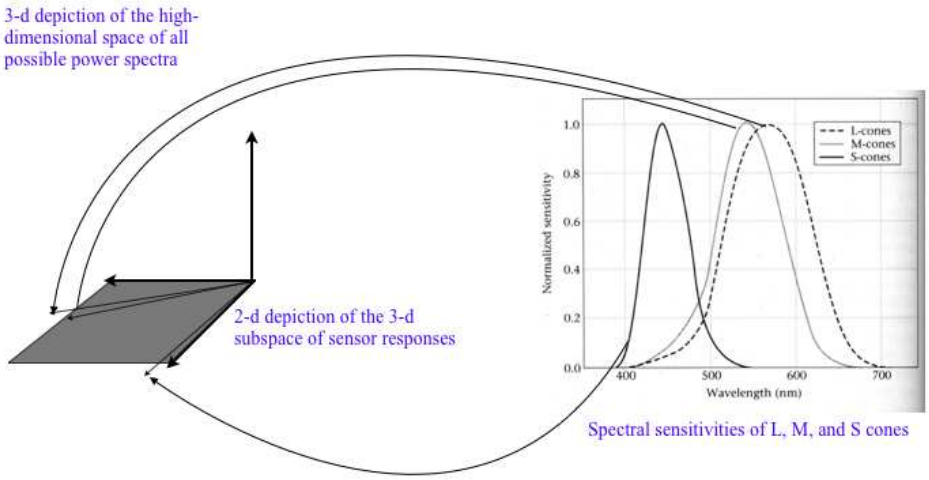
But in practise, the projections we observe do a pretty good job of capturing much of the interesting action in images. Given how much information is not captured by our eyes, hyperspectral images (recorded at many different wavelengths of analysis) add some, but not a lot, to the pictures formed by our eyes.

Let us summarize our discussion of color so far. Under certain viewing conditions, the perceived color depends just on the spectral composition of light arriving at the eye (we move to more general viewing conditions next). Under such conditions, there is a simple way to describe the perceived color: project its power spectrum onto a set of 3 vectors called color matching functions. These projections are the color coordinates. We standardize on particular sets of color coordinates. One such set is the CIE XYZ system.

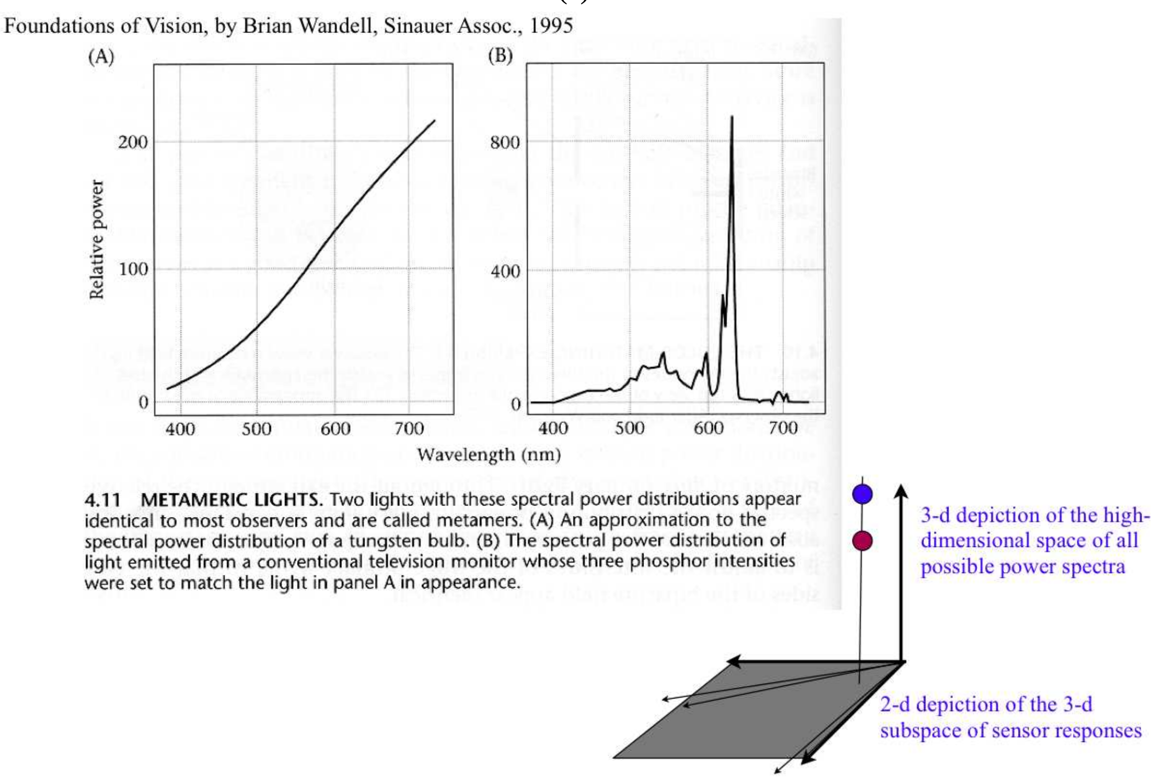
How do you translate from one set of color coordinates to another, say, from the color coordinates in a unprimed system to those in a primed system? Place the spectra of a set of primary lights into the columns of a matrix $\mathbf{P}$. If we take the color coordinates, $\vec{x}$, as a 3x1 column vector and multiply them by the matrix $\mathbf{P}$, we get a spectrum which is metameric with the input spectrum whose color coordinates were $\vec{x}$. So to convert $\vec{x}$ to its representation in a primed coordinate system, we just have to multiply this spectrum by the color matching functions for the primed color system:

$$\vec{x}' = \mathbf{C}'\mathbf{P}\vec{x} \tag{6.6}$$

The color translation matrix $\mathbf{C}'\mathbf{P}$ is a 3x3 matrix.



(a)



(b)

Figure 6.17: (a) The cone spectral sensitivity curves can be thought of as three basis vectors onto which the observed light spectrum is projected. This takes the input spectrum from a high-dimensional space (depicted as 3-d here) into a three-dimensional subspace (depicted as a 2-d plane here). Any two spectra with the same projection into the 3-d subspace (with the same cone responses) will look the same. (b) Two spectra with the same projection onto the cone response curves, and a depiction of the two spectra as distinct points in the high-dimensional space of all possible power spectra, projecting onto the same point in the subspace of possible cone responses.

6.3 Other color coordinate systems

To measure color, we just need to describe point locations within the subspace of human cone responses, but we are free to use different projection bases that span that same 3d subspace. Once we have made the projection into one coordinate system spanning that 3-d subspace, we are free to apply a 3x3 coordinate transformation matrix to employ a different coordinate system. There are many such coordinate systems that have been used to describe colors.

6.3.1 RGB

There are many different standards for color bases called RGB (sRGB, Adobe RGB, etc). Here are the transformation matrices between the CIE coordinates X, Y, Z described above and sRGB:

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 3.24 & -1.54 & -0.50 \\ -0.97 & 1.88 & 0.04 \\ 0.06 & -0.20 & 1.06 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{6.7}$$

The inverse coordinate transformation is:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.41 & 0.36 & 0.18 \\ 0.21 & 0.72 & 0.07 \\ 0.02 & 0.12 & 0.95 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \tag{6.8}$$

6.3.2 YIQ

It is often useful to have one color component correspond to grayscale, or luminance, image variations. A color basis that does this is the YIQ system, used in the old NTSC television standard. Here, Y is a luminance component (not the CIE Y component), and I and Q represent chromatic variations. The translation from an RGB system is given here:

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \tag{6.9}$$

6.3.3 Uniform color spaces

All the color systems described have a common drawback: equal perceptual differences between colors do not correspond to equal numerical distances in the color representations. Nonlinear transformations are required to achieve that, typically a cube root. See

`http://en.wikipedia.org/wiki/Lab_color_space`

for the formulas.

6.4 Spatial Resolution and Color

One reason to transform between color coordinate systems is because of the human visual response to different colors. Figure 6.18 shows the contrast sensitivity to sinusoids of different spatial frequencies for a luminance (Y) grating, and two different color components (labeled R/Y and B/Y in the plot).

We are much more sensitive to variations in luminance, and this property is often exploited in image compression, processing, and display algorithms. Figures 6.19 through 6.22 show the effect on the full color image of blurring different color channels, within different color representations. Blurred chromatic components have very little effect on the full-color image, while a blurred luminance component is quite noticable.

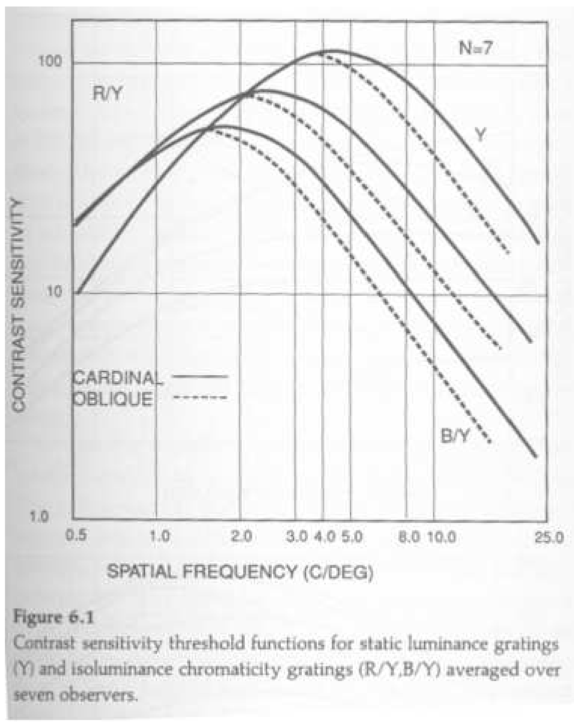
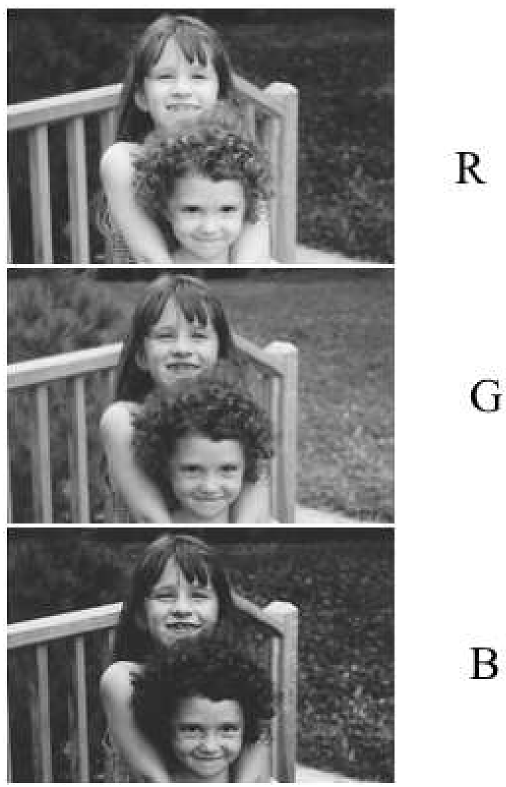
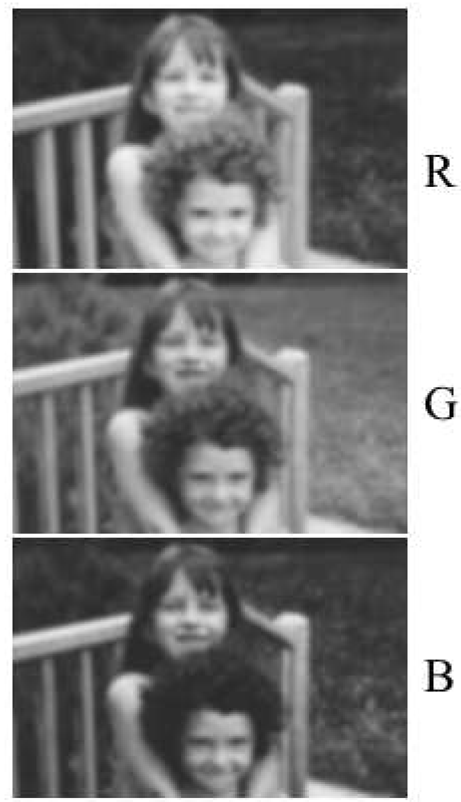


Figure 6.18: Human spatial frequency sensitivity in R, G, B and L, a, b color representations

(a) Original



(b) R, G, B components



(c) blurred R, G, B components

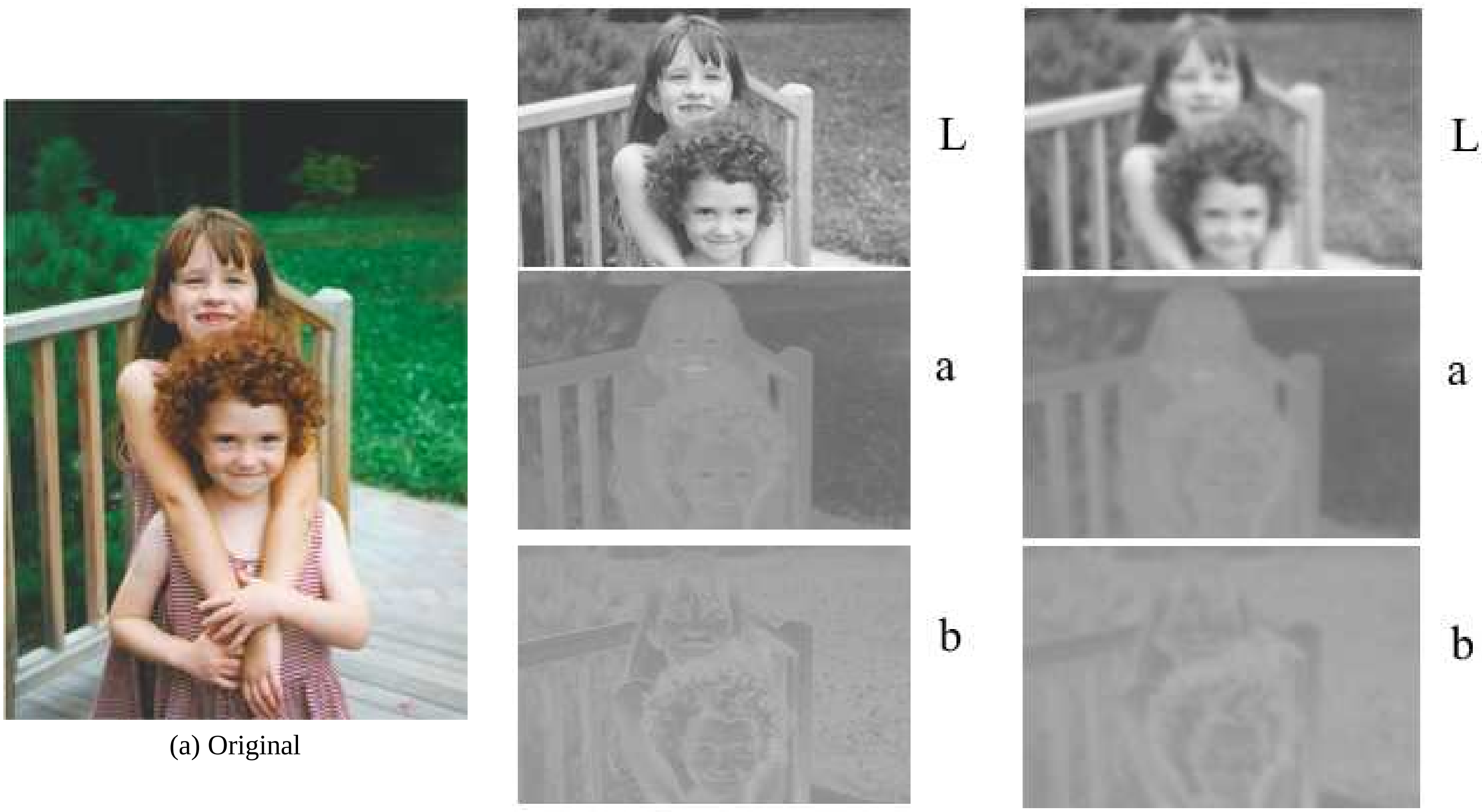
Figure 6.19: (a) Original image. (b) RGB components (c) RGB components, each blurred.

(a) R component blurred

(b) G component blurred

(c) B component blurred

Figure 6.20: (a) R component blurred, G and B components sharp. (b) G blurred, R and B sharp. Green is the dominant component of the luminance signal, and blurring G has the most effect on the color image. (c) B blurred, R and G sharp.



(a) Original

(b) L, a, b components

(c) blurred L, a, b components

Figure 6.21: (a) Original image. (b) Lab components (c) Lab components, each blurred.

(a) L component blurred (b) a component blurred (c) b component blurred

Figure 6.22: (a) L component blurred, a and b components sharp (big effect!). (b) a component blurred, L and b sharp. (c) b blurred, L and a sharp.

6.5 Color Constancy

Color perception depends strongly on the power spectrum of the light arriving at the eye, but it does not depend only on that. Now we address the assumption that a given spectral power distribution always leads to the same color percept.

In the demonstration of Fig. 6.24, identical spectral distributions arriving at your eye lead to different color percepts. What's going on? The eyes receive the product of the illumination and surface reflectance spectra, but the visual system may want to let us "see" the color of the surface color, independent of the spectrum of the illumination. So our visual system needs to "discount the illuminant" and present a percept of the underlying colors of the surfaces being viewed, rather than simply summarizing the product spectrum arriving at the eye. The visual system uses the context of the other colors is used to perform that calculation.

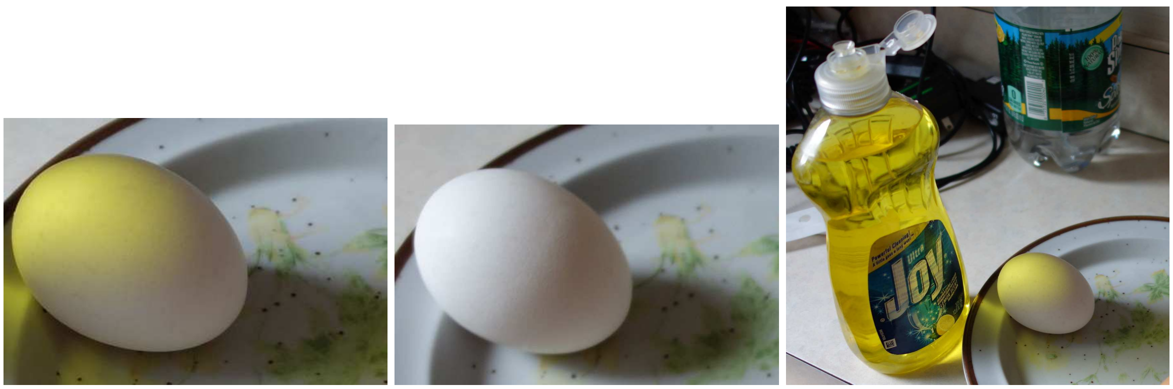


Figure 6.23: How do we distinguish an egg that is yellow because a yellow illuminant is falling on it, from a yellow egg?

The ability to perceive or estimate the surface colors of the objects being viewed, and to not be fooled by the illumination color, is called "color constancy"–you perceive a constant color, regardless of the illumination. People have some degree of color constancy, although not perfect color constancy.

For the case where there is just one illumination color in the image, if we know either the illumination spectrum or any of the surface color reflectance spectra, we can estimate the other from the data. So, from a computational point of view, you can also think of the color constancy task as that of estimating the illuminant spectrum from an image.

The rendering equation

Let's examine the computation required to achieve color constancy. Here's the rendering equation, showing, in our model, how the L, M, and S cone responses for the jth patch are generated:

$$\begin{pmatrix} L_j \\ M_j \\ S_j \end{pmatrix} = \mathbf{E}^T(\mathbf{A}\vec{x}_j^s .* \mathbf{B}\vec{x}^i) \tag{6.10}$$

In the above, L_j, M_j, and S_j are the cone responses of the j patch of color. In this matrix equation, we divide the visible spectrum into N bins. The three rows of the Nx3 matrix, $\mathbf{E}^T$, are the spectral sensitivity curves of the three cone classes. The columns of the matrix $\mathbf{A}$ are the surface reflectance spectra basis functions and $\vec{x}_j^s$ contains the surface (^{s}) reflectance basis function coefficients for the jth color patch.

(a) (b) (c) (d) (e)

Figure 6.24: Color constancy demonstration (made by Prof. David Brainard, U. Penn) (a) a set of colors. (b) The “nothing up my sleeves” picture: the tiny blue square at the left, and the large blue square at right are made from the same filter material. If we cover some of the colors with the small blue square, the colors change their appearance. (c) the white square (3rd row, 3rd column) goes to blue, and (d) orange (1st column, 5th row) goes to green-brown. (e) But if we cover all the colors with the large blue filter, the colors maintain their original appearance, for the most part. White stays white, orange stays orange. This is despite the fact that the same spectral signal is reaching your eye for those two patches as when the small blue squared covered each of them.

".*" represents term-by-term multiplication. Similarly, The columns of the matrix $\mathbf{B}$ are the illumination reflectance spectra basis functions and the vector $\vec{x}_j^i$ contains the illumination (^{i}) spectral basis function coefficients.

Figure 6.25 shows a graphical diagram showing the vector and matrix sizes in the above equation. We have some unknown illuminant, described by, say, a 3-dimensional vector of coefficients for the illumination spectrum basis functions. For this jth color patch, we have a set of surface reflectance spectrum basis coefficients, let's say also 3-dimensional. The term-by-term product of the resulting spectra (the quantity in parenthesis in the top equation) is our model of the spectrum of the light reaching our eye. That spectrum then gets projected onto spectral responsivity curves of each of the three cone classes in the eye, resulting in the L, M, and S response for this jth color patch. (An equation for the RGB pixel color values would be the same, with just a different matrix $\mathbf{E}$). If we make N distinct color measurements of the image, then we'll have N different versions of this equation, with a different vector $\vec{x}_j^s$ and different observations $\begin{pmatrix} L_j \\ M_j \\ S_j \end{pmatrix}$ for each equation.

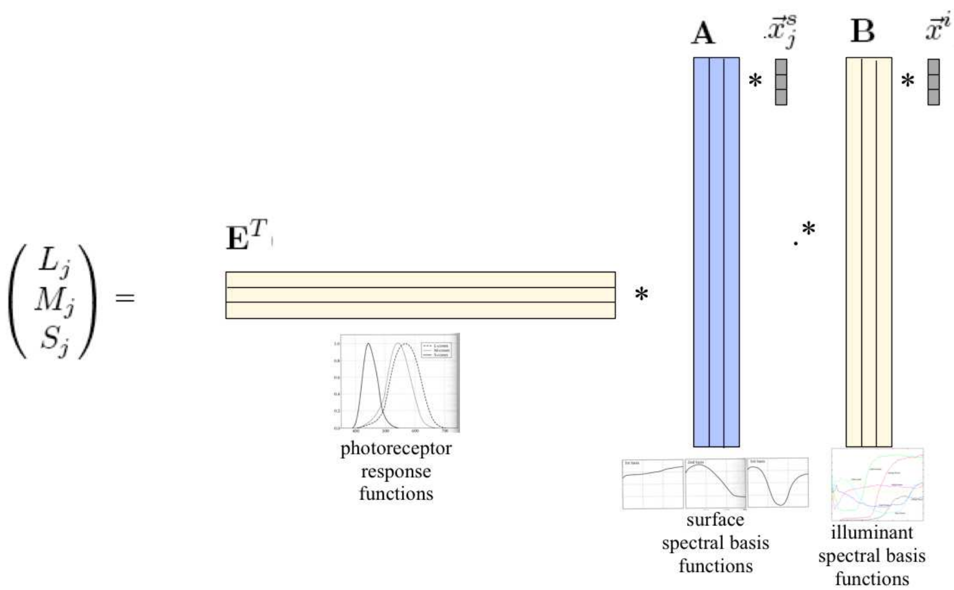


Figure 6.25: Graphical depiction of Eq. 6.10.

Like various other problems in vision, this is a bilinear problem. If we knew one of the two sets of variables, we could find the other trivially by solving a linear equation (using either a least squares or an exact solution). It's a very natural generalization of the a b = 1 problem that Antonio talked about last week.

Let's notice the degrees of freedom. We get 3 numbers for every new color patch we look at, but we also add 3 unknowns we have to estimate (the spectrum coefficients $\vec{x}_j^s$), as well as the additional three unknowns for the whole image, the illumination spectrum coefficients $\vec{x}^i$. If only surface color spectra had only two degrees of freedom, we'd catch up and potentially have an over-determined problem if we just looked at enough colors in the scene. Unfortunately, 2-dimensional surface reflectance models just don't work well in practice, so that approach doesn't work.

6.5.1 Some color constancy algorithms

So how will we solve this? Let's look at two well-known simple algorithms, and then we'll look at a Bayesian approach.

Bright equals white If we knew the true color of even a single color patch, we'd have the information we needed to estimate the 3-d illumination spectrum. One simple algorithm for estimating or balancing the illuminant is to assume that the color of the brightest patch of an image is white. (If you're working with a photograph, you'll always have to worry about clipped intensity values, in addition to all the non-linearities of the camera's processing chain). If that is the kth patch, and $\vec{x}^W$ are the known spectral basis coefficients for white, then we have

$$\vec{y}_k = \begin{pmatrix} L_k \\ M_k \\ S_k \end{pmatrix} = \mathbf{E}^T(\mathbf{A}\vec{x}^W \;.*\; \mathbf{B}\vec{x}^i) \tag{6.11}$$

which gives a linear equation that we can solve for the unknown illuminant, $\vec{x}^i$.

How well does it work? It works sometimes, but not always. The bright equals white algorithm estimates the illuminant based on the color of a single patch, and we might expect to get a more robust illuminant estimate if we use many color patches in the estimate. A second heuristic that's often used is called the **grey world assumption**: the average value of every color in the image is assumed to be grey.

Figure 6.26: An image that violates the grey world assumption.

We take the sum over all samples j on both sides of the rendering equation, Eq. (6.10). Letting $\vec{x}^G$ be the spectral basis coefficients for grey, which we equate to the average of all the basis coefficients in the image, $\frac{1}{M}\sum_j \vec{x}^s_j$ we have

$$\frac{1}{M}\sum_j \begin{pmatrix} L_j \\ M_j \\ S_j \end{pmatrix} = \mathbf{E}^T(\mathbf{A}\frac{1}{M}\sum_j \vec{x}^s_j \;.*\; \mathbf{B}\vec{x}^i) \tag{6.12}$$

$$= \mathbf{E}^T(\mathbf{A}\vec{x}^G \;.*\; \mathbf{B}\vec{x}^i), \tag{6.13}$$

where we have assumed there are M color patches in the image. Now, again, that leaves us with a linear equation to solve for $\vec{x}^i$.

This assumption can work quite well, although, of course, we can find images for which it would completely mess up, such as the forest scene of Fig. 6.26.

Using just part of the data (the brightest color, or even the average color) gives sub-optimal results. Why not use all the data, make a richer set of assumptions about the illuminants and surfaces in the world, and treat this as a **Bayesian estimation problem**? That's what we'll do now, and what you'll continue in your homework assignment.

To remind you, in a Bayesian approach, we seek to find the posterior probability of the state we want to estimate, given the observations we see. We use Bayes rule to write that probability as a (normalized) product of two terms we know how to deal with: the likelihood term and the prior term. Letting $\vec{x}$ be the quantities to estimate, and $\vec{y}$ be the observations, we have

$$P(\vec{x}|\vec{y}) = kP(\vec{y}|\vec{x})P(\vec{x}) \tag{6.14}$$

where k is a normalization factor that forces that the integral of $P(\vec{x}|\vec{y})$ over all $\vec{x}$ is one. $P(\vec{x}|\vec{y})$ is the posterior probability–in this case, the probability of a vector $\vec{x}^i$ of illumination spectral basis coefficients, and of all the vectors $\vec{x}^s_j$ of surface spectral basis coefficients, given the data, $\vec{y}$, of observations L_j, M_j, and S_j from each color patch j. $P(\vec{y}|\vec{x})$ is called the likelihood term, and $P(\vec{x})$ is the prior probability of any given illuminant and set of surface reflectance basis spectra.

The likelihood term tells us, given the model, how probable the observations are. If we assume additive, mean zero Gaussian noise, the probability that the jth color observation differs from the rendered parameters follows a mean zero Gaussian distribution. Remembering that the observations $\vec{y}_j$ are the the L, M, and S cone responses,

$$\vec{y}_j = \begin{pmatrix} L_j \\ M_j \\ S_j \end{pmatrix} \tag{6.15}$$

we have

$$P(\vec{y}_j|\vec{x}^i, \vec{x}^s_j) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{-|\vec{y}_j - \vec{f}(\vec{x}^i, \vec{x}^s_j)|^2}{2\sigma^2}, \tag{6.16}$$

For an entire collection of N surfaces, we have

$$P(\vec{x}|\vec{y}) = P(\vec{x}^i) \prod_j P(\vec{y}_j|\vec{x}^i, \vec{x}^s_j) P(\vec{x}^s_j) \tag{6.17}$$

reminder: The rendering function, $\vec{f}(\vec{x}^i, \vec{x}^s_j)$, comes from Eq. (6.10). We assume diffuse reflection from each colored surface. Given basis function coefficients for the illuminant, $\vec{x}^i$, and a matrix $\mathbf{B}$ with the illumination basis functions as its columns, then the spectral illumination as a function of wavelength is the column vector $\mathbf{B}\vec{x}^i$. We also need to compute jth surface's diffuse reflectance spectral attenuation function, the product of its basis coefficients times the surface spectral basis functions: $\mathbf{A}\vec{x}^s_j$ In our diffuse rendering model, the reflected power is the term-by-term product (we borrow Matlab notation for that, .*) of those two. The observation of the jth color is the projection of that spectral power onto the eye's photoreceptor response curves. If those photoreceptor responses are in the columns of the matrix, $\mathbf{E}$, then the forward model for the three photoreceptor responses at the jth color is:

$$\vec{f}(\vec{x}^i, \vec{x}^s_j) = \mathbf{E}^T(\mathbf{A}\vec{x}^s_j \;.*\; \mathbf{B}\vec{x}^i). \tag{6.18}$$