Contents

1 Color and color constancy

There are many benefits to sensing color. Color differences let us check whether fruit is ripe, tell whether a child is sick by looking at small changes in the color of the skin, and find objects in clutter.

We'll first describe the physics of color, then our perception of it–both the physiology and psychophysics. Finally, we'll discuss what one can infer about surfaces in the world from the color of reflected light.

1.1 Color physics

Isaac Newton revealed several intrinsic properties of light in experiments summarized by his drawing in Fig 1.1. A pinhole of sunlight enters through the window shade, and a lens focuses the light onto a prism. The prism then divides the white light into many different colors. These colors are elemental: if one of the component colors is passed through a second prism, it doesn't split into further components.

Our understanding of light and color explains such experiments. Light is a mixture of electromagnetic waves of different wavelengths. 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 different wavelengths of light as different colors. Our eyes are sensitive to only the narrow band of that electromagnetic spectrum, the visible light, from approximately 400 nm to 700 nm, from blue to deep red, respectively.

The bending of light at a material boundary is called *refraction*, and its wavelength dependence lets the prism separate white light into its component colors. A second way to separate light into its spectral components is through *diffraction*, where constructive interference of scattered light occurs in different directions for different wavelengths of light. Fig. 1.2 (a) shows a simple spectrometer, an apparatus to reveal the spectrum of light, based on diffraction from a compact disk (CD) (2004). 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 will occur at a different angle for each wavelength of

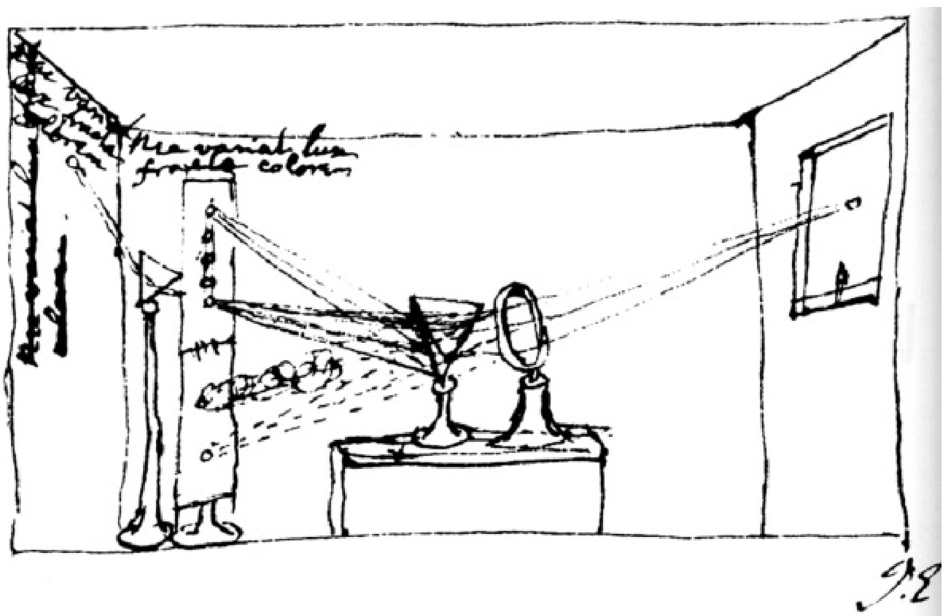
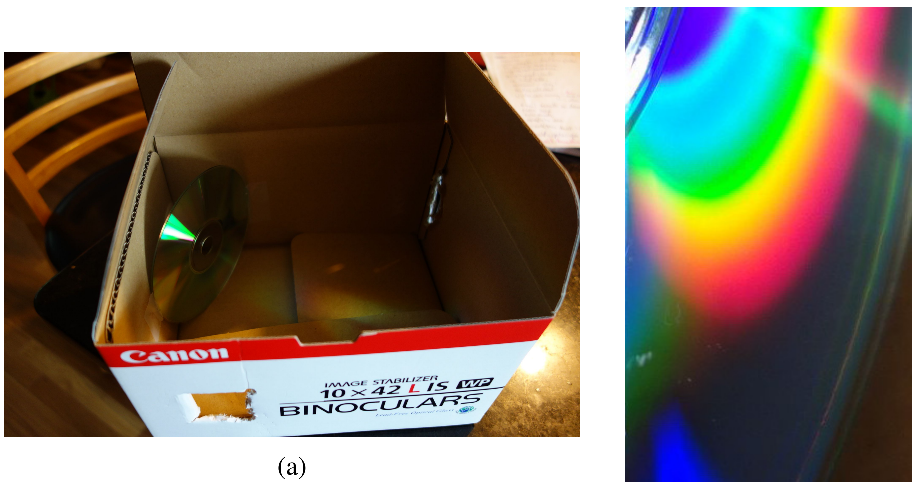


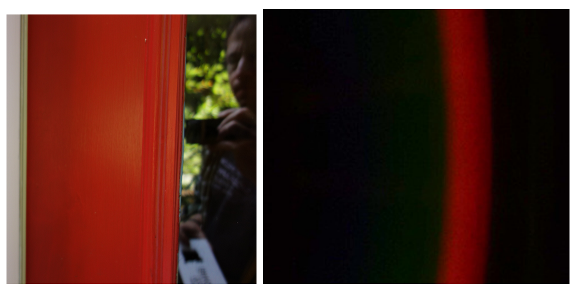
Figure 1.1
Isaac Newton's illustration of experiments with light. White light enters from a hole in the window shade at the right, where it is focused with a lens and then passes through the first prism. The prism separates the white light into different colors by bending each color a different amount. The second prism in the drawing demonstrates that those colors are elemental: as an individual color passes through the second prism, the light doesn't break into other colors. See also Wandell (1995)



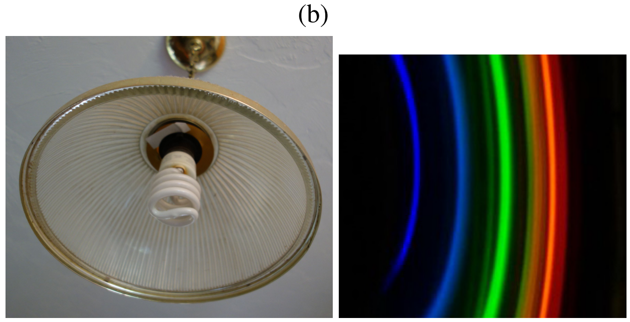
(a) (b)

Figure 1.2
(a) A simple spectrograph. Slit at right accepts light from primarily one object in the world. Light diffracted by the CD is viewed from the hole at the bottom left. The bending by diffraction is wavelength dependent, and the light from a given direction is broken into its spectral components, shown in (b).

(a)



(b)



(c)

Figure 1.3
The light spectra from some everyday objects, analyzed by the spectometer of Fig. 1.2. (a) Green leaf, with some yellowish elements, shows primarily green, with a little red (red and green make-up yellow). (b) A red door (c) A fluorescent light (when turned on) shows the discrete spectral wavelengths at which the gas fluoreses.

the light, yielding a separation of the different wavelengths of light according to their angle of diffraction. The diffracted light can be viewed or photographed through the hole at the bottom left of the photograph. The spectrometer gives an immediate visual representation of the spectral components of colors in the world. Some examples are shown in Fig. 1.3.

1.1.1 Light Power Spectra

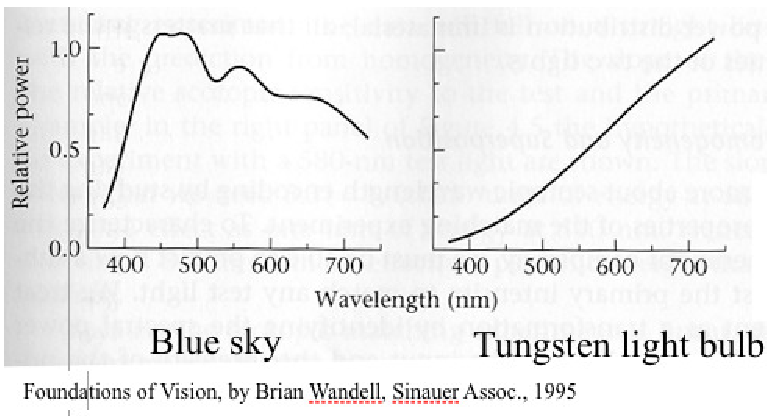


Figure 1.4
Two sources of light, and their spectra.

The light intensity at each wavelength is called the *power spectrum* of the light. It is an approximation, but a good one, to say that light's power spectrum determines its perceived color. Fig. 1.3 shows a spectrograph visualization of some light power spectra. Fig. 1.4 shows other light spectra, plotted as intensity as a function of wavelength. The spectrum of blue sky is on the left, and the spectrum of a tungsten light bulb (which will appear orangish) is on the right.

1.1.2 Light reflecting off surfaces

When light reflects off a surface, the power spectrum changes in ways that depend on the characteristics and geometry of the surface. It is those changes to the light that let us see objects and surfaces by observing how they influence the reflected light.

The interaction of light with a surface can very rich. Reflections can be specular or diffuse, and the reflected power spectrum can depend on the relative orientations of the incident light, the surface, and the observed reflected ray. The reflection of light from a surface, in such full generality, is described by the "bi-directional reflectance distribution function", or BRDF, which is covered in most computer graphic textbooks. For this discussion of reflections, we will consider only diffuse surface reflections, where the power spectrum of the reflected light, $r(\lambda)$, is proportional to the wavelength-by-wavelength product of the power spectrum of the incident light, $i(\lambda)$, times a *reflectance spectrum*, $s(\lambda)$, of the surface:

$$r(\lambda) = i(\lambda)s(\lambda) \tag{1.1}$$

This diffuse reflection model describes most matte reflections. Such wavelength-by-wavelength scaling is also a good model for the spectral changes to light caused by transmission through an attenuating filter. The incident power spectrum is then multiplied at each wavelength by the *transmittance spectrum* of an attenuating filter.

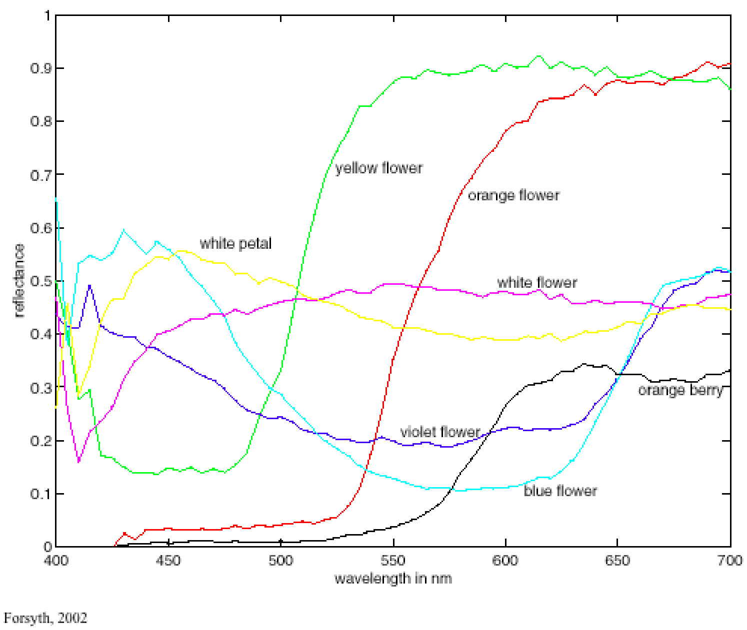


Figure 1.5
Observed spectra of light reflecting off the surface. Source: Forsyth and Ponce.

Some reflectance spectra of real-world surfaces are plotted in Figure 1.5. A white flower reflects spectral power almost equally over all visible wavelengths. A yellow flower reflects in the green and red.

When the illumination is white light, with equal power in all spectral bands, the reflected spectrum is proportional to the reflectance spectrum of the material itself. Under more general conditions, when the illumination color is unknown, a visual system needs to estimate the surface reflectance spectrum, taking the context of other observed spectra into account to estimate the overall illumination. We address that computational problem in Section 1.4.

1.2 Color Perception

Now we turn to our perception of color. We first describe the machinery of the eye, and motivated by that understanding, describe methods to measure color appearance.

1.2.1 The machinery of the eye

Figure 1.6 a is a drawing of the photoreceptors, called the rod and cones, in the retina of the eye. The tall receptors are the rods, used in low-light levels, and the short ones are the color-selective cones. Interestingly, the light enters from the bottom of the drawing, passing through the nerve fibers and blood vessels before reaching the photosensitive detectors at the top of the image.

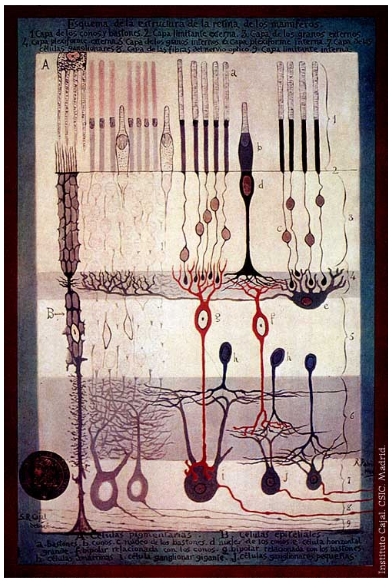
An instrument can measure the light power spectrum at hundreds of different wavelengths within the visible band, yielding hundreds of numbers to describe the light power spectrum. But a useful description of the visual world can be obtained from a much lower dimensional description of the light power spectrum. The human visual system analyzes the incident light power spectrum with only three different classes photoreceptors, called the L, M, and S cones because they sample at the long, medium, and short wavelengths. This gives the human visual system a 3-dimensional description of light, with each photoreceptor class taking a different weighted average of the incident light power spectrum.

Figure 1.6 b shows a colorized image of a live human retina. The black-and-white photograph of the retinal mosaic has been colored by the experimenters to indicate which cone type is at which location in the subject's eye Hofer et al. (2005). The L cones are colored red, the M cones green, and the S cones are colored blue. Note the hexagonal packing of the cones in the retina, and the stochastic assignment of L, M, and S cones over space. Figure 1.6 c shows the spectral sensitivity curves for the L, M, and S cones.

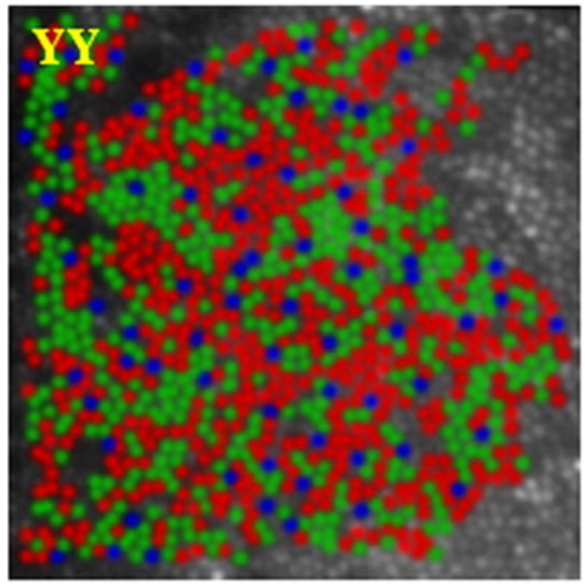
If the matrix C consists of the spectral sensitivity curves of the L, M, and S cones in 3 rows, and the vector $\vec{t}$ is a column vector of the spectrum of light incident on the eye, then the L, M, and S cone responses will be the product,

$$\left\{ \begin{array}{c} L \\ M \\ S \end{array} \right\} = C\vec{t} \tag{1.2}$$

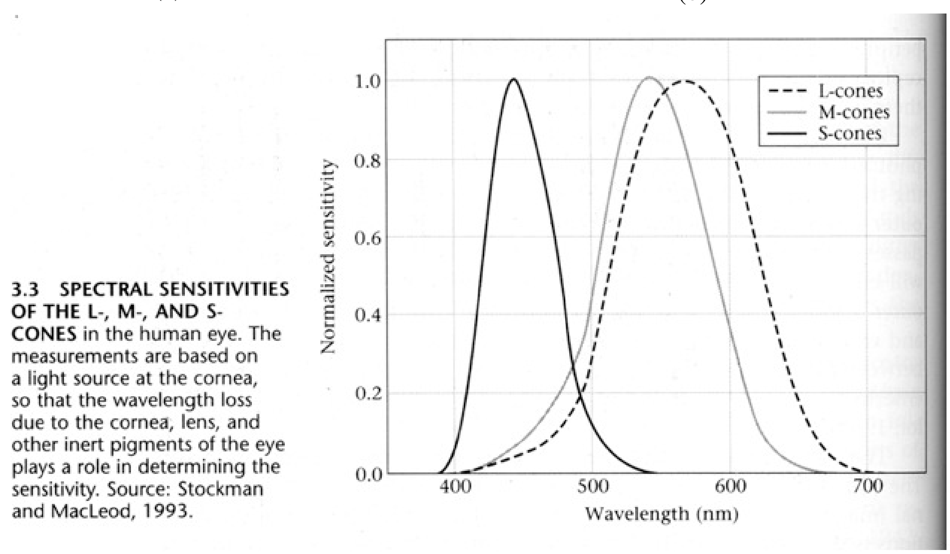
The fact that our eyes have three different classes of photoreceptors has many consequences for color science. It determines that there are three primary colors, three color layers in photographic film, and three colors of dots needed to render color on a display screen. In the next section, we describe the color matching experiments that led to much of our understanding of human color vision.



(a)



(b)



(c)

Figure 1.6
(a) Drawing of the eye's photoreceptors by ramon y cajal. (b) Lateral slice through an eye, with cone receptor class identities indicated by red, green, or blue overlay, to signify the long, medium, and short wavelength receptors. From Hofer et al. (2005). (c) Photoreceptor sensitivities as a function of light wavelength (from Wandell (1995)).

1.2.2 Color matching

Color science tells us how to analyze and reproduce color. We can build image displays so that the output colors match those of some desired target, and manufacture items with the same colors over time. Much industry revolves around the ability to repeatably control colors. Colors can be trademarked (Kodak Yellow; IBM Blue, etc) and we have color standards for foods. Figure 1.7 shows a USDA kit for evaluating food color, and a chart showing french fry color standards.

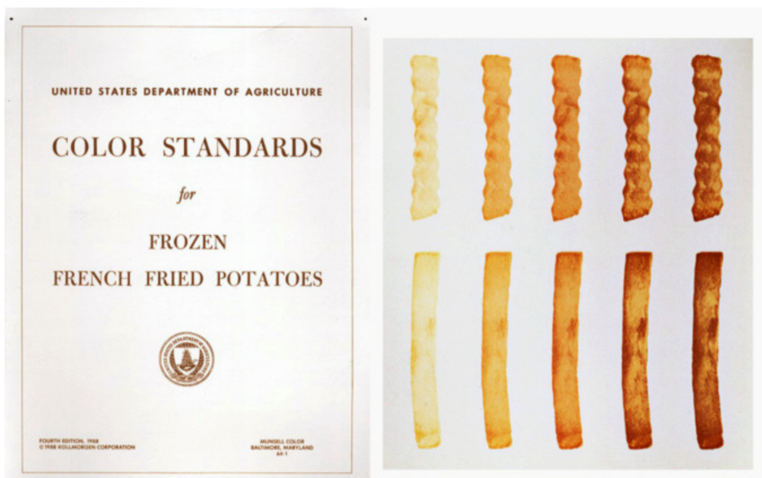


Figure 1.7
The USDA color standards for French fried potatoes, one of many color standards.

One of the tasks of color science is to predict when a person will perceive that two colors match. For example, we want to know how to adjust a display to match the color reflecting off some colored surface. Even though the spectra may be very different, the colors can often be made to match.

It is possible to infer human color matching capabilities by knowing the spectral sensitivity curves of the receptors shown in Fig. 1.6 (c), taking into account coloration of the cornea and lens. However, it was through measurements of human color matching judgments that procedures for matching colors were originally derived.

We try to match a color with a combination 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. This stems from the fact that we have three classes of photoreceptors in our eyes. It has been found that these color matches are transitive–if color A matches some particular combination of primaries, and color B matches the same combination of primaries, then color A will match color B.

Thus the amount of each primary requred to match a color can serve as set of coordinates indicating color.

In this section, we're assuming that color appearance is determined by the spectrum of the light arriving at the eye. In practise, color appearance can be influenced by many other factors, such as the eye's state of brightness adaptation, the ambient illumination, and the surrounding colors. To control for those variables, care must be taken to view the color comparisons under repeatable, controlled surrounding colors. 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. A grey surround field is placed around the viewing aperture, giving a view to the subject that looks something like that of Fig. 1.8 (a), right hand side (from Wandell (1995)). We assign a triplet of numbers to the color appearance of a surface: the amount of each of the three primary colors that is required to match a given test color.

1.2.3 Box: Color matching experiments

Here is the procedure for matching an arbitrary color with an additive combination of primary lights. We shine the test light on the left in Fig. 1.8 (a), and originally all the primary lights are turned off, so the right hand side is black.

Now we light up some combination of the primaries and we adjust their amounts until we get a color match. 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.

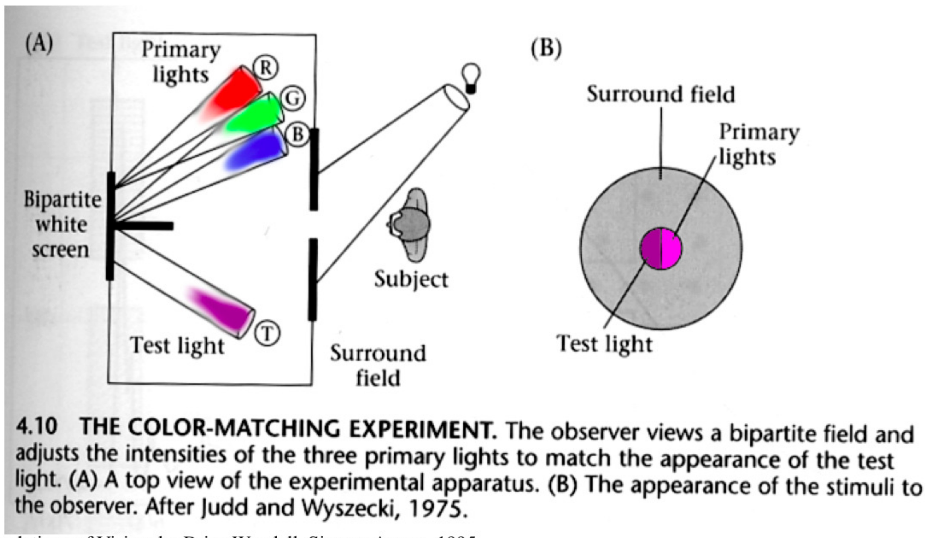
What if no combination of the three selected primiaries matches the test color? Fig. 1.8 (b) shows an example of that. We can exploit another property of color matching: if two colors match, and we add the same amount of any other light spectrum to both colors, the resulting modified colors will also match. In other words, if color A_1 matches color B_1, and color A_2 matches color B_2, then the sum of spectra of colors A_1 and A_2 will match the sum of the spectra of colors B_1 and B_2. This linearity follows from Eq. (1.2).

Exploiting that fact, we can always match any input test color if we allow "negative light", adding light from one or more of the primary lights to the test color in order to match the modified test color to some combination of the remaining primary lights.

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, Fig. 1.9.

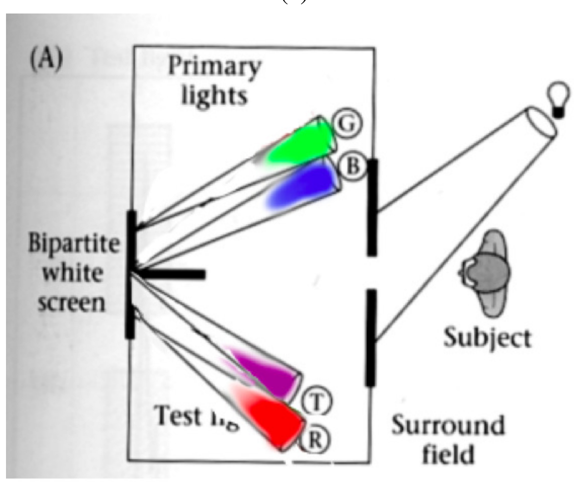
1.2.4 Linear algebraic interpretation of color perception

From Eq. (1.2), the task of color measurement is simply the task of finding the projection of any of the possible spectrum into the special 3-d subspace defined by the cone spectral response curves. Any basis for that 3-d subspace will serve that task, so the three basis functions do not need to be the color sensitivity curves of Eq. (1.6) themselves. They can be any linear combination of them, as well. We seek to predict the cone responses



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

(a)



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

(b)

Figure 1.8
Color matching method (a) Left: showing a controlled combination of primary colors shining on a screen, for an observer to compare against a test light. Right: Example of what the observer sees while making a “match” judgement . (b) Sometimes to create a match with a given test light, one of the primary lights needs to be added to the test light. Mathematically, that can be modelled as requiring a negative amount of the given primary to create a match.

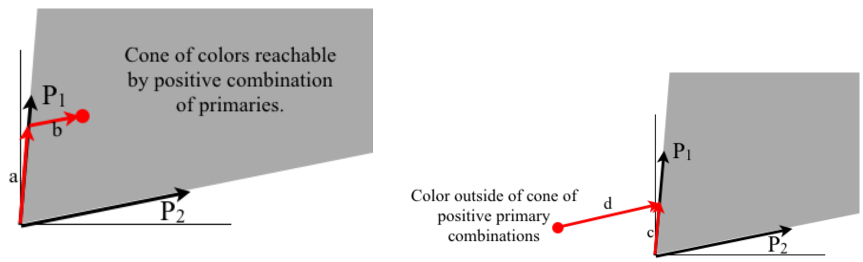


Figure 1.9
How to treat negative colors: 2-d illustration of color spectra space. With positive combinations of the primary light basis functions, we can only directly match colors within the cone of the primary light basis functions. The left hand side shows a spectrum, denoted by a point, being matched by a sum of only positive amounts of the primaries $\vec{P_1}$ and $\vec{P_2}$. To represent a color spectrum (point) outside the positive cone of vectors $\vec{P_1}$ and $\vec{P_2}$, we first add positive amounts of the vectors $\vec{P_1}$ or $\vec{P_2}$ to the out-of-cone color to bring it to the cone. In the right-hand figure, we say that the out-of-cone color is represented as $c\vec{P_1} + d\vec{P_2}$.

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.

We start with a set of *any* three linearly independent primary lights, ie, none of the three spectra can be written as a linear combination of the other two. The idea is this: if we find spectral curves which, when taking the projection of an input spectrum, give us the controls for each primary to match the input color, then we have found a basis for the 3-dimensional cone response space. This is because 3-d projection vectors that always lead to matched colors must be basis vectors for that same 3-d space.

1.2.4.1 Color matching functions Here's how we can find such basis vectors, called "color matching functions", for any give 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.

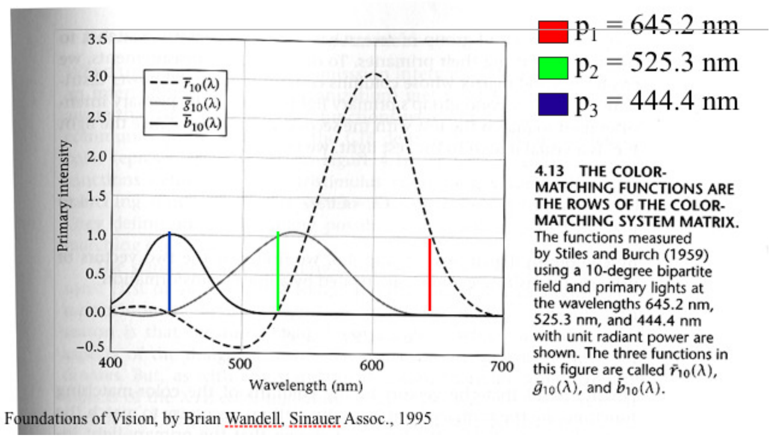


Figure 1.10
Psychophysically measured color matching functions. Figure from Wandell (1995).

Figure 1.10 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 matric bfC be the color matching functions arranged in rows,

$$C = \begin{pmatrix} c_1(\lambda_1) & \dots & c_1(\lambda_N) \\ c_2(\lambda_1) & \dots & c_2(\lambda_N) \\ c_3(\lambda_1) & \dots & c_3(\lambda_N) \end{pmatrix} \tag{1.3}$$

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 $C\vec{t}$.

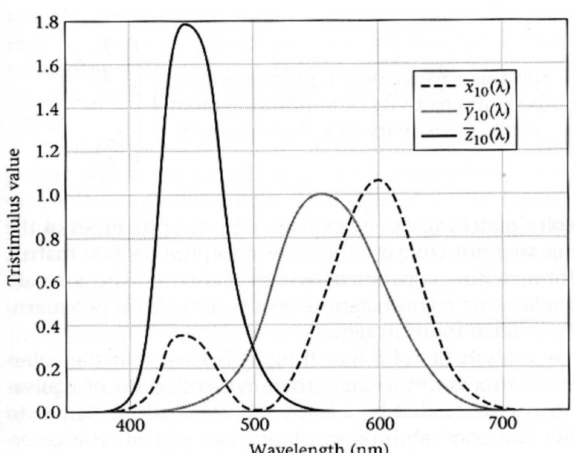
If the model of Eq. (1.2) is correct, then measured curves of Fig. 1.10 (and analogous ones, measured using different primary lights) must be a linear combination of the human eye's color sensitivity curves. Let the spectra of the primary lights be in the columns of the matrix P_0, and let the corresponding measured color matching curves be in the rows of the matrix C_0. Then for any light spectrum column vector, $\vec{t}$, the color matching experiments assure us that the combination of primary lights given by the 3 by 1 vector $C_0\vec{t}$ will be a perceptual match to the spectrum $\vec{t}$, or

$$C\vec{t} = CP_0C_0\vec{t} \tag{1.4}$$

This holds for all vectors $\vec{t}$, so we can omit $\vec{t}$ from both sides of the above equation. Denoting the 3x3 matrix, CP_0 as R, we have

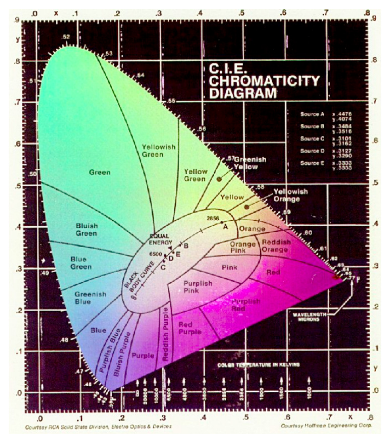
$$C = RC_0, \tag{1.5}$$

showing that the rows of the physically measured color matching functions, C_0 must be a linear combination of the human eye's spectral sensitivity curves, C.



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

(a)



(b)

Figure 1.11
(a) CIE color matching functions. (b) The space of all colors (intensity normalized), as described in the CIE coordinate system.

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.

1.2.5 CIE color space

One color standard 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. 1.11.

An unfortunate property of the CIE color matching functions is that there is no all positive set of color primaries 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 calculates $x = \frac{X}{X+Y+Z}$ and $y = \frac{Y}{X+Y+Z}$.

1.2.6 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.

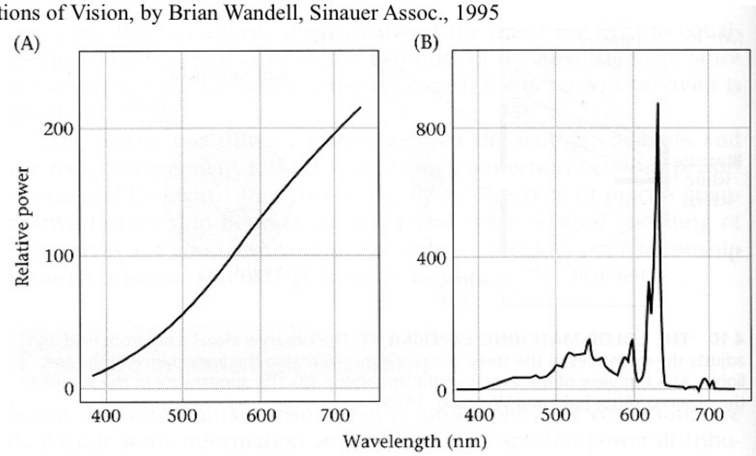
There's a sense that our eyes are missing much of the possible visual action. There's a high-dimensional space of colors out there, and we're only viewing projections onto a 3-d subspace of that.

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, for notation, 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

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



4.11 METAMERIC LIGHTS. Two lights with these spectral power distributions appear identical to most observers and are called metamers. (A) An approximation to the spectral power distribution of a tungsten bulb. (B) The spectral power distribution of light emitted from a conventional television monitor whose three phosphor intensities were set to match the light in panel A in appearance.

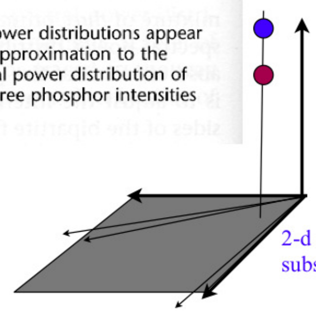


Figure 1.12
The two spectra in the figure above, from Wandell (1995), are colors that match perceptually: daylight, and the spectrum a monitor adjusted to match daylight. The figure at right shows a graphical rendition of the projection from the high-dimensional space of power spectra onto a lower-dimensional subspace, representing the 3-d space of human color perception. The red and blue dots in the higher-dimensional space are “metameric” in that they project to the same location in the lower-dimensional subspace.

its representation in a primed coordinate system, we just have to multiply this metameric spectrum by the color matching functions for the primed color system:

$$\vec{x}' = \mathbf{CP}'\vec{x} \tag{1.6}$$

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

From the reflected or transmitted light, we seek to learn about what did the attenuating or the reflection. Either for light coming directly from a source, or for light reflecting off a surface, it's very useful to characterize it's spectrum in terms of the color appearance.

1.2.6.1 Simplified Color Spectra–include??? It is helpful to develop a feel for the color appearance of different light spectra. Later, after we describe human color sensing, we can make the intuitions precise as we will be able to compute the color appearance of a given light power spectrum.

The visible spectrum lies roughly in the range between 400 and 700 nm, see Fig. 1.13. We can divide the visible spectrum into three one-hundred nm bands, and study the appearance of light power spectra where power is present or absent from each of those three bands, in all of the eight (2^3) possible combinations.

Light with spectral power distributed in just the 400 to 500 nm wavelength band will look some shade of blue, the exact hue depending on the precise distribution. Light in the 500 - 600 nm band will appear greenish. Most distributions within the 600 - 700 nm band will look red.

White light is a mixture of all spectral colors. A spectrum of light containing power evenly distributed over 400 - 700 nm would appear white. Light with no power in any of those three bands, that is, darkness, appears black.

There are three other spectral classes left in this simplified grouping of spectra: spectral power present in two of the spectral bands, but missing in the third. Cyan is a combination of both blue and green, or roughly spectral power between 400 and 600 nm. In printing and color film applications, this is sometimes called "minus red", since it is the spectrum of white light, minus the spectrum of red light. The blue and red color blocks, or light in the 400-500nm band, and in the 600-700nm band, is called magenta, or minus green. Red and green together, with spectral power from 500-700 nm, make yellow, or minus blue.

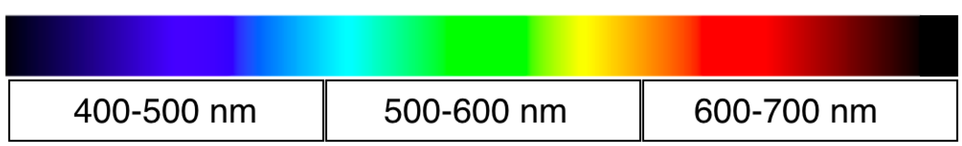


Figure 1.13
The approximate color appearance of light over different spectral regions.

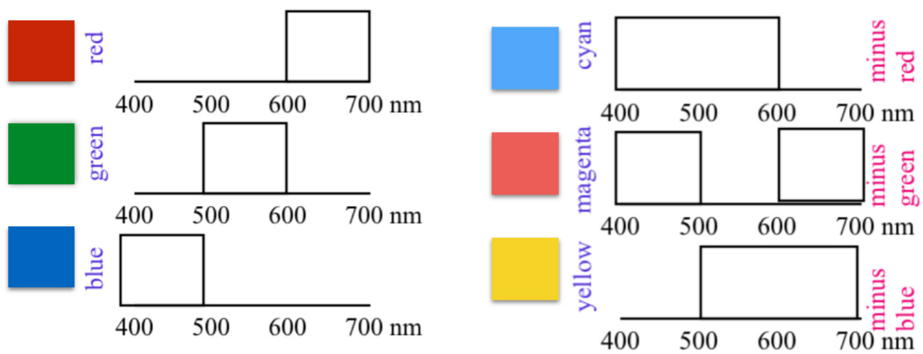


Figure 1.14
Cartoon model for the reflectance spectra of observed colors

1.2.7 Color mixing

In art or photography, we often talk of ... (relate to color primaries).

Figure 1.15
The cyan of Windex and the yellow of Joy combine to give a green color.

Colors and color mixing are a delightful aspect of our visual experience. Fig. 1.15 shows an everyday example of color mixing: yellow and cyan colors combining to give a green.

Physically, color mixing involes combining the spectra of the two colors to be mixed to create the power spectrum of the mixed color. Various processes cause colors to mix, but the results can be divided into two broad classes of mixing, called additive and subtractive. Under additive color mixing, the power spectra add together to form the spectrum of the resulting color. This model of mixing covers the case of many small color elements for which the appearance is fused, such as tiny elements of a display, or the case of several light projectors pointing at the same screen. 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, red and green combine to give yellow, as can be seen from the cartoon models of Figs. 1.14 and 1.15.

A second way colors combine is called subtractive color mixing, but might better be called multipliciative color mixing. 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 of one color reflects diffusely off a surface of another color, or passes through a sequence of attenuating spectral filters, such as with photographic film, paint, optical filters, and crayons. Under the subtractive color mixing 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. Under subtractive color mixing, red and green combine to give black.

Figure 1.16 shows examples of color mixing.

Because the spectrum depends on surfaces, it is very useful to measure the spectrum of reflected light, and we discuss how the eye does that in the following section.

1.3 Spatial Resolution and Color

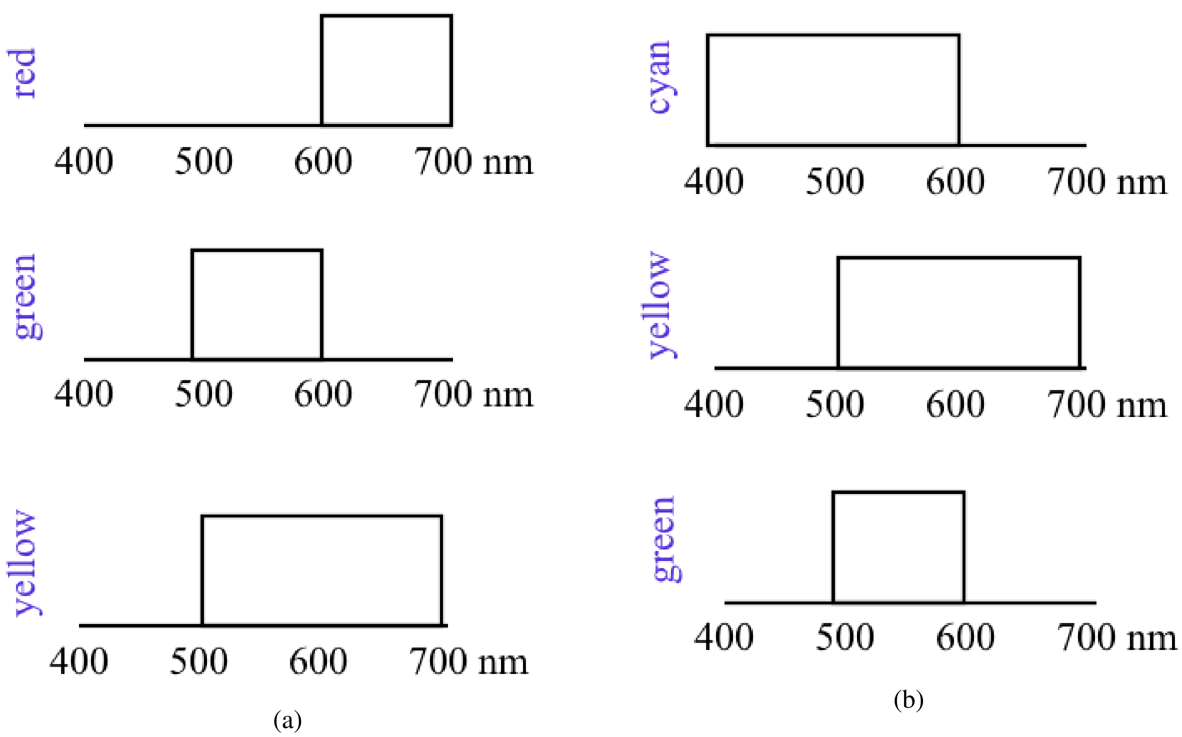
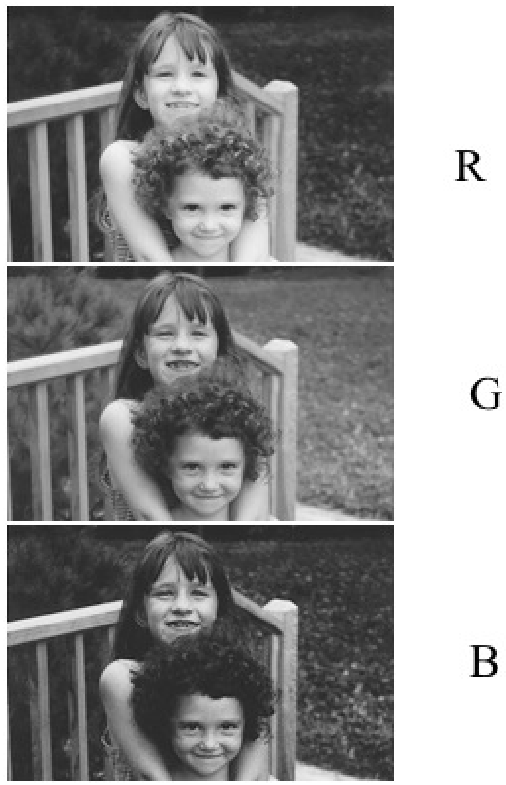
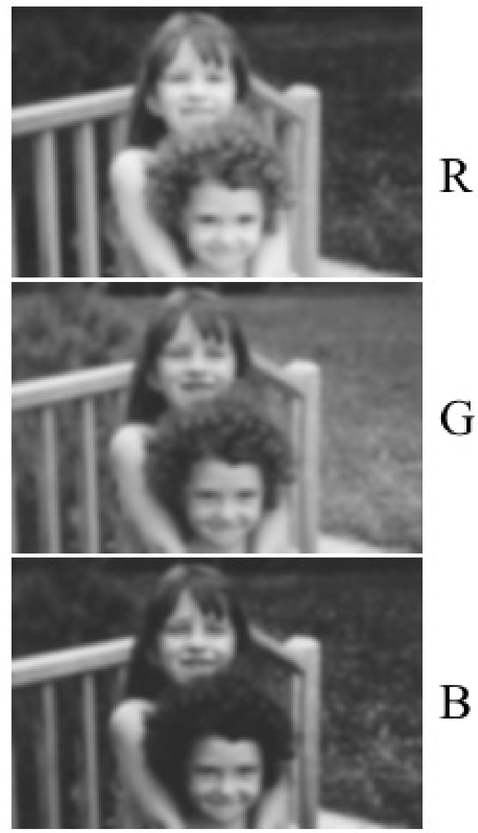


Figure 1.16
Examples of color mixing, in the world of cartoon color spectra. (a) Additive, (b) subtractive. The spectra and the resulting colors.

(a) Original



(b) R, G, B components



(c) blurred R, G, B components

Figure 1.17
(a) Original image. (b) RGB components (c) RGB components, each blurred.

(a) R component blurred (b) G component blurred (c) B component blurred

Figure 1.18
(a) R component blurred, G and B components sharp. (b) (c)

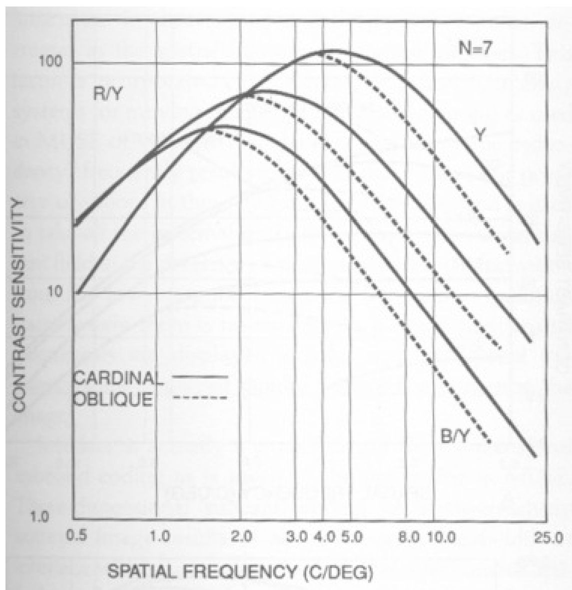
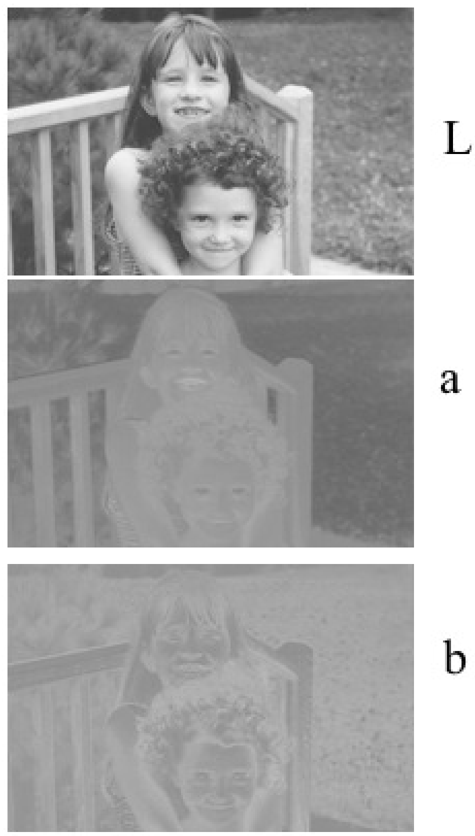


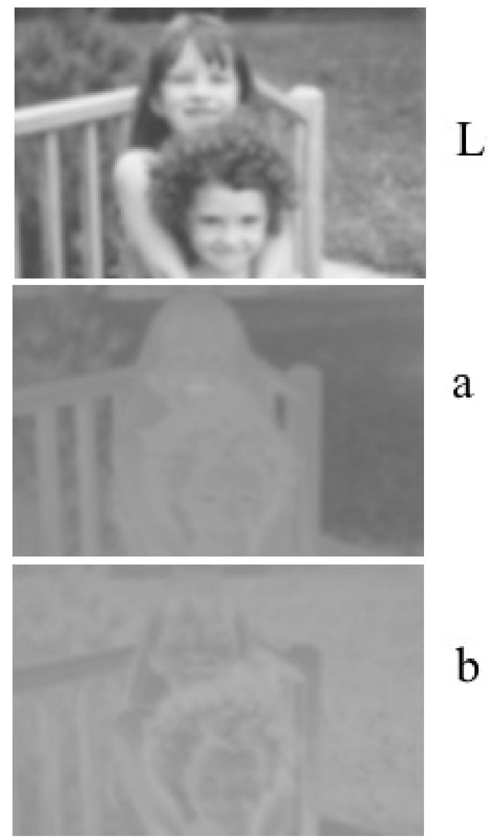
Figure 6.1
Contrast sensitivity threshold functions for static luminance gratings (Y) and isoluminance chromaticity gratings (R/Y,B/Y) averaged over seven observers.

Figure 1.19
Human spatial frequency sensitivity in R, G, B and L, a, b color representations

(a) Original



(b) L, a, b components



(c) blurred L, a, b components

Figure 1.20
(a) Original image. (b) Lab components (c) Lab components, each blurred.

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

Figure 1.21
(a) L component blurred, a and b components sharp. (b) (c)

1.3.1 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 at 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

$$\begin{pmatrix} \vdots \\ S(\lambda) \\ \vdots \end{pmatrix} \approx \begin{pmatrix} \vdots & \vdots & \vdots \\ u_1(\lambda) & u_2(\lambda) & u_3(\lambda) \\ \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \tag{1.7}$$

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

$$D = U * \Lambda * V' \tag{1.8}$$

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.

Here's 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, made the same way year after year. (So iconic that this woman, a dedicated color scientist, I presume, has tatooed a Macbeth color checker on her arm! Alas, I'm sure the tatoo colors are only an approximation to the real Macbeth colors).

The reflectance spectra of each Macbeth color chip has been measured. All those spectra are pretty well approximated by a 3-dimensional linear model, as you can see from these plots.

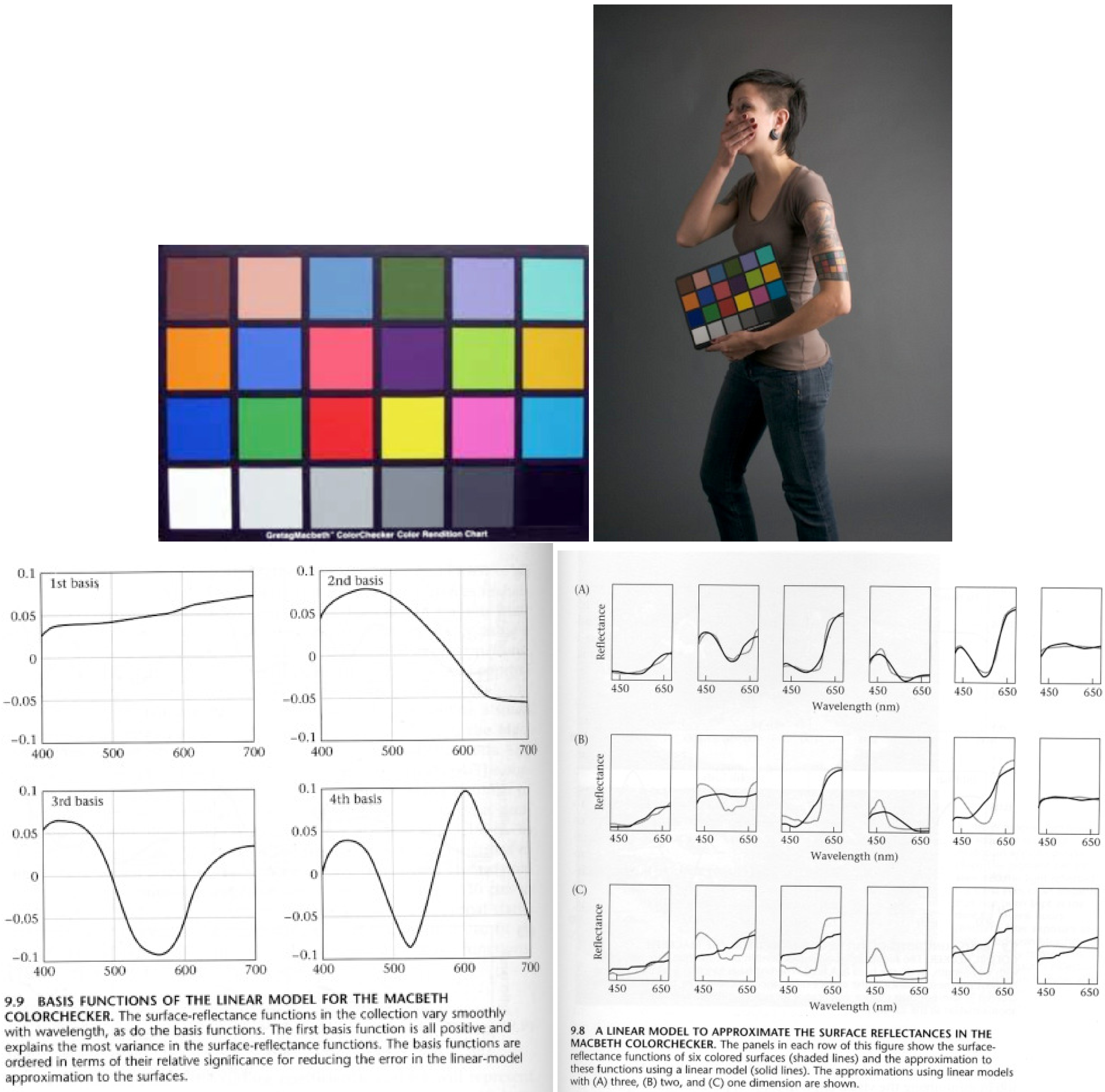


Figure 1.22
Macbeth color checker. iconic status. spectra. bottom two figures from Foundations of Vision, by Brian Wandell, Sinauer Assoc., 1995

1.4 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 color constancy demo of Fig. 1.23, we'll show an example where the identical spectral distribution arriving at your eye leads to a very different color percept. What's going on? The visual system needs to perceive the color of surfaces, but the data it gets is the wavelength-by-wavelength product of the surface color and the illuminant color. 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 spectrum arriving at the eye.

Figure 1.23
color constancy demo

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 illuminant or all the surface colors, 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.

1.4.0.1 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:

$$\begin{pmatrix} L_j \\ M_j \\ S_j \end{pmatrix} = \mathbf{E}^T(\mathbf{A}\vec{x}^s_j .* \mathbf{B}\vec{x}^i) \tag{1.9}$$

Figure 1.24 shows a graphical diagram showing the vector and matrix sizes that I hope makes things a little clearer. 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 $\vec{x}^s_j$ and different observations $\begin{pmatrix} L_j \\ M_j \\ S_j \end{pmatrix}$ for each equation.

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 $\vec{x}^s_j$), 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.

1.4.1 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,

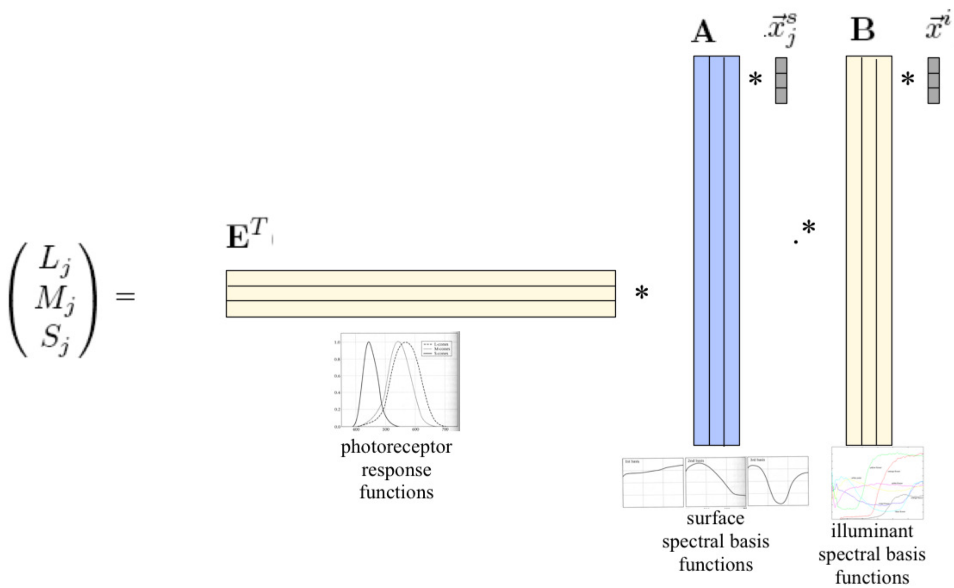


Figure 1.24
Graphical depiction of Eq. 1.9.

then we have

$$\vec{y}_k = \begin{pmatrix} L_k \\ M_k \\ S_k \end{pmatrix} = \mathbf{E}^T(\mathbf{A}\vec{x}^W \; .* \; \mathbf{B}\vec{x}^i) \tag{1.10}$$

which we can solve for the unknown illuminant, $\vec{x}^i$.

How well does it work? It works sometimes, but not always. On the left is a picture for which the bright equals white algorithm would probably work (although I haven't checked it. On the right is one where I don't think it would work.

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 1.25
An image that violates the grey world assumption.

Taking the sum over all samples j on both sides of the rendering equation, and letting $\vec{x}^G$ be the spectral basis coefficients for grey, gives

$$\frac{1}{N}\sum_j \begin{pmatrix} L_j \\ M_j \\ S_j \end{pmatrix} = \mathbf{E}^T(\mathbf{A}\frac{1}{N}\sum_j \vec{x}^s_j \; .* \; \mathbf{B}\vec{x}^i) \tag{1.11}$$

$$= \mathbf{E}^T(\mathbf{A}\vec{x}^G \; .* \; \mathbf{B}\vec{x}^i) \tag{1.12}$$

$$\tag{1.13}$$

Then, again, we just have 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 this forest scene here.

Using just part of the data (the brightest color, or 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

$$P(\vec{x}|\vec{y}) = kP(\vec{y}|\vec{x})P(\vec{x}) \tag{1.14}$$

where k is a normalization factor that forces that the integral of $P(\vec{x}|\vec{y})$ over all $\vec{x}$ is one.

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,

$$\vec{y}_j = \begin{pmatrix} L_j \\ M_j \\ S_j \end{pmatrix} \tag{1.15}$$

we have

$$P(\vec{y}_j|\vec{x}^i, \vec{x}^s_j) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{-|\vec{y}_j - \vec{f}(\vec{x}^i, \vec{x}^s_j)|^2}{2\sigma^2}, \tag{1.16}$$

For an entire collection of N surfaces, we have

$$P(\vec{x}|\vec{y}) = P(\vec{x}^i) \prod_j P(\vec{y}_j|\vec{x}^i, \vec{x}^s_j) P(\vec{x}^s_j) \tag{1.17}$$

reminder: Here's what's inside the rendering function, $\vec{f}(\vec{x}^i, \vec{x}^s_j)$. 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:

$$\vec{f}(\vec{x}^i, \vec{x}^s_j) = \mathbf{E}^T(\mathbf{A}\vec{x}^s_j \; .* \; \mathbf{B}\vec{x}^i). \tag{1.18}$$

1.4.1.1 Eq. (1.18), in component form We can find the linear solution of Eq. (1.18), for a given illuminant vector, $\vec{x}^i$, and assuming no noise in the observations. Let's write everything out in component form, in order to do the calculation carefully. Let's assume we're only fitting the x^s_{kj} to the jth color patch observation, and therefore omit all subscripts j, for simplicity. So x^s_k will mean the kth reflectance basis component (of the jth patch). w indexes wavelength values.

$$\begin{aligned} y_n &= \sum_w E_{nw} \sum_k A_{wk} x^s_k \sum_m B_{wm} x^i_m && (1.19) \\ &= \sum_k x^s_k \sum_w E_{nw} A_{wk} \sum_m B_{wm} x^i_m && (1.20) \\ & && (1.21) \end{aligned}$$

If we define the n, k components of a matrix D to be

$$D_{nk} = \sum_w E_{nw} A_{wk} \sum_m B_{wm} x^i_m, \tag{1.22}$$

then we have

$$\vec{y} = \mathbf{D} \vec{x}^s \tag{1.23}$$

If $\mathbf{D}$ is invertible, then we have

$$\vec{x}^s = \mathbf{D^{-1}} \vec{y} \tag{1.24}$$

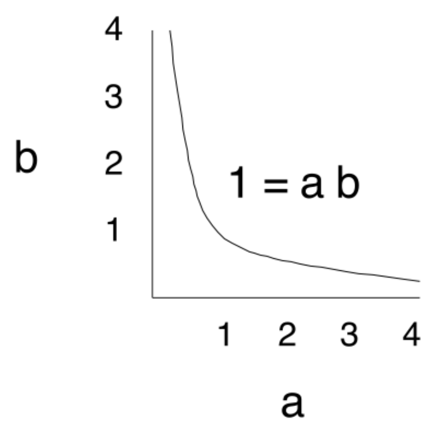


Figure 1.26
y = ab problem

I want to remind us about the similarities between the 1 = ab problem, and our color constancy problem. While we can't draw out the high-dimensional likelihood function for the color constancy problem, conceptually it's very similar to that of this 1=ab problem (and various other vision problems share these same characteristics). Many different settings of the parameters can explain our data, giving what we call a "likelihood ridge". For the 1=ab problem, that gives a 1-d ridge of parameter settings that explain the same data. For the color constancy problem, with 3-d data, surface parameterizations and illuminant parameterization, the likelihood ridge is 3-dimensional.

How do we pick from the many feasible solutions on the ridge, all with the same likelihood value? The priors will let us distinguish different values on the likelihood ridge.

In the problem set, you'll fit Gaussians to model the observed prior distribution of surface and illuminant basis function coefficients.

Another difference for different positions along the likelihood ridge is the "width" of the ridge. At some positions, only a very precise specification of all the parameter values will explain the observations. At other positions, we have more slop in the parameters, and many different nearby parameter settings also explain the data. In a Bayesian framework, this is most naturally quantified with the *loss function*, which specifies the penalty for guessing wrong. Let $\hat{\vec{x}}$ be your estimate of the parameters, $\vec{x}$. Then $L(\hat{\vec{x}}, \vec{x})$ is the loss incurred by guessing $\hat{\vec{x}}$ when the true value was $\vec{x}$. With the posterior probability, we can calculate the expected loss, $\bar{L}(\hat{\vec{x}}, \vec{x})$

$$\bar{L}(\hat{\vec{x}}, \vec{x}) = \int_{\vec{x}} L(\hat{\vec{x}}, \vec{x}) P(\vec{x}|\vec{y}) \tag{1.25}$$

We often use a loss function which is only a function of $\hat{\vec{x}} - \vec{x}$.

Bayes rule lets us find a posterior probability for $\vec{x}$, given the observations $\vec{y}$. The final stage of Bayesian estimation is to go from that function, $P(\vec{x}|\vec{y})$, to a single best guess value, $\hat{\vec{x}}$, a point estimate.

The two most commonly used point estimates are called the MAP and MMSE estimates, and in your homework for this material, you'll use either one of those for this color constancy problem.

The MAP estimate stands for "maximum a posteriori", Latin for "just take the max of the posterior distribution". While quite a natural thing to do, it can suffer from various problems, since it only depends on the single maximum valued point of the posterior probability. The loss function implied by the MAP estimate assigns a constant penalty for all guesses, except for a precisely correct guess, which receives high reward. This penalty structure doesn't make sense for perceptual tasks, for which nearly the right answer is often just as good as precisely the right answer.

Another very common estimate is MMSE, "minimum mean squared error". This is the mean of the posterior probability, which, as the name implies results in an estimate which minimizes the expected squared error in the estimated parameter. The homework assignment, excerpted on these slides, gives details of how to find each of those estimates for this problem. Again, for perceptual problems, this loss function often doesn't make sense, although in practice an MMSE estimate is often very good. (Although for the 1 = ab problem, limited to $0 < a, b < 4$, the MMSE estimate is relatively far away from any feasible solution to the problem).

Bibliography

Hofer, H., J. Carroll, J. Neitz, M. Neitz, and D. R. Williams. 2005. Organization of the human trichromatic cone mosaic. *The Journal of Neuroscience* 25 (42): 9669–9679.

Wandell, Brian. 1995. *Foundations of vision*. Sinauer Assoc..

2004. A CD spectrometer, http://www.cs.cmu.edu/ zhuxj/astro/html/spectrometer.html. `http:\/\/www.cs.cmu.edu\/\\texttildelowzhuxj\/astro\/html\/spectrometer.html`.