Digital Camera Sensor Colorimetry

Douglas A. Kerr

Issue 1.1
March 15, 2010

ABSTRACT AND INTRODUCTION

We look to a digital camera sensor to discern the color of the light landing on it and report that in some system of “color coordinates”. It does this by essentially emulating the mechanism used by the human eye to discern color. But for practical reasons, that emulation is imperfect. In most practical sensor designs, two “instances” of light having different spectrums but nevertheless having the same color but may be reported by the sensor as having different colors.

The story of this involves many concepts. Near its end, we see the motivations for, and the nature, certain compromises in practical sensor design, and we see how their adverse affects are mitigated.

Language used to describe the color response properties of a sensor is explained.

Included is a discussion of the role in all this of the concept of “white balance color correction”.

Extensive background review is given to several areas pivotal to the overall presentation, such as the nature of color, the concept of color spaces, and so forth.

BACKGROUND

Our work on this topic requires an accurate grasp of a number of fundamental matters in the area of colorimetry. Here I will review some of those topics and other critical matters.

The color filter array (CFA) digital camera sensor

We will shortly begin speaking of a digital camera sensor array in which there is, at each pixel location of the image on the sensor, an “organ” for determining the color of the light at that point, comprising three photodetectors of differing “spectral response”.

In fact, for the preponderance of digital cameras today, this is not quite so. Rather, these cameras use a color filter array (CFA) sensor. It has photodetectors of three differing spectral responses “interleaved” in a repeating pattern across the pixel locations of the image.

By a clever technique (a process called *demosaicing*), the camera extracts from this array a "best estimate" of the color of the light at each pixel location.

The colorimetric matters discussed in this article, revolving around the spectral sensitivity of the three kinds of photodetectors, apply equally to either arrangement.[1] It is perhaps easiest to follow the action by thinking in terms of the less-common "three photodetectors at each pixel location" model, and I will in general speak as in those terms.

Color

Color is a property that distinguishes among different kinds of light. It is defined wholly in terms of human perception.

I'll state this mantra now, and in bold type, as it must be kept clearly in mind during the work to come:

> **If two instances of light appear to a viewer to be the same color[2], they are the same color.**

Color, as we use the term in the technical sense, is usually recognized by the viewer as having two aspects:

- L*uminance*, which we can think of for our purposes as an indication of the "brightness" of the light.[3]
- C*hromaticity*, the property that distinguishes red from blue, and red from pink. (This is the property that lay people typically think is meant by "color", not realizing that formally the concept embraces luminance as well.)

The dimensionality of color

It has been long recognized that, as perceived by most human observers, any color of light can be specified my merely stating three

---

[1] Some cameras, such as the Sigma models using the "Foveon" sensor, actually do have a three-photodetector organ at each pixel location.

[2] I have to add, for rigor, "if observed under the same conditions".

[3] There is a subtle but important formal distinction between *luminance* and *brightness*, but for our purpose here we can ignore it.

numerical values.[4] That is, color is three-dimensional in the mathematical (not geometric) sense.

There are, however, many different schemes under which these three numerical values can be defined. These schemes, when fully specified, are called *color spaces*.

What determines the color of light?

Color is not a direct physical property like the temperature or pressure of a gas. We can, however, ascertain the color of a “sample” of light by physical measurements which will predict for us the eye’s response to it.

The physical property of the light that gives it its color is its “spectrum”, the “plot” of distribution of the power in it over the range of wavelengths that can affect the eye (the “visible wavelengths”).[5]

The “shape” of the plot determines the *chromaticity* of the light; its overall “vertical scale” determines the *luminance*. That is, if we have two different instances of light, whose spectrums have the same shape, but for one instance is proportionately “stretched” vertically, the two instances have the same chromaticity, but the second one has a greater luminance.

In the other direction, things are not nearly so tidy. We can have two instances of light with the same color that nevertheless have different spectrums. In fact there are an infinity of spectrums that will have any given color.

This situation is called *metamerism*, and different spectrums having the same color are called *metamers*.

How the eye determines color

It has been determined that (for fairly substantial luminance) the eye observes each tiny element of the image on the retina with three kinds of “cones”, which are “photodetectors” Each kind has a different *spectral response*, by which we mean a curve that tells how much “output” the cone delivers from light of a fixed “potency” at each wavelength over the visible range.

---

[4] For almost all humans; there are a few humans, all women, whose perception of color requires four values to describe.

[5] The formal name of this is the power spectral distribution (PSD) of the light.

When a certain spot on the retina is bathed in light with a certain spectrum, in effect, for each of the three kinds of cones:

- The spectrum of the light is multiplied by the spectral response of the cone, meaning that, for each wavelength, the "potency" of the light at that wavelength is multiplied by the value of the spectral response at that wavelength.

- All these products are added together[6], giving the output of the cone.

The three types of cone are called "L", "M", and "S", referring to the fact that the peaks of their spectral responses are at different wavelengths, which we arbitrarily consider to be "long", "medium" and short. The spectral response curves of the three types of cone are called l, m, and s [7] (the usual typography is an overbar, but that is a pain to produce in this word processor, so I will underline them instead).

Figure 1 shows these three response curves (scaled so that their peaks are all at 100 units). (They are labeled here with the cone names, L, M, and S, not the curve names.)

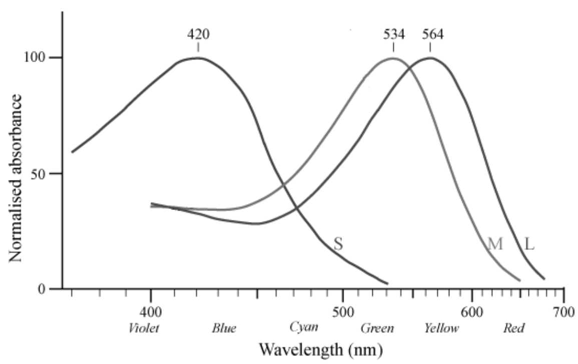

Figure 1. Eye cone response curves

GENERATING COLOR IN A DISPLAY

When we have created a digital representation of an image in our digital camera, stored that in a data file, and loaded that file into our

[6] Since both the spectrum of the light and the spectral response function of the cone are continuous, the process is actually *integration*, not *summation*, but the concept is identical.

[7] The mathematician would write then as $l(\lambda)$, $m(\lambda)$, and $s(\lambda)$, reminding us that all of them are functions of wavelength, λ.

computer to view the image, we look to our display system, at each pixel location of the image, to generate light of the appropriate color for the benefit of the viewer's eye.

We ordinarily do this by having, at each pixel location, a light-generating organ consisting of three parts, each capable of generating light of a certain chromaticity. By controlling the "potency" of these three light outputs at each pixel, we can make the overall light there have any color we want, within a certain range.

That means that the overall joint spectrum of the three emitters, under this situation, is a metamer of the color desired. Remember, any color has an infinite number of different spectrums that are metamers; we use the one we are be able to make with the three ingredients at hand.

We call the three "kinds" of light that can be emitted by these three types of emitters the *primaries* of the display system. Each has a different chromaticity.

Note then that a primary does not have a "color"; it only has a chromaticity. If it had a color, it would have a certain (fixed) luminance.

There is no "inherent" set of three primaries. A set for use in a display can be chosen under several different criteria. The minimum requirements are that they have different chromaticities, and, when plotted as points on a "chromaticity diagram" (a two-dimensional graphic presentation of chromaticity) do not fall in a straight line.

Regardless of the set of three we chose, by invoking appropriate amounts of each, we can construct light having any color at all so long as its chromaticity, on the chromaticity diagram, does not fall outside the triangle formed by the points indicating the chromaticity of the three primaries (and of course does not call for any of the ingredients to be included at a potency beyond the capability of its emitter).

In almost all display systems we encounter, including in color television and computer displays, the primaries have chromaticities that we call "red", "green", and "blue". These words do not describe chromaticities except in a broad way. But in any particular system, the primaries have certain specified chromaticities.

Color models and color spaces

Of great importance in this area is the matter of exactly how we quantitatively describe a particular color with a set of three numbers.

There are a number of general schemes by which a color can be quantitatively expressed (just as there are different schemes for stating the location of a point on the Earth). A particular scheme concept is called a *color model*. Most of the color models of interest to us fall in one of two genres:

- *Luminance-chromaticity* models. Here, one numerical value states the luminance of the color, and two more together state the chrominance. There are in fact several subdivisions of the genre, differing as to how the chromaticity is stated. And there is a "cousin" genre, the *luminance-chrominance* genre. (Sounds almost the same, doesn't it?)

- *Additive* models. These define a color by stating the amounts of three primaries (of defined chromaticity) that should be added together in a display device to produce that color. Additive models can rightly be called "tristimulus" models, since they work on the basis of defining three luminous stimuli[8] to be given to the eye. But by custom, the term "tristimulus" is reserved for a particular additive model, called the "CIE XYZ" model, used in scientific work. An important model of the additive type is the "RGB" model, in which the assumed primaries have chromaticities that we call "red", "green", and "blue". As before, these words do not describe chromaticities except in a broad way.

If we are going to state a color, perhaps as a description of a pixel of an image, with the expectation that the receiving system will be able to direct its display to reproduce that color, we have to get very specific about the details. Once we have done that, we have defined a specific *color space*. In the case of color spaces based on the RGB model, this includes:

- Precisely defining, in scientific notation, the chromaticity of the three primaries upon which the description of color will be predicated.

- Defining exactly how the "amounts" of each primary are to be stated: how much is "one unit" of a primary; are the numerical

---

[8] I normally eschew Latin plurals (such as "spectra") unless I am writing in Latin, but "stimuluses" is just more than I can expect the reader to bear.

values the actual amounts or some nonlinear representation of them, and, if so, following exactly what nonlinear function; and so forth.

One widely used example is the “sRGB” color space (“standard RGB”), widely used for the representation of color in digital image files.

The dual role of a set of primaries

We have seen above that a set of defined primaries has two separate, but related, purposes:

- As ingredients to compose any color of light over a certain range.
- As the underlying premise for describing color in an “additive” color space.

One convenient way these two functions can be unified is by establishing a preferred set of primaries for display devices, and then predicating our standard color space on that same set. One benefit of this is that the “coordinates” of each pixel’s color, as sent to the “viewing” end, can be directly fed into the “throttles” of the three primary emitters of the display at that pixel location.

In fact, in the early days of color television, exactly that was done, in part because there was no good reason to do otherwise.

And because of the relatively seamless progression from color television technology to computer graphic technology, we have that situation today. The primaries upon which the sRGB color space is predicated are the same that were considered desirable for computer displays, again allowing a verbatim application of the incoming color description coordinate values to the “throttles” of the display emitters.

Notation

In the case of each color space, we will be speaking of several sets of three things that all seem to have the same names, such as “R”, “G”, and “B”. I will use a consistent notation to identify them.

To identify the three primaries, I will use Roman letters: R, G, B (for example).

To identify the three values that tell us the amounts of the three primaries to be added to produce a certain color (the three *coordinates* of the color space), I will use italic letters (consistent with mathematical practice for variables, which these are): R, G, B.

Note that these are expressed in a linear fashion; that is, the amount of the primary to be included is proportional to the value of the coordinate. In many color spaces (such as sRGB), the coordinate values we actually store in a file are nonlinear transforms of these linear values (they are often said to be *gamma precompensated*). I will not need to speak of these in our work here. But I note that, to add to the confusion, in normal practice, we call those nonlinear coordinates R, G, and B.

Transformation of primaries

We have seen that we can generate a color using any of numerous sets of primaries (again subject to the fact that for any set of primaries, the colors produced must have chromaticities lying in a certain range).

Correspondingly, we recognize that, in the realm of additive color spaces, we can use different sets of primaries as our premise for defining any given color (again subject to the limitation just mentioned).

If we have a color defined in terms of an additive color space predicated on one set of primaries, but want to have it described in terms of a different set of primaries, we can do that with a straightforward (if not necessarily simple) mathematical procedure, involving matrix multiplication. This is called “transformation of primaries”. We’ll get a closer look at the mathematical concept later.

The CIE XYZ color space

I glancingly mentioned the CIE XYZ color space above. And it is frequently referred to in the literature of our area of interest. I thought it would be good to describe it.

This is a color space defined by the CIE. Those are the initials of its French name; its English name is *International Commission on Illumination*. It is a prominent body in the realm of standards for colorimetry.

The XYZ color space is an “additive” color space: it describes a color in terms of the amounts of three “primaries” that would need to be combined to create light of that color.

In this case, though, there is a wrinkle that seems extraordinary. The three primaries (called X, Y, and Z) are imaginary; we cannot actually generate any of them, and if we could, they would be invisible. On a chromaticity diagram, they are represented by points outside the region of visible light.

This notwithstanding, they follow the rules for determining the color resulting from the addition of certain amounts of a set of primaries. Thus, we can see on the chromaticity diagram that if we combine equal “amounts” of primaries X, Y, and Z, we will get light of a certain visible chromaticity (which turns out to be a certain kind of “white”).

The XYZ color space is attractive for scientific work because it has several convenient mathematical properties. One especially fascinating one is that the luminance of any combination of amounts of the three primaries (that is, for any set of the coordinates *X*, *Y*, and *Z*), the luminance of the light indicated is exactly *Y*, the amount of primary Y in the “recipe”. This is in fact why the symbol “Y” is used for luminance in much colorimetric work.

One way to look at this is that primaries X and Z are “impotent” insofar as contributing to luminance. If this is all beginning to sound too fantastic, remember that this entire XYZ story is fictional, and these primaries, like superheroes in comic books, can have any properties the wonks wanted to give them, so long as the math works out properly in connecting the story to the real world.

Because the XYZ color space is an “additive” one, it can properly be said to be a “tristimulus” color space. In fact, by custom, when the term “tristimulus color space” is used without further elaboration, it is taken to mean the XYZ color space (it is the “mother of all tristimulus color spaces”).

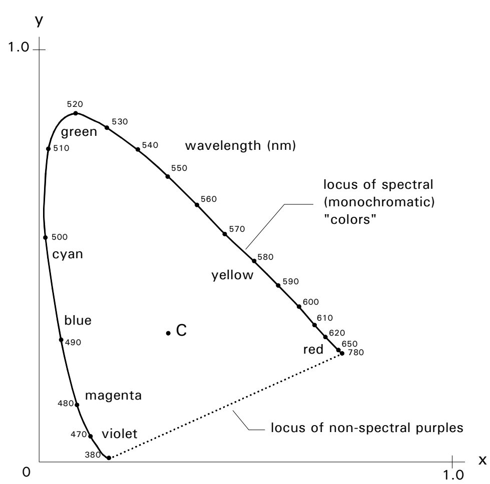

Figure 2. The CIE x-y chromaticity diagram

From the XYZ color space flows another one, called xyY. This is not an "additive" color space, but rather a "luminance-chromaticity" one. here *Y* represents luminance and *x* and *y* together define chromaticity. In fact, a chart on which we plot the *x* and *y* values of colors is said to be a "CIE x-y chromaticity diagram". Figure 2 shows the CIE x-y chromaticity diagram.

Any chromaticity corresponds to a point on this diagram. The "horseshoe" curve is the set of all the points for the chromaticity of light whose spectrum consists of only a single wavelength. These are called "spectral" chromaticities.[9]

The dotted line at the bottom completes the region "enclosed" by the horseshoe. The chromaticities along it are not "spectral" (there is no light with only a single wavelength component that exhibits a color with such a chromaticity). They are called the *nonspectral purples*.

All colors of visible light have chromaticities represented by points inside the region bounded by the horseshoe (and the locus of nonspectral purples) are visible. Since chromaticity is an aspect of color, and color is defined in terms of human perception, radiation that is not visible does not have a color, nor a chromaticity. Thus, strictly speaking, points outside the horseshoe do not represent chromaticities.

Nevertheless, we can have "imaginary" primaries, and we suspend rigor to consider them to have a chromaticity, which is then represented by a point on the diagram. We can safely do this because these chromaticities have all the mathematical properties of real chromaticities, including reckoning what happens when we add together two or three of them to "give" a visible color.

This having been said, and now that we understand the chromaticity diagram, circling back to an earlier point, figure 3 shows, on this diagram, the chromaticity of the three imaginary primaries of the CIE XYZ color space, X, Y, and Z.

---

[9] This comes from a different use of the word "spectrum" than we are using here. It refers to the "spectrum" of daylight generated by a prism, which spatially separates the components of the light by wavelength. At any point along this spectrum, the light contains only a single wavelength. Thus we give the name "spectral" for light whose spectrum (in the sense we have been using the term) contains only a single wavelength. They are sometimes called "monochromatic" chromaticities (meaning "single color"). This is a different meaning of *color* than we have been using.

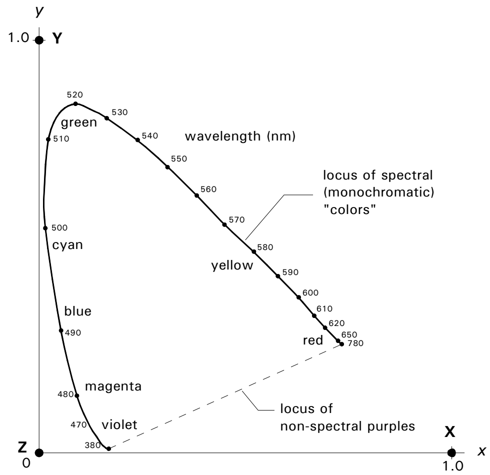

Figure 3. The CIE primaries X, Y, and Z

THE DIGITAL CAMERA SENSOR

The task of the sensor

Recall that we are ignoring the fact that the sensor array of our camera is likely a CFA type, and is not organized by pixel, Rather, we will imagine a sensor array that has a "color-determining" organ at each pixel location.

We look to that organ to ascertain the color of the light landing at its location and report that in terms of three coordinates in accordance with some color space.

Make an eye?

How might we construct such a color-determining organ? One fairly obvious way would be to make it of three photodetectors, each provided with a filter whose spectral response is l, m, or s—exactly emulating the behavior of the eye.

The outputs of these three "sensor channels", which we will call *L, M,* and *S,* can describe any color. (We are tempted to say, "any color the eye can see", but in fact light that cannot be seen, owing to lack of any light within the range of visible wavelengths, has no color, so we need not bother with the qualifying words.)

Maybe not

But for various reasons it is not attractive to do this. For one thing, the overall height of those response curves would make the sensor, overall, unattractively "insensitive" to meet some of our photographic objectives. And the fact that the l and m functions have their peaks very close together leads to the need for coordinate transformations with large coefficients, which can exacerbate noise in the overall image process.

So instead, we equip the three kinds of photodetectors with three other spectral response curves. Can such a sensor actually consistently discern the color of the light falling on pixel location?

The colorimetric researchers (von) Luther and Ives[10] showed mathematically that the outputs of a set of photodetectors will consistently describe the color of the light, regardless of the specific spectrum involved, if (and only if):

> The response curves of the three types of photodetectors are linear combinations of the eye cone response functions $\underline{\mathrm{l}}$, $\underline{\mathrm{m}}$, and $\underline{\mathrm{s}}$ (there being a couple of other requirements, some pesky stuff about orthogonality and so forth).

This includes the possibility that the response curves **are** just *l*, *m*, and *s*, but opens the door to the use of other sets of response curves.

These requirements are called the Luther-Ives conditions.

Suppose we exploit the Luther-Ives principle in our sensor design, adopting three responses that are each linear combinations of the response curves $\underline{l}$, $\underline{m}$, and $\underline{s}$. If we do that, our sensor will operate as an instrument of a color space (although not a “recognized” one, such as sRGB—we’ll se later that it can’t be).

We could then transform the sensor outputs into (linear) coordinates of some standard color space (for example, sRGB) by using a straightforward mathematical transform.

We will use *R*, *G*, and *B* to represent the (linear) coordinates in the sRGB color space, and *D*, *E*, and *F* to represent the three outputs of the sensor. [11]

---

[10] “Ives” is Herbert E. Ives, of Bell Telephone Laboratories, who conducted an early demonstration of television in 1927. Much of his work was in the theory of color imaging.

[11] We generally label the three “channels” of a sensor R, G, and B, but as we’ll see later, this is misleading, so I purposely avoid it here.

The transform comprises these three equations:

$$R = m_{11}D + m_{12}E + m_{13}F \quad (1)$$

$$G = m_{21}D + m_{22}E + m_{23}F \quad (2)$$

$$B = m_{31}D + m_{32}D + m_{33}D \quad (3)$$

where the nine "m" coefficients (they are constants) define the transform.

But this can be written in matrix notation this way:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} D \\ E \\ F \end{bmatrix} \quad (4)$$

or, more compactly, with [M] as a symbol for the entire matrix, implying exactly what is shown in detail in equation 4:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = [M] \begin{bmatrix} D \\ E \\ F \end{bmatrix} \quad (5)$$

In a later section we will get some further insight into why it works this way.

Maybe we won't even follow the Luther-Ives conditions

In fact, a design following the Luther-Ives conditions may not be attractive either. Thus we may well compromise even further and use a design that does not follow those conditions.

If the Luther-Ives conditions are not met, then colors having metameric spectrums (that is, different spectrums but nevertheless having the same color) will in general give different sets of outputs from the three sensor channels. In other words, such a sensor will not give a reliable indication of the color of the light. It would be said to have *metameric error*. As a consequence, the behavior of such a sensor cannot be said to be consistent with any color space (not even a "private" one only applying to that sensor design).

Fortunately, if we concentrate on light having the kinds of spectrums we most often encounter in photography, we can minimize the "average" amount of metameric error we encounter with such a sensor. If we choose "properly" a certain arbitrary transformation matrix [M], the average metameric error for a collection of “representative” light spectrums will be held to a minimum. And this is in fact what is done with many camera sensors today.

SENSOR CHANNEL SENSITIVITY

We often hear that the three “channels” of a sensor have different sensitivities, with the G channel sensitivity usually being highest and the R channel sensitivity being lowest. (We often hear, for example, about the need to “scale” the sensor outputs as a first stage of processing the data to compensate for this difference in sensitivity.)

But this concept is too simple. Any one of the three channels has many “sensitivity” values, one for each wavelength. This is a result of their individual response curves (without which they could not do their work). At some wavelengths, the R channel has a greater sensitivity than the G channel.

So when we speak of the sensitivity of a channel in a simple way, we must mean its average sensitivity to a test light source of some particular spectrum (the same source for all three channels).

One light source that is useful for this is “illuminant E”, whose spectrum is “flat” (equal power in any increment of wavelength).

Other times, what is used is one of the standard “daylight” spectrums, such as that of “illuminant D50”.

Now, that having been said, the realities are such that we probably do want to multiply the outputs of the R and B channels by certain constants in order to get them ”in the same ballpark” for typical scene light spectrums. But these constants are pragmatic, not the result of there being such a thing as an actual single “sensitivity” value for each channel.

The transformation matrix used will be affected by the choice of scaling constants. Another way to look at it is that the scaling constants are “built into” the matrix, rather than being applied by a separate multiplication step before the matrix transform is applied.

SENSOR RESPONSE DECOMPOSITION

The colorimetric behavior of a sensor is often described in terms of “the decomposition of its response”. To explain this, we will first need to do a little preparation.

A Luther-Ives sRGB sensor

Imagine that we had a sensor that:

- Met the Luther-Ives conditions, so that any two instances of light having the same color, even with different spectrums, would be reported with the same output values.
- Had outputs would be the (linear) coordinates of the sRGB color space.

To meet the Luther-Ives conditions, the response curves of all three channels of this sensor would have to be (different) linear combinations of the eye cone response curves, *l*, *m*, and *s*. We can in fact determine the "coefficients" of that combination (the constants by which one, two, of three of the cone response curves were to be multiplied before they were added to give each of our response curves) with a straightforward mathematical process.

So, why don't we just build such a sensor? It would be so handy. There would be no need to transform its outputs into linear sRGB coordinates—they already are.

Aye, but here's the rub. Such a sensor's response curves would have to have negative values in parts of the wavelength range, which is of course impossible to implement in a real sensor (it is not even physically meaningful). Figure 4 shows these response curves (again, labeled with the names of the primaries, not the names of the curves).

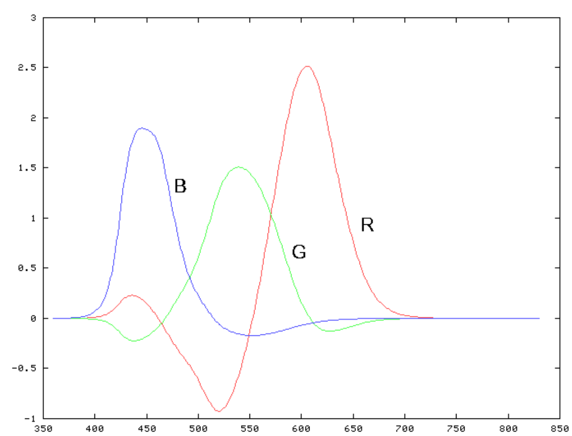

Figure 4. sRGB response curves

So we must consider another sensor design for our real camera.

Another Luther-Ives sensor

Now consider another sensor of interest. It also meets the Luther-Ives conditions (so it would be free of metameric error), and its response curves are nowhere negative, so they can actually be implemented. But the implied primaries cannot be the sRGB primaries (or any other real primaries), a consequence of the “nowhere negative” response curves.

However, we can transform its outputs into linear sRGB coordinates using a matrix transform such as we saw in equations 1-5.

Transformation of coordinates

Since the response curves of the “sRGB” sensor are linear combinations of the eye cone response curves, and the response curves of the “actual” sensor are (different) linear combinations of the eye cone response curves, then the response curves of the “actual” sensor are linear combinations of the sRGB sensor response curves.

From that we can show that each of the outputs of the sensor is a linear combination of the three (linear) sRGB coordinates for the color being observed. (I will spare you the proof.)

Thus, in effect:

$$D = \mathrm{k}_{11}R + \mathrm{k}_{12}G + \mathrm{k}_{13}B \qquad (6)$$

$$E = \mathrm{k}_{21}R + \mathrm{k}_{22}G + \mathrm{k}_{23}B \qquad (7)$$

$$F = \mathrm{k}_{31}R + \mathrm{k}_{32}G + \mathrm{k}_{33}B \qquad (8)$$

where *D*, *E*, and *F* are the outputs of the sensor, and *R*, *G*, and *B* are the (linear) sRGB coordinates of the same color.

If we take the “k” coefficients and write them as a matrix, [K], then the inverse of that matrix ($[K]^{-1}$) will be matrix [M], the one we need to transform our sensor outputs into sRGB (linear) coordinates (as seen in equation 5).

Also note, for future reference, that if we know [M] for a sensor, then its inverse ([M]-1) will be the matrix [K], consisting of the “k” values.

Now suppose that, to test the *R* channel of our real sensor, we expose it to sRGB “R” primary light. There are many light spectrums that would be that color, any of which could be used for the test.

For that color, in the sRGB coordinate system, R would have a nonzero value, but G and B would be zero.

Thus, equation 6 would degenerate into this:

$$D = \mathrm{k}_{11}R \qquad (9)$$

So by observing the output of our real sensor when exposed to light whole color is that of the sRGB R primary, we can discern the coefficient k_{11}. Eight other similar exercises would reveal to us the other "k" values for our sensor.

Based on this, when sensor characteristics are being discussed, presentations of, for example, k_{11}, k_{12}, and k_{13} often describe them as "the response of the sensor's D channel[12] to the three sRGB primaries, R, G, and B." And we have just seen that this description is justified.

Another way to look at this is that the output of the D channel would ideally be *R* ("ideally" based on a preoccupation with an sRGB final result), but it is corrupted by the presence of the terms $k_{12}G$ and $k_{13}B$. Thus, *D* is said to be "impure" (and language based on that is often found in the "colloquial" aspect of descriptions of sensor performance).

Consistent with this, the transform matrix ([M]) can be thought of as starting, for example, with *D* (the "impure *R* output" of the sensor) and "backing out" of it the "corrupting" terms, $k_{12}G$ and $k_{13}B$.

Keep in mind that this specific conceit of "purity" of a sensor channel response is based on the predicate that the sRGB color space is how we really want to specify color.

If we always set our camera to deliver its outputs in the Adobe RGB color space (which has a different G primary), then this definition of "purity" is entirely fanciful.

What if we don't have no stinkin' Luther-Ives conditions?

I spoke above of a Luther-Ives sensor that was "buildable", unlike an sRGB Luther-Ives sensor. But I also said earlier that for various reasons, sensors in actual use rarely follow the Luther-Ives conditions.

---

[12] Normally spoken of as the "R" channel.

Just for interest, figure 5 shows the response curves of the three kinds of detectors in the sensor of a Canon 500D digital SLR camera (in the U.S., the EOS Rebel T1i), as measured by DxO Labs.

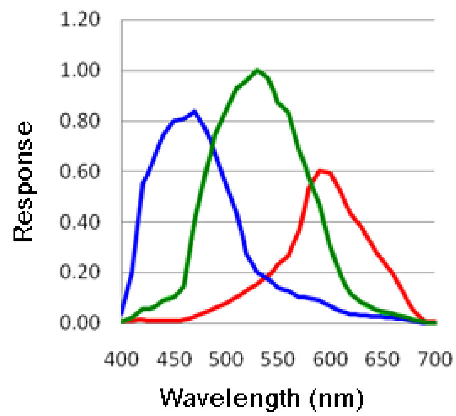

Figure 5. Sensor response, Canon EOS 500D

Such sensors do not really operate in any color space (even a "private" one). That is because they will not necessarily deliver a consistent set of outputs for light of the same color but with different spectrums (metamers).

How does the approach discussed above for characterizing the response of a sensor apply to that situation? Well, it really doesn't. But we pretend that it does.

As discussed above, for such sensors, we adopt an arbitrary transformation matrix [M] to convert the sensor outputs to sRBG linear coordinates. If we were working with a Luther-Ives sensor, the inverse of that matrix, $[M]^{-1}$, would be [K], the matrix of the coefficients of the relationship between a color (as described by its sRGB linear coordinates) and the sensor outputs (thought of as the coordinates of the color in the sensor's own private color space). And in that case, those coefficients can be described (quite properly) as "the response of the sensor D, E, and F channels to light whose color is that of the sRGB primaries R, G, and B."

But that's not actually true here. Remember that [M] was crafted based on "minimization of average metameric error". It didn't flow directly from any basic properties of the sensor response.

But we act as if the description is true. We take the inverse of our adopted matrix [M], consider it to be [K], the matrix of the coefficients "k", and state those coefficients, saying that they are "the response of the sensor D, E, and F channels to the sRGB primaries R, G, and B."

They aren't (even for primaries having the "standard" spectrums that are defined for each). But "we" say they are.

PRIMARIES WE DON'T MENTION

I suggested that any set of linear combinations of the eye cone sensitivity response curves can define a color space of the "additive" type. And any additive color space is defined by a set of primaries.

Now the first additive color space we encountered here is that inside the eye, defined directly by the three eye cone response functions (the simplest linear combination of them, the mother of all Luther-Ives compliant systems). Why do we not ever hear, even as a matter of curiosity, about the primaries implied by this color space?

Well, it's because they are imaginary (just like the primaries of the CIE XYZ color space).

In fact, for any color space defined by three response curves that are not at any wavelength negative, the primaries must be imaginary. (I'll spare you the proof.)

Another class of color space is that implied by a "buildable Luther-Ives" sensor, whose three channels have response curves that are linear combinations of the three eye cone response curves (but are nowhere negative). Why do we not ever hear of the set of three primaries that would be implied by some illustrative Luther-Ives sensor?

Yep—it's because they are imaginary.

Now what about our "real" sensors, not conforming to the Luther-Ives conditions? Well, because of that "shortcoming", they don't really report in a true color space. But, by adopting a matrix to transform the sensor outputs into the (linear) coordinates of some real color space, we pretend that the sensor operates in a true color space. And why do we not ever hear about its primaries?

Because they would be imaginary, a consequence of the sensor's response curves being "nowhere negative", the converse of the situation stated above in bold.

SENSITIVITY METAMERISM INDEX

The concept at issue

Let's review the concept of the "metameric accuracy" of a sensor system.

The spectrum of an instance of light determines its color: any given spectrum implies a certain color.

(We must resist the temptation to say, “appears to the human eye to be a certain color”. That is true, but in fact, that color **is**, by definition, the color of the light, so we don’t need the more complicated expression.)

Recall, however, there can be many different spectrums of light that will nevertheless have the same color, a situation called *metamerism*. Different spectrums that have the same color are called *metamers* (or sometimes, “metamers of that color”).

We would like our digital camera sensor to “report” the color of the light it detects, consistently, regardless of the specific spectrum (metamer) involved. That is, for any spectrum having a certain color, the outputs of the sensor would be the same. (These three output values will probably represent the color in an unfamiliar color space, but we can transform them to a representation in a standard color space, such as sRGB.)

A sensor that meets this expectation can be said to be “metamerically accurate”. (And we saw above that, for a sensor to do so, it must meet the Luther-Ives conditions regarding the spectral sensitivity curves of its three “channels”.)

But in fact typical sensors do not meet that expectation. In that case, if we subject the sensor first to light of one spectrum having a certain color, and then to light of a different spectrum having the same color (two metamers of the same color), the sensor may well deliver different sets of three outputs—may report two different colors for the two instances of light, whose actual colors are identical.

This is of course not desirable as part of our overall quest for “accuracy” of imaging.

We cannot “correct” for this “flaw in measurement” by any form of mathematical processing of the sensor output data. That is, we cannot massage the data so that, after a transformation of coordinates, it will constantly describe the actual color of the light in our chosen output color space.

The discrepancy between the actual color of the light and the indicated color (after applying to the sensor outputs whatever transformation matrix we have adopted) is called the *metameric error* of the sensor for that particular light spectrum.

We can mitigate the impact of this (as was described above) by choosing a transform between the outputs of the sensor and what we treat as the (linear) coordinates of the “color” under a certain color space (for example, sRGB) such that for some arbitrary collection of representative light spectrums, the overall average metameric error will be as small as possible. In fact, ISO standard 17321 gives a procedure for constructing this “optimal” transformation matrix based on measurements of sensor response for a certain such collection of light spectrums, predicated on a certain way of “scoring” the metameric error in each case.[13]

The definition in the standard is actually for constructing the “optimum transformation” matrix for transforming the sensor outputs into coordinates of the CIE XYZ color space. However, we can adapt that matrix, by a straightforward mathematical procedure, into one for use to transform the sensor outputs to any specific color space we wish (such as sRGB).

“Scoring” the sensor metameric error

After a camera manufacturer has adopted a transformation matrix (whether by the strategy suggested by ISO 17321 or some other), we will still have metameric error. We will probably have some error even for light having any of the spectrums used as “models” in the test, and we will in general have (likely greater) metameric error for other spectrums.

In light of this, metrics for “rating” the degree of “residual” metameric error have been devised, and ISO 17321 presents one such method. The result is called the *digital still camera sensitivity metameric index*, or DSC/SMI.

As with the development of the “optimum” matrix, it involves comparing:

- The “indicated” color of each of a set of light spectrums, after transforming them with the “optimum” matrix to the CIE XYZ color space, to
- The “known” color the spectrum should indicate (as described in the CIE XYZ color space).

The differences (metameric errors) are numerically assessed using a metric of color difference called “ΔE” (the definition of this metric is

[13] Not all experts would agree that this is the “best” way to “score” metameric error.

beyond the scope of this article). The average of the errors, over the collection of test spectra, is used as the basis of the DSC/SMI. It is 100 for no error, and decreases with increasing average error.

This "rating" is actually called the "average DSC/SMI". We can also state a "special DSC/SMI", based on the metameric error for some single particular light spectrum of interest to us.

WHITE BALANCE COLOR CORRECTION

The spectrum of the light coming from any reflective object is a joint result of the spectrum of the incident illumination and the "reflective spectrum" of the object (the "plot" of the fraction of the incident light it reflects, as a function of wavelength).

Thus, the spectrum of the light from a favorite Fiestaware dinner plate, illuminated by the incandescent lighting in our breakfast room, mat be quite different than the spectrum of the light from the plate when illuminated by shaded sun on our patio table.

Clearly, based on the concept introduced at the beginning of this article, those two instances of light have different color.

Yet we find that the human observer sees the plate as "the same color" in both settings. How can this be?

This is the perceptual phenomenon of *color constancy*, and it is explained by a perceptual phenomenon known as *chromatic adaptation*.[14]

In effect, the human eye "determines" the chromaticity of the incident illumination in the setting where the object is being observed, by processing the color of the light reflected by many familiar objects. Then, the eye "discounts" the effect of the chromaticity of the incident light on the spectrum of the object of interest before concluding "what is its reflective color".

When the eye has come to a conclusion about the chromaticity of then incident illumination in a particular setting, it is said to be "chromatically adapted" to that illumination.

---

[14] This is one of (at least) three meanings of the term *chromatic adaptation* in color science, so we need to be quite careful.

Now suppose that we take a photograph of our object. The color recorded in the image will represent the color of the light reflected from the object (not its reflective color).

When we show the image to a viewer, in general the viewer’s eye has not been chromatically adapted to the illumination on the object (which was in a different place, at a different time). Rather, it will be chromatically adapted to the illumination at the site, and time, of viewing.

Thus the eye’s “chromatic adaptation” mechanism incorrectly “discounts” the effect of the incident illumination on the object, and will not necessarily perceive the object with its familiar color (a situation often referred to as a “color cast” in the image). (“I don’t remember that plate being so reddish.”)

To overcome this, in photography, we transform the colors in the image, before we “deliver” it, so that they will be properly perceived by the viewer.

But how can we do that if we do not know the chromaticity of the ambient light at the viewing location. Well, we can’t.

But what we do is to transform the colors so that they would be properly perceived if viewed in an environment with a certain presumed illumination. And in fact, in such a standard color space as sRGB, the colors are supposed to be represented transformed so that they would be perceived properly if the image were viewed under an illuminant matching, in chromaticity, a “CIE standard illuminant” called “D50”. This transform process, often done in the camera, is called “white balance color correction”.

D50 is a so-called “daylight” illuminant, corresponding to daylight at dusk or dawn. Thus, it would seem, the viewer will properly perceive an ideal image, carried in the sRGB color space, by viewing it in “dusk/dawn” daylight (that is, with his eye chromatically adapted to such light).

Is that reasonable? Aren’t most images viewed indoors?

Yes, but there is a wide range of spectrums of the incident light “indoors”. And because of that (or perhaps rationalized by that), and for complicated historical and pragmatic reasons, the assumed viewing illumination of the sRGB color space is D50.

Now, suppose that the viewer is a sophisticated image editor, and it is vital that he be able to perceive the colors on the image precisely as intended. One way to do that is to light the viewing room with light whose chromaticity is that of D50.

Now to be able to do white balance color correction, we need to have a "mathematical model" of the way the eye changes its response when it has been "chromatically adapted". It turns out that, essentially, this works by the eye changing the sensitivity of its three kinds of cones. In effect, for any particular value of what the eye concludes is the chromaticity of the ambient light, the "basic" response curves of the three types of cones are just scaled by a set of three constants.[15]

Now, how can we outguess this at the camera as we do white balance color correction? Assume we know the chromaticity of the incident light used for the photograph (how we do that is another big story, beyond the scope of this article).

In theory, if we had the image colors described in terms of the three coordinates used in the eye (the outputs of the three types of cone), we could then:

- multiply those coordinates by three constants that would represent the adaptation the eye would take on if aware of the chromaticity of the incident illumination, then
- divide the resulting coordinates by three constants that would represent the adaptation the eye would take on in adapting to the illumination presumed by our output color space (e.g., sRGB).

(The two sets of three constants could of course be consolidated into a "color correction vector", so we could do the whole process in one step.)

Finally, we would transform that new "eye-based" description of the colors into a description under sRGB for delivery.

But we don't ordinarily transform the colors as described by the three sensor outputs into an "eye-based" coordinate system on their way to an sRGB representation. So we can't conveniently follow that ideal scenario.

Could we do the whole thing by using some matrix on the sensor outputs, which would, in effect, convert the color representation to

---

[15] This can be called the "chromatic adaption vector", *vector* in this sense meaning a set of several values (a one-dimensional matrix).

“eye-based”, apply the consolidated color correction vector, and transform it back to “sensor output” form? Yes, but this would be clumsy, and would introduce some practical problems.

Could we just multiply the three sensor outputs by the set of three values of the consolidated color correction vector? Not properly. That simple manipulation would only work “exactly” if we had the colors at that point in “eye-based” coordinates, not in the color space (or “quasi-color space”) implied by the sensor outputs.

But keep in mind that using the sensor outputs as an indication of color is already a compromise (because of metameric error).

So, pragmatically, we typically do what was suggested just above. The three sensor outputs are multiplied by a consolidated color correction vector—a set of three constants—dependent on the known or assumed chromaticity of the incident light on the subjects.

In fact, in many cases, this step is consolidated with the “rescaling” of the sensor outputs to compensate for their different “sensitivities”.

If we look at the proprietary part of the Exif metadata for an image output file for a Canon dSLR camera, we will find a table of these consolidated color correction vectors, one for each of the “preset” color corrections based on “familiar” types of illumination (“full daylight”, “shaded daylight”, “incandescent light”, etc.), as well as one for a particular kind of illumination whose chromaticity had earlier been measured using the camera (the “custom white balance” vector). These have the “sensitivity compensation” factors built in.

Thus, for an illumination very nearly D50 (which we can say “requires no color correction”), in my EOS 40D, that vector is (using the D, E, F channel labeling):

D: 2131, E:1024, F:1498

The ratios there, relative to the “E” channel, 2.08:1,00:1.46, essentially reflect the inverse of the relative “sensitivities” of the D, E, and F channels, in the relevant sense.

IN A CFA CONTEXT

At the outset, I noted that in many cameras of interest, we did not actually gave three photodetectors (with different spectral response curves) at each pixel site, but rather photodetectors with three different response curves deployed in a repetitive pattern across the different pixel sites, a so-called CFA (color filter array) sensor.

We then derive from the collection of outputs of these “single response” photodetectors a “full-color” image, with a representation (perhaps in the sRGB color space) of a color for each pixel. Note that I said “a color”, not “the color”, since we realize that this can only be a “crafty estimate” of the color of the light at each pixel location. The process by which this happens is called *demosaicing*.

The question has been raised as to whether the characterization of the colorimetric behavior of a sensor by way of the response curves of its three types of photodetectors is entirely meaningful in this situation. After all, no photodetector discerns the color of light at its location. nor do three photodetectors collaborate to discern the color of the light at any pixel.

Thus perhaps, it has been suggested, we should characterize the sensor itself as reflected in the outputs of some particular demosaicing algorithm. (Ah, but which one?)

Indeed, we **are** ultimately interested in the colorimetric behavior of the entire imaging chain, including a particular sensor and a particular demosaicing algorithm, and we do often measure and characterize that.

However, in order to craft a demosaicing algorithm, we need to know (from a colorimetric standpoint) the “low-level” behavior of the sensor on whose outputs it will operate.

And we can most directly characterize that in terms of the behavior of the sensor’s “low-level” elements, the photodetectors, as I describe in this article.

ACKNOWLEDGEMENT

I would like to recognize the authors of the numerous technical papers and doctoral dissertations from which I gleaned much of my knowledge and understanding of this complex and confusing field.

Special appreciation goes to one doctoral candidate, who recognized in his acknowledgements the contributions of his wife, not only in assisting him in his work, and being supportive of it, but, earlier in their relationship, having been responsible for his moving to Paris, where he ended up taking his doctoral studies and doing the important research he reported. *Cherchez la femme, et peut-être trouvez aussi bien la connaissance*.

Special thanks to my wife, Carla, for her patient, insightful, and speedy copy editing of this difficult manuscript, and for creating an environment for my life in which this kind of study and exposition can flourish. *Vraiment, j'ai trouvé **la** femme!*

#