Metameric Error in Digital Photography

Douglas A. Kerr

Issue 1
March 22, 2010

ABSTRACT AND INTRODUCTION

The color of light is defined wholly in terms of visual perception. The color of an instance of light perceived by the human visual system—by definition, that is the color of the light—is determined by the spectrum of the light. Any particular spectrum will "have" a specific color. But there can be many spectrums that will have the same color, a situation called *metamerism*.

Ideally, a digital camera imaging system would honor this, "recording" the same color for light of any spectrum having the same color. But various compromises in sensor design make our cameras fall short of this; they exhibit some degree of *metameric error*.

In this article we investigate human perception of color, the nature of metamerism, the operation of digital camera sensors, and why metameric error exists. We also discuss ways in which the impact of metameric error can be mitigated, and the way in which the residual metameric performance of a digital camera sensor can be "scored".

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 transformation of sensor outputs into the colors of pixels described in some standard color space (such as sRGB) occurs in the course of demosaicing. This makes more complicated the impact of the matter discussed in this article on overall camera performance, and it may complicate our vision of the matter, but it does not actually disrupt the concept. For that reason, it is perhaps best if the reader thinks for a while in terms of a sensor in which there is a photodetector of each type at each pixel location.[1]

Color

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

If two instances of light appear to a viewer to be the same color[2], they <u>are</u> 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 numerical values. 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 as to their details, are called *color spaces*.

---

[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". The situation this respects will not come up in this article.

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

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”).[4]

The “shape” of the plot determines the *chromaticity* of the light; its overall “vertical scale” determines its *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”.[5] 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.

When an area 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

---

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

[5] There are a few humans, all women, who have four kinds of cones. Accordingly, their perception of color requires four values to describe.

"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 $\underline{l}$, $\underline{m}$, and $\underline{s}$[7] (the usual typography is an overbar, but that is a pain to produce in this word processor, so I will underline them instead).

We cannot actually determine these curves (formally described as response *functions*, since they are functions of wavelength). We can derive related curves by tedious visual experimentation, and from these indirectly come to conclusions as to the curves themselves. But different researchers have come to different conclusions.

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

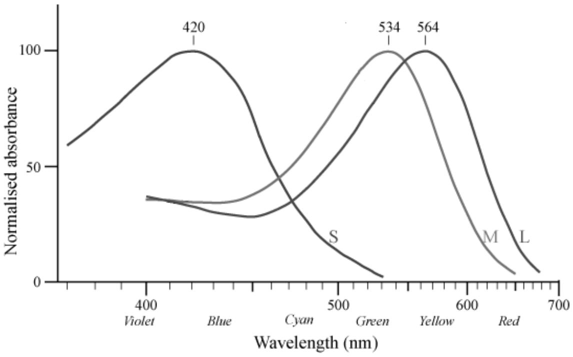

Figure 1. Eye cone response curves

---

[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 $\underline{l}(\lambda)$, $\underline{m}(\lambda)$, and $\underline{s}(\lambda)$, reminding us that all of them are functions of wavelength, λ.

THE DIGITAL CAMERA SENSOR

The task of the sensor

As I urged at the outset, we are ignoring the fact that the sensor array of our camera is likely a CFA type, with photodetectors of three different types arrayed across the image in some recurrent pattern. Rather, we will imagine a sensor array that has a complete "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 this 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. (They would be in the same place—the center of a pixel site.)

The outputs of these three "sensor channels", which we will call *D, E,* and *F,*[8] 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 certain needs in the "processing" of the sensor outputs, 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?

---

[8] It is customary to call these *R*, *G*, and *B*, suggesting that they are the coordinates of some (perhaps specialized) RGB-family color space. In fact, this is a little misleading, and so I avoid that notation here.

The colorimetric researchers (von) Luther and Ives[9] 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 $\underline{l}$, $\underline{m}$, and $\underline{s}$ (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 $\underline{l}$, $\underline{m}$, and $\underline{s}$, 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 $\underline{s}$ (we can call those $\underline{d}$, $\underline{e}$, and $\underline{f}$. If so, then the sensor outputs D, E, and F will have consistent values for light of any given color, regardless of which of the possible metameric spectrums the light has.

We could then transform the sensor outputs into (linear) coordinates of some standard color space (for example, sRGB) by using a straightforward mathematical transform. It can be explicitly determined from the functions $\underline{d}$, $\underline{e}$, and $\underline{f}$ that we chose for the sensor.

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, having 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 is said to have *metameric error*.

---

[9] "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.

Metameric error cannot be “corrected” by any mathematical processing of the sensor outputs.

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 “cleverly” a transformation matrix to be used for conversion of the sensor outputs into the coordinates of our “delivery” color space, 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.

IMPLICATIONS OF METAMERIC ERROR

There are several implications of metameric error in digital camera systems, some of which are very subtle. I will describe one of the clearer, and more potent, impacts.

Early on, designers and manufacturers of consumer goods early keen students of the realties of human color perception. Two objects, whose paint had different spectrums, might (in a side-by-side comparison), under one illuminant appear, to the viewer as having the same color, but under a different illuminant have visibly different colors.[10]

That is not just because the two items have different reflective spectrums, and thus, under the same illuminant, the light from them has different spectrums. It is because those two (different) spectrums are not metamers—they do not have the same color. They are not only different but “different”.

If we want the object to be “neutral” (that is, appear to be white or gray), there is a trivial solution: make the paint have a uniform reflectance spectrum across the visible wavelength band. But that’s not easy to do, and in any case we don’t want all our products to look white (although at one time, in the refrigerator business, that would have worked).

So paint scientists developed a very sophisticated discipline of planning the reflectance spectrums of paints so that two products, intended to “look the same color”, but painted with different kinds of paint (perhaps one high-durability, one not) would appear the same

---

[10] This is known here as *the haberdashers’ lament*. A jacket and trousers might look the same color under the light of the store, but not so in the buyer’s living room.

color **under commonly-encountered illuminants.** ("Under all illuminants" would be better, but is much harder to achieve.)

As plastics came into prominence in product design, that discipline was enlarged to lead to practical ways to attain this consistency of appearance between painted and plastic items or components. We might see this at work in a hand kitchen mixer, with a painted metal body and a integral-color plastic end cap.[11]

Now, suppose we take a picture of this mixer with our digital camera under one of the "cooperating" illuminants. Its two parts look to us to be the same color as we examine the subject in place for the shot. However, as a result of metameric error, in the image the two portions of the mixer will be recorded with different colors, and will be clearly seen as different when the image is viewed or printed.

Note that this is not a matter that can be cured by "white balance color correction". The problem here is not that the color of the mixer does not appear "correct" to someone who views the image with his eyes "chromatically adapted" to a different illuminant than that at the site of the photography. It is that the two parts of the object do not, in the viewed image, look the same color.

SENSITIVITY METAMERISM INDEX

As we have just seen, it is not attractive that our digital camera sensors (and thus the cameras overall) do not consistently respond to instances of light having a certain color but different spectrums. It would be nice to have a "metric" for the overall degree of metameric error committed by a sensor, just as we have metrics for other properties of camera performance, such as resolution, dynamic range, noise performance, and so forth. This "rating" could be one factor in choosing one camera over another.

ISO standard 17321 presents one such metric, called the *digital still camera sensitivity metameric index*, or DSC/SMI, and gives a protocol for determining it.

The protocol is **conceptually** as follows:

a. The sensor is made to observe the light from a number of standard "test patches", each having a known "reflectance spectrum", all illuminated by a standard illuminant, of known spectrum, so that

---

[11] I know that construction sounds rather nostalgic today—perhaps we should visualize the color as "avocado".

we have a known spectrum of light from each of the patches. For each patch, we determine (mathematically), from that spectrum, the “true” color of the light reflected from it (described in some color space, most often the “CIE XYZ” color space, used to represent color in much scientific work).

b. By “mathematical” simulation, based on the response functions of the sensor (which we have presumably measured earlier), we determine what the sensor outputs would be for the light from each patch.

c. Using a “trial transformation matrix”, we convert the calculated sensor outputs for the light from each patch to a representation of the color in the XYZ color space. If the sensor is free from metameric error, and if the “trial” matrix is ideal for the sensor behavior, those results will be the same as the XYZ representations of the light from the sample (as developed in step a).

d. We reckon and note the errors between the “recorded” color and the “true” color for each patch (metameric errors), using a metric of color difference called “ΔE” (the definition of this metric is beyond the scope of this article).

e. Mathematically, we determine an alternate (“optimum”) transformation matrix (to the XYZ color space) that would, if put into effect rather than the “trial” one, result in a collection of color errors whose average would be the least achievable. The thought is that this matrix would actually be used in a camera with the sensor type we had been testing.

f. Additionally, that least average of the errors (with the “optimum” matrix in use) is used as the basis of the “metric” of metameric error.

In reality, we do not do steps c-e. Rather, a complex matrix calculation takes us directly from the values determined in steps a and b to the optimum matrix mentioned in step e. Then, we determine the average error that would occur in actually using it.

The scale of the final metric is such that a value of 100 represents no average error (“100% metameric accuracy”), and lower values represent greater average error (“lesser accuracy”).

This metric 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.

It is important to note that this property pertains to the behavior of the sensor, not to the entire camera. While we of course are ultimately interested in the performance of the entire camera in this regard, this involves both the performance of the sensor and the impact of all the subsequent processing, including demosaicing, with its own inherent transformation of coordinates, further transformation of coordinates, tonal scale mapping, color correction, and many other things.

And we may choose to apply various schemes of processing, other than those practiced by the camera, to the sensor data (as when we take the sensor data in “raw” form in a special output file and process it in our choice of external “raw conversion” or “image development” software).

Thus, in planning these other stages, both as may be done in the camera and as may be done by external software, we must work from knowledge of the “low-level” behavior of the sensor itself in various respects, including the matter of metameric error.

We may, in the final camera design, actually use the “optimum” transformation matrix developed in the course of the procedure described above (modified so that the “destination” color space is not CIE XYZ but instead our desired “output” color space, such as sRGB or Adobe RGB.)

In the case of a CFA sensor, such a transform does not appear as a distinct stage of processing, but is (at least in part) inherent in the details of the demosaicing algorithm. We can craft that aspect of the demosaicing algorithm so the overall result, for areas of non-trivial size of a constant color, is the same as if that “optimum” matrix were used in the non-CFA context.

THE ROLE OF THE ILLUMINANT

The MSI

If we look at the DSC/MSI results given for various cameras by, for example, DxO laboratories, we find that they have tested the cameras for metameric error under two illuminants, standard illuminant A and standard illuminant D50. The result is typically different.

This should not be surprising. Recall that the MSI is an indicator of the degree of metameric error committed by the sensor. The test is run over a repertoire of spectrums, those of the light reflected by a certain set of test patches (each with a known reflectance spectrum), as illuminated by the illuminant being used. Metameric error conceptually refers to the inconsistent report, by the sensor, of the color of light onto the sensor for different instances of light having the same color but different spectrums.

It would not be surprising that the error would depend on which of the specific spectrums we chose to use. This in turn depends on the illuminate used for the tests (just as much as it would depend on the particular test patches used).

The optimum transformation matrix

In those same reports by DxO Laboratories, for each camera tested, and for each of the two test illuminants, the report gives the "optimum transformation matrix" determined with the procedure in ISO 17321.[12] This matrix is again different for the two test illuminants.

But that is hard to grasp. If the sensor outputs consistently reported the color of the light (albeit in a parochial color space), then the matrix from that color space to, say, sRGB should be definitive and fixed. The matrix should not depend on where the light landing on the sensor came from.[13]

The key to this mystery is in the underlined "if". In fact, the sensor outputs are not consistent with the color of the impinging light (which is what metameric error is). The matrix crafted with the ISO 17321 procedure is an arbitrary one that seeks to minimize the overall average of the metameric error committed for a certain set of input spectrums.

If in fact, between two test series, we change that set of input spectrums—whether by changing the collection of "test patches" used, or merely by changing the illuminant on the patches—then the metameric error situation changes, and typically the "optimum" matrix will change, to "try and beat" the new situation.

How can we then exploit our craftily-devised optimum transformation matrix in the camera design? By choosing one as a compromise (it is a compromise compromise).

---

[12] Not quite; it seems to give the matrix that is not the optimum matrix determined per ISO 17321, which works from the sensor outputs to a CIE XYZ representation, but rather that matrix transformed so it works from the sensor outputs to an sRGB representation. But the relationship between those two is definitive.

[13] It was in fact my unease with this mystery that set me off on the long investigation of digital camera colorimetry matters that led to this article and the earlier one, its spiritual godfather, the one mentioned below.

FOR MORE INFORMATION

More detailed information on this topic, on some of the collateral matters discussed above, and on several related issues may be found in the article, “Digital Camera Sensor Colorimetry”, by this same author, probably available where you got this article.

#