

White Balance Color Correction in Digital Photography

Douglas A. Kerr

Issue 2
December 13, 2009

ABSTRACT

Variations in the chromaticity of the light under which a scene is photographed result in variations of the chromaticity recorded for different objects, often leading to an “unnatural” color appearance in the finished image as viewed. Avoiding this effect is the task of the *white balance color correction* process. In this article, we discuss the concepts involved, as well as some of the ways in which the process is performed in digital photography. We also discuss various “measurement tools” used in this connection. The discussion is detailed, but not mathematical.

BACKGROUND

Chromaticity

Chromaticity is that property of light that distinguishes red from green, and distinguishes red from pink. It is the property that many lay people think is meant by “color”. However, in color science, *luminance* (which we can think of as indicating the “brightness” of the light) is also an aspect of *color*.

The chromaticity exhibited by objects

The light that reaches us from a visible object has a certain spectrum, the relative distribution within the light of components having different wavelengths, across the range of visible wavelengths. This can be plotted as a curve of spectral density versus wavelength.¹

Note that there can be many different spectrums that are all perceived as having the same chromaticity, a situation called *metamerism*.

It turns out that color is a “three-dimensional” property in the mathematical (not geometric) sense: any color can be described by three numerical values. If we think in terms of the subordinate properties of luminance and chromaticity, then, since luminance can be described by a single number, chromaticity must require two numbers to describe it. There are many sets of two “coordinates” that we can use for that purpose: *hue* and *saturation*, the *x* and *y* coordinates on the CIE chromaticity diagram, and so forth.

¹ The formal name for this is *spectral power density* (SPD).

If we consider a reflective object (that is, one that is not self-luminous), the spectrum of the light coming from it is the product of the spectrum of the incident light (the “illumination”) and the *reflectance spectrum* of the object. The reflectance spectrum is the “plot” of the fraction of the arriving light that is reflected as it varies with wavelength. (In this case, *product* means the product of the value of the incident light spectrum and the value of the reflectance spectrum of the object for every frequency across the visible range.)

If the spectrum of the illumination and the reflectance spectrum of the object are relatively “smooth”, and if the chromaticity of the illumination is somewhere in the region we can consider as “white”, then it is possible to approximately calculate the chromaticity of the reflected light from the chromaticity of the illumination and what we can call the “reflective chromaticity” of the object.

Chromatic adaptation

Imagine that we have a sheet of “white” paper, illuminated by a “green” floodlight. The paper will of course appear “green” to us. Now we illuminate it instead with a “red” floodlight, and it appears “red” to us.

But in more subtle cases, things don’t work out quite that way. For example, the spectrum of light from the mid-day sky is quite different from the spectrum of the light from the incandescent ceiling fixture in our kitchen. Yet, if we examine our favorite green Fiestaware dinner plate out on the patio, or on the kitchen table, it seems to us to be the same “color”. How can this be?

The answer lies in an extraordinary capability of the human visual system.

To rather oversimplify a rather complex situation, we can note that when looking at a “scene” that contains “familiar” objects (objects whose “reflective” chromaticity we know, subconsciously, from experience), the human visual system, from the detected chromaticity of the light from these familiar objects, essentially deduces the chromaticity of the illumination. It then interprets the chromaticity of the light from all objects in the scene in that vein and “reports” to the higher levels of our perceptual process the presumed reflectance chromaticity of the objects. The result is that the dinner plate seems to us to be the same “color” regardless of under what light we view it (assuming that this light is sort-of “white”).

Does this extraordinary ability of the eye (spoken of as *chromatic adaptation*²) bring the same benefit to the after-the-fact viewing of photographic images, delivering to our consciousness the same impression of the color of an object regardless of the chromaticity of the illumination on the object when it was photographed?

² “Chromatic adaptation” also refers to a mathematical adjustment made to the “measured” chromaticity of a light source to take into account the eye’s chromatic adaptation, and to another related mathematical adjustment involved in the measurement of reflective chromaticity.

Unfortunately not. The eye does not, in this situation, have the opportunity to examine the light reflected from a range of objects (outside the image frame) illuminated by the light in the environment where the photograph was taken. Rather, the eye bases its chromatic adaptation on the light from familiar objects in the environment in which we are viewing the photo image—the kitchen table, the wall calendar, our own hands, etc.

The consequence is that our eye will not properly correct for the fact that the chromaticity of the different objects in the photo depends on the chromaticity of the illumination under which the photo was taken. The result is a potentially “unnatural” color appearance of familiar objects in the photo (sometimes called a “color cast”).

To overcome this undesirable effect is the purpose of a process called “white balance color correction” in our overall photographic chain.

The end of the chain

With some simplification of what is a terribly complex issue, we can say that the viewer of a photographic image will “get the ‘proper’ impression of the color of objects” if the chromaticity of any object in the delivered image is the same as the “reflective chromaticity” of the real object.

To clarify this with an example, suppose we are working with an actual object whose reflective chromaticity is “neutral”; that is, the chromaticity of the light it reflects is the same as the chromaticity of the light incident on it. For example, this might be the “pure white” background of a letter we are photographing.

Suppose that, in the digital image, the background parts of the letter have an RGB code of 180,180,180. That implies a specific chromaticity—the chromaticity of the “white point” of the color space. (The white point of an RGB-type color space is the chromaticity implied by any color code for which $R=G=B$. For the sRGB color space, widely used for images in JPEG form, that is the chromaticity of a certain “standard illuminant”, called “D65”.)

Then, the viewer will perceive the letter background on the display as being “white”, just as if viewing the letter directly.³

Is this a problem?

If we view the image as a print (and we’ll assume that the printing chain faithfully renders the reflectance chromaticity as represented in the image file), the eye will adapt to the chromaticity of the ambient illumination where the print is viewed. Won’t this adaptation spoil the ideal result we went to so much trouble to achieve? Fortunately not.

³ This is only strictly true if the incident light at the location of the display has the same chromaticity as that of the white point of the color space, but we won’t obsess over that here.

Remember, the print is also illuminated by the ambient illumination of the viewing environment. Thus the reflected light from the print will have its chromaticity varied according to the chromaticity of that illumination. But the chromatic adaptation of the eye corrects for that variance. As a result, the impression we receive is still of the reflective chromaticity of the object—it's "real color" (assuming that the correction described above has been properly made for the chromaticity of the ambient light for the shot).

On-screen viewing in an unknown environment

When we send off a color-corrected image in JPEG form, we have no idea if the viewer will view it in an environment in which the incident light has the same chromaticity as the white point of the sRGB color space. If not, then the chromatic adaptation of the eye will work counter-productively, and cause the viewer to perceive properly-recorded objects "improperly". We really can't do anything about that.

If the viewer is a sophisticated worker in the field of images, though, and knows that his work station area has an incident light chromaticity that is not the same as the white point of the sRGB color space,⁴ he can take certain steps in his image "pipeline" for our image to "outsmart" the effect of his chromatic adaptation on his perception of our image⁵. These are beyond the scope of this article.

The bottom line

Accordingly, in most cases, when we perform "white balance color correction" on a taken image, our objective is to have the chromaticity for any object in the delivered digital image be the same as the "reflective chromaticity" of the object (a simple conclusion for a complicated chain!).

Therefore, we can easily check for the proper working of our color correction process. We include in a test scene an object that we know has a "neutral" reflective chromaticity (such as a calibrated "white target" or "gray target"). We photograph the scene and then apply our color correction process (as we will discuss shortly). We then load the "corrected" image into our favorite image editor and examine the RGB coordinates of our white or gray object. If we find that $R=G=B$, then we say, "ahh—the color correction worked just perfectly". We will refer to that result as "theoretically ideal white balance color correction".

⁴ In much high-level graphics art work, the images are handled in a color space whose white point is the chromaticity of "standard illuminant D50", and accordingly, often work station areas in that field are illuminated with light matching that illuminant.

⁵ This is also described as *chromatic adaptation*, in the second of the three senses mentioned previously.

SCIENCE AND ART

By defining “theoretically ideal white balance color correction”, I do not mean in any way to suggest that this is what we should strive for in every, or even any, real photographic project. Often, various “artistic” considerations lead us to apply a different color correction. In portrait work, we may intentionally apply a “bias” to the process in order to produce a “warmer” skin tone. In some cases, the whole idea of color correction is counter-productive. If we take a scene illuminated by the beautiful light of sunset, indeed all the objects in the delivered image will look “more yellow” than their reflective chromaticities. But we may want the viewer to perceive just that, a result that would be negated by the application of “theoretically ideal white balance color correction”.⁶

Still, in the part of the field in which people argue about the chromatic neutrality of one color balance measurement diffuser or another, we have no choice but to think in terms of some quantifiable “ideal” result. This article proceeds in that vein, with the desired result as defined just above.

WHITE BALANCE CORRECTION “IN CAMERA”

The battle zone

Here, we will assume that the photographer takes the “developed” image from the camera (in JPEG or TIFF form), in which case the white balance correction process is conducted by the camera, and has its effect on the output image.

If, instead, the photographer takes the “raw” image from the camera (the essentially-unprocessed output of the sensor array), and has the image “developed” outside the camera in raw conversion software, then the white balance correction process is conducted there. The concept is identical, but there are some practical differences, which we will address in a later section.

Can we “not do it at all”?

Is there a concept of “not doing white balance color correction at all”? There is, although it is a little arbitrary. To me, that would mean that the delivered image would, for each pixel, represent the actual chromaticity of the **light received by the camera** from that point in the scene.

But the chromaticity represented by any given digital “code” for a pixel is in fact dependent on the color space used for the image. It then turns out that “no color correction at all” is equivalent to performing white balance color correction based on an assumed illumination whose chromaticity is the same as that of the reference illuminant defined for the color space in use (which we can think of as the same as the “white point” for the color space).

⁶ Thanks to Asher Kelman for reminding me of this particular, and telling, example.

So, we might think that, at least in more sophisticated cameras, there would be some kind of “preset” that produces this result. But there isn’t (at least in any of the cameras I know about).

But that’s all right, as in fact there would rarely be any need for this special result in normal photography (there could, though, be a need for it in various kinds of scientific or forensic photography).

The concept of in-camera white balance color correction

Suppose that the camera has somehow been made aware of the chromaticity of the illumination on the scene when the shot is taken.⁷ (Don’t worry now about how this can happen—we will discuss that at length a little later.)

The camera then reckons the difference between that chromaticity and the “reference white” chromaticity of the “color space” under which the image is recorded, such as sRGB or Adobe RGB.

Then, the camera shifts the chromaticity of every pixel in the captured image by the opposite of that difference.⁸ The result (ideally) is that for each object, its coded chromaticity in the delivered (“color corrected”) image is that of the object’s reflective chromaticity—it’s “real color”, in lay terms.

HOW DO IT KNOW?

How does the camera become aware of the chromaticity of the incident illumination on the scene at the time the photograph is taken? There are several possible ways.

We can tell the camera

We can inform the camera of what we know (or guess) about the chromaticity of the illumination in one of several ways; which are available depends on the camera model.

Named light types

We can choose a named type of light from a list on the camera (“full daylight”, “full shade”, “incandescent light”, “photoflash, etc.). For each, the camera assumes a preordained chromaticity, typical for that “kind” of illumination.

⁷ You may wonder why I say “the chromaticity of the **illumination**” rather than “the chromaticity of the **incident light**”. There is an important distinction, but I will not explain it at this point, in the interest of moving forward on the principles. The matter is discussed in Appendix A.

⁸ The process is really a little more complicated than that, but this description will serve to illustrate the essence of the process.

Set the chromaticity

We might have measured the chromaticity with a special instrument (a *chromatic photometer*).

We can then tell the camera the chromaticity of the illumination by entering two values (on some system of coordinates, perhaps the axes of the CIE x-y chromaticity diagram). There can be many mysteries in this mode as implemented on different cameras (the details of which are beyond this article).

Set by “color temperature”

We can enter the known, supposed, or measured “color temperature” of the illumination, a measure that denotes the specific chromaticity of a certain class of illuminants. It doesn't really apply to, or precisely describe, other types of illuminant. But, with an added wrinkle, it is nearly applicable to a wide range of natural light chromaticities as well as to the chromaticity of the light from various types of incandescent lamps. (Appendix B gives an extensive discussion of color temperature, including what it is and what it isn't.)

Let the camera measure it—the “target” approach

We can let the camera measure the chromaticity of the ambient illumination, for example, in this way. (The specifics described are for the scheme as practiced in Canon EOS digital cameras.)

We place in the scene, where it will be illuminated by the light that will illuminate the subject, a “white balance target”: a diffuse reflecting “plate” with a neutral reflective chromaticity (in fact, ideally more than that: a uniform spectral response over the range of visible wavelengths). Its color might be what we would think of as “white” or “gray”, depending on its absolute reflectance.

We then take a “white balance reference frame” shot in which the white balance target occupies at least a certain portion of the center of the frame (prescribed by the camera manufacturer).

We then tell the camera to determine the average chromaticity of that region of the image, which will be considered to be the chromaticity of the ambient light. The camera stores that chromaticity, and can then be told (for any subsequent shot) to use that “learned” chromaticity as the assumed incident light chromaticity for white balance correction during image development. (This is often described as the “custom white balance” setting.)

A detailed issue here is in what direction the target should face. This is a complex issue, which I discuss in Appendix A. It involves the distinction between “the chromaticity of the illumination on the subject” and “the chromaticity of the incident light”, alluded to earlier.

Let the camera measure it—the “diffuser” approach

Alternatively, we can temporarily fit the camera with a “white balance measurement diffuser” a disk-like device placed in front of the lens. Its job is to accept all the light that lands on it (coming from various directions) and convert it to a “uniform luminous disk” facing the lens. In effect, it temporarily converts the camera to a chromatic photometer, able to determine the chromaticity of the incident illumination.

We place the camera where the diffuser will receive the same illumination as does the subject. We then tell the camera to determine the average chromaticity of (the central region of) the image of that luminous disk, which will be considered to be the chromaticity of the ambient illumination. We then proceed as in the previous mode.

There are many ramifications of this technique, and it will be discussed in considerable detail in a later section.

The camera can guess

The camera can be told to try and deduce the chromaticity of the ambient illumination at the time of the shot by analyzing the distribution of recorded chrominance in the captured image. (The camera is asked to do the same kind of chromatic adaptation the human eye does. News report: it is rarely as good at it as the eye is, and in any case it is handicapped by only being able to see a narrower view of the environment than could a “live” human observer.) I have no idea how the algorithms used for this purpose work. This mode is often called “Automatic White Balance” (AWB).

WHITE BALANCE CORRECTION WITH RAW DATA FROM THE CAMERA

In many cases, photographers do rely on the camera to “develop” an image from the suite of sensor data. Rather, they take a camera output file that almost verbatim presents that suite of sensor data (called a “raw” file) and use an external software package to “develop” the image. This allows better control over how the development is done, allowing (for example) the ability to mitigate importune exposure. Many photographers also feel that the development algorithms in their favorite software package “do a better job” than what is used inside the camera.

Normally, the application of white balance correction is closely integrated with image development, and thus in this situation white balance correction is done, subject to the user’s control, during external image development.

As before, to set this up, the software needs to know the chromaticity of the incident illumination on the scene. There are three ways to do this:

- We can instruct the software of the chromaticity of the incident light (as we know or assume it), typically by setting it in two coordinates. In many software

packages, these are the *correlated color temperature* and the *Planckian offset*.⁹ (The reader is referred to Appendix B for an explanation of this.)

- We can have the camera make a determination, in the way previously described, from a white balance reference frame, using a neutral test target, and store that for use as the basis for internal white balance correction (even though we may not take advantage of that). When the raw output is taken from the camera (and the raw data is not modified by the in-camera color correction—only a JPEG output would be), the camera will nevertheless report the correction it *would have made* in the “header” of the raw data file. The user can then just tell the image development software, “do the white balance correction just as the camera would have done it.”
- We can have the camera take a reference shot of a scene including a neutral target (placed as discussed earlier) and take the raw data output for that shot (as well as for the actual scene shot). In the image development software, the recorded chromaticity of the image of the test target is read (often the user uses an “eyedropper” tool to probe the image in the proper region) and adopted as the basis for the white balance correction of the actual scene shot(s).
- We can include in the scene, at the time of the actual scene shot, a white balance target that will be innocuous in the finished scene shot. We then take the raw output, and (in our “development” software) point to the neutral target with our “eyedropper” and tell the software, “this is the chromaticity of the incident illumination—rework this image accordingly.” We might then “airbrush out” the target during final image editing.

WHITE BALANCE CORRECTION WITH DEVELOPED IMAGES FROM THE CAMERA

If the photographer has taken “developed” image data from the camera (typically in JPEG form) but has not performed white balance color correction in the camera, such color correction can still be applied in an image editor. However, there are various impediments to full success in doing so. The principal problem is that the dynamic range of one or more color “channels” in the sensor system has already been diluted by the scaling needed in image development.

Because of the often-disappointing results attained with this process, many image editors do not provide a full suite of tools to facilitate its use.

⁹ In raw development software, these are often called “temperature” and “tint”.

MEASUREMENT OF THE INCIDENT LIGHT WITH THE CAMERA AND A DIFFUSER—DETAILED DISCUSSION

The white balance measurement diffuser—basic technique

We spoke before about using the camera to measure the chromaticity of the incident light with the aid of a neutral reflective target (“gray card”) placed at the subject location.

We also spoke before fitting the camera temporarily with a “white balance measurement diffuser”, a translucent disk, reminiscent of a filter, whose properties we will discuss shortly. In effect, it temporarily converts the camera to a chromatic photometer. We then (in the preferred technique) place the camera with diffuser at the subject location.

There is a complicated issue of in exactly what direction to orient the camera; I discuss this in detail in Appendix A. For the moment, just assume that the camera is oriented so the diffuser faces generally where the camera will be when the actual shot is taken.

Basically, the diffuser collects the incident light falling on its face and from that collection generates a uniform “luminous disk” on its back side (toward the lens).

The lens takes that luminous disk and from it creates a pattern of light on the sensor. (It isn’t really an “image” of the luminous disk, since the disk is in a “gravely out-of-focus” situation, given its proximity to the lens.)

We now have the camera record that “image” and analyze it just as before to determine the average chromaticity over some portion of the frame. This is then considered to be the effective chromaticity of the incident illumination, and is used as the basis for in-camera color correction of the subsequent actual photographs.

The diffuser just does, in a “transmissive” mode, what the neutral target does in a “reflective” mode.

Ideally the diffuser should exhibit the following properties:

- A. It should combine light from “streams” arriving from different directions weighted by the cosines of their respective angles of incidence.
- B. It should deliver a luminous disk whose color (luminance and chromaticity) is uniform over its entire area, and uniform with respect to the angle from which it is viewed.
- C. The chromaticity of the luminous disk should be the same as the effective net chromaticity of the illuminance afforded by the collection of incident light “streams”.

Requirement A is needed so the diffuser will collect the different incident light streams in a way that mimics their impact on the surface. The underlying technical concept is discussed in Appendix A.

Requirement B is needed since we have no idea how a the particular lens on the camera, set in a particular way (focal length, aperture, and focus distance) will take the light it collects from the rear of the diffuser and deposit it on the sensor, nor what parts of the sensor the particular camera considers in making white balance measurements.

Requirement C is needed since we rely on the light being delivered to the sensor for measurement having the same chromaticity as the incident illumination on the subject.

Requirements A and B together can be said to require that the device be “transmissively Lambertian”. Requirement C is often spoken of as calling for its “chromatic neutrality”.

Manufacturers of (most of these) diffusers go to great lengths to attain these objectives. For example, it is difficult to attain requirement A with a simple diffuse transmissive element—we just can’t find real stuff that is “transmissively Lambertian”. Thus, for example, such techniques as a domed diffuser, or an array of little prisms or lenses in front of the diffusing element proper, are often used to force this behavior.

Often requirement C is met by the inclusion in the device of a transparent filter layer, in some cases selected (perhaps by manufacturing batch) to null out the departure from neutrality of the diffuser proper.

An alternate technique

Some photographers, and the manufacturers of some white balance measurement diffusers, may advocate (sometimes with reservations) an alternate technique.

Here, we leave the camera at the position to be used to photograph the subject, aimed at the subject. We then fit the diffuser to the lens and make the chromaticity measurement. This is delightfully handy! It saves all that walking over to the subject position. And certainly better if the subject is a black bear. (They won’t hold a gray card for you, either.)

But how could this reveal the chromaticity of the incident light illumination upon the subject? It seems paradoxical.

Still, surprisingly enough, we find that in many actual cases, this technique produces a measurement that, used as a basis for white balance color correction, yields an acceptable result. How can this happen?

The most credible theory is that this works if (and only if) the incident illumination upon the front of the camera (and thus the diffuser) happens to be nearly the same

in chromaticity as the incident illumination upon the subject. Examples of where we might expect this to be closely true would be a large ballroom, where essentially all the illumination comes from a large number of identical chandeliers, uniformly spaced across the entire ceiling, or many “open” outdoor settings.

Proponents of this technique say that this condition (near equivalence of the incident light on the camera position and on the subject) is not necessary for the “alternate technique” to reliably yield usable measurements. Still, I haven’t heard from any of them any “blackboard story” about why that should be—just exhortations that I should, in effect, hark to Hamlet’s words, “There are more things in heaven and earth, Horatio, than are dreamt of in your philosophy”.

Many diffuser manufacturers today do not even mention this technique, or urge that it only be used when the basic technique is not practical. On the other hand, the manufacturer of one “non-cosine” diffuser says that it is specially suited to this technique. I am unable to relate that assertion to what I know of its technical properties. I’ll talk a little about it later.

Naming the two techniques

Recall that, briefly, the two “diffuser” techniques are:

- A. Place the camera, with the diffuser in place, at the subject location
- B. Leave the camera, with the diffuser in place, where it will be for the actual shot, and aimed in the same direction.

Often we hear these described as the “incident light” and “reflected light” techniques, respectively. This is an attempt to draw a parallel with two techniques used for exposure metering, but the parallel is not apt.

The description of technique “B” as “reflected light” is intended to imply that the light whose chromaticity is observed is the light reflected from the subject. However, if the diffuser has a fairly broad “acceptance pattern”, only a small part of the light it accepts is that reflected from the subject; most of it is light from other sources: for example, from the sky almost overhead, or to the left, or right, or reflected from a red barn just to the left of the subject). It is in fact essentially the incident light **upon the diffuser** where it is right now—at the camera position for the shot. And that’s why this technique works at all.

Now note that in technique “A” (“incident light technique”), part of the incident light measured there is also reflected, for example, from the photographer’s assistant standing at the tripod where the camera will go for the shot, or from the wall of the blue barn behind the tripod. Yet we do not speak of this as a “reflected light” measurement.

Now suppose that the diffuser in use has a “very narrow” acceptance pattern. Isn’t it now fair to say that, under technique “B”, the measurement is basically of “the light reflected from the subject”? Sadly, yes.

“Sadly” because the chromaticity of that light is greatly affected by the reflective chromaticity of the subject, while what we seek to determine is the chromaticity of **the illumination upon the subject.**

If in fact we could reliably determine the chromaticity of the incident illumination from the light reflected from the subject, we would not need a diffuser at all. We would just zoom in on the subject with the “bare lens” (so the subject filled the “white balance measurement region” of the frame) and ask the camera to take a white balance measurement exposure.

A non-cosine diffuser

One widely marketed diffuser does not have an acceptance pattern that even approximates the “cosine” pattern called for by the classical theory of white balance measurement, but rather a somewhat narrower pattern.

The manufacturer of this diffuser says it is specially suited for what I call above “technique B” (measurement from the camera position). The manufacturer is not very forthcoming as to what unique properties confer that ability on it or how. But the intimation is that it comes from the relatively narrow acceptance pattern of this device. I am unable to make that connection on any conceptual grounds.

In any case, how does the manufacturer of that device attain this relatively-narrow acceptance pattern?

Supposedly, it comes from the fact that the actual working diffuser is of relatively small diameter (much smaller than the overall diameter of the unit)—most of the “diffuser disk” is masked by an opaque black ring.

There is no theoretical reason that this masking should reduce the breadth of the acceptance pattern. And in fact tests done here show that it has almost no effect on that property.

In fact, what causes the acceptance pattern of this diffuser to be narrower than that of “cosine” diffusers is the fact that no special features are incorporated to broaden it; the inherent acceptance pattern of typical practical diffusing materials is rather narrower than the “cosine” pattern.

To give a cosine pattern to a diffuser using a practical diffusing material as its basis, as mentioned above, special features have to be incorporated—often an array of small prisms or lenses in front of the diffusing material proper. This is in fact part of the reason for the relatively high cost of typical “cosine” diffusers.

By merely declining to incorporate such features, the manufacturer of the diffuser being discussed achieves its “relatively narrow” acceptance pattern. And, because of misconceptions about the premise of white balance measurement from the camera position (that it is intended to “capture the light reflected from the subject”), that narrower pattern is made to sound beneficial.

I have discussed above why, in many real cases, a useful white balance measurement can be made from the camera position. But there is no reason to believe that a narrower acceptance pattern will enhance this prospect—in fact, the theory tells us to the contrary.

SKIN TONE ADJUSTMENT

Often, in portrait or fashion photography, we may wish to alter the skin color from that produced by the application of “theoretically ideal” white balance color correction, for “artistic” reasons. (A common one is that the subject wants to look “more tanned” in the portrait.)

If an overall chromaticity shift of the image will produce the effect we want (or can live with), then this process can be integrated with the white balance color correction process.

For example, many manufacturers of white balance measurement diffusers offer alternate models whose transmission is not “neutral”. If, for example, the transmissive chromaticity is in the “bluish” direction, then an image color corrected using a measurement made through such a diffuser will be, overall, rather “warmer” than the image given a valid “theoretically ideal” color correction.

Also available are “quasi-neutral” measurement targets, not actually neutral, but with a reflective chromaticity that will have the same effect as the non-neutral measurement diffusers mentioned above.

For use in color correction during image processing, one manufacturer offers a target disk with multiple non-neutral sectors. The photographer can choose one of these sectors during image processing to be the “white” target, thus providing an overall shift in image chromaticity from that which would have resulted from the application of “theoretically-ideal” color correction. Through practice, the photographer can learn which of these sectors produces the desired result for certain situations and certain skin “types”, thus allowing consistent replication of the technique.

#

APPENDIX A

ORIENTATION OF THE TARGET OR DIFFUSER

When using a neutral target to determine the chromaticity of the incident illumination, the question arises, “which way should the target face?” We hear such guidelines as:

- It should face the camera.
- It should face the light source.
- (If there are multiple light sources) It should face the “principal” light source.
- (If there are multiple light sources) It should face some place in the center of the light sources.
- It should face halfway between the camera and the dominant light source.

There is in fact an almost identical issue with regard to the orientation of the face of a diffuser in the basic diffuser technique (“at the subject”). We will treat that at the end of this Appendix.

If all the light comes from a single source, or from multiple sources of the same chromaticity, or from an extended source of uniform chromaticity, the orientation of the target doesn’t matter.¹⁰ All that is necessary is that the target be illuminated by that light (and, of course, that it be visible to the camera).

But suppose that there are two different sources of presumably-different chromaticity—perhaps daylight entering through a window from our right, and light from a table lamp from our left. The subject is significantly illuminated by two light “streams” of different chromaticity. What happens then?

Our concern now becomes what I call the “net effective chromaticity” of the overall illumination on the subject. (Note that I pointedly speak of “the *illumination*” rather than “the *incident light*”. The distinction will be seen shortly.) By this I means the chromaticity that a single “stream” of light on the subject would need to have to produce the reflected light chromaticity that we actually have.

As we consider this notion, let’s first review an important principle of photometry, not directly related to chrominance. The impact of a “stream” of light on a surface (which impact we quantify via the photometric property *illuminance*) depends not only on the “potency” of the light stream (which we quantify via the photometric property *luminous flux density*, about which we don’t really hear much) but also on the angle at which the stream strikes the surface (the *angle of incidence*, which we measure from a line perpendicular to the surface). In particular, the *illuminance*

¹⁰ We assume here that the target has “Lambertian” behavior, that of an “ideal diffuse reflecting surface”.

caused on the surface by the stream is the product of the luminous flux density of the stream and the cosine of the angle of incidence. Hold that thought.

Let's spend a moment considering, in general, the effect of "mixing" the illumination from two streams of light of different chromaticities falling on some surface. The chromaticity of the overall illumination on the surface can be determined (conceptually) on the CIE chromaticity x-y diagram if we know the chromaticity of each stream and the relative **illuminance** each causes on the surface (not its relative **luminous flux density**—ah, those pesky cosines). This "chromaticity mixing" is shown on figure 1.

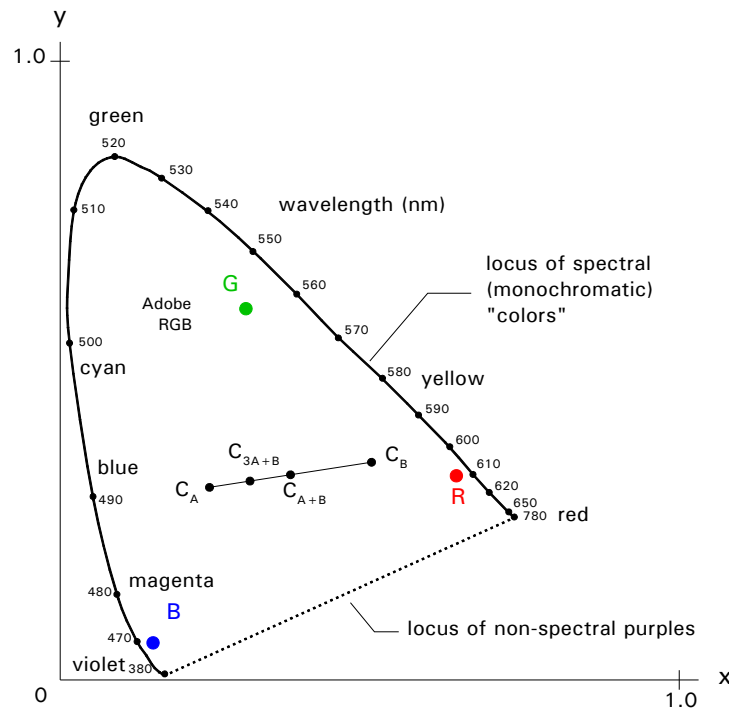


Figure 1. Chromaticity "mixing"

In the example, we show the mixing of light of two different chromaticities, C_A and C_B . First, we will assume equal illuminance caused by two light streams having those chromaticities. Then, the chromaticity of the "overall illumination" would be located at the center of the line (on the CIE x-y plane) joining the two "ingredient" chromaticities (point C_{A+B}). We can call this the "net effective chromaticity" of the joint illumination.

Next we will assume that the illuminance from the stream of chromaticity C_A is three times that from the stream of chromaticity C_B . Then the chromaticity of the joint illumination would be 1/4 the way from C_A to C_B on the line joining them, (point C_{3A+B}).

Now, consider the illuminance of our gray card by two streams of light arriving from different directions and having (maybe) different luminous flux densities and

(certainly) different chromaticities. Recall that the illuminance afforded by each stream (and thus the “weight” given to its chromaticity as an ingredient of the overall net effective chromaticity) is dependent both on its luminous flux density (a characteristic of the stream itself) and **the cosine of its angle of incidence**.

The bottom line of this is that the chromaticity of the net illuminance on the surface will vary with the orientation of the card (because of the effect of that orientation on the two angles of incidence). Thus, the chromaticity of the reflected light (as seen by the camera in a “measurement” shot) will also vary with the orientation of the card.

Won’t it also vary with the direction from which the card is observed? Not if the surface is in fact an “ideal diffuse reflecting surface” (a so-called “Lambertian” surface, since it obeys Lambert’s law and its corollaries). There, the observed luminance is constant, regardless of the direction of observation. Thus, the contribution to luminance of each of the two streams is independent of the angle of observation, and thus so is the observed chromaticity.

Now, let’s return to our actual problem, how to orient our target card. First, consider the case in which the subject is a planar (“flat”, in the geometric sense) surface—perhaps a poster). And we will assume this surface is Lambertian.

Then the answer is simple: the target should face the same way the subject surface does. Then the chromaticity of the composite illumination on the target will be the same as the chromaticity of the composite illumination on our photo subject, which clearly is what we want to determine for white balance purposes.

But now consider a case where the subject isn’t a planar surface—for example, a human face. Different portions will face in different directions. Thus, in our “mixed light” situation, they will receive composite illumination having different net chromaticities. What direction should our test card face?

Well, note that since the different surfaces of the face will receive illumination of differing chromaticity, we just cannot have proper white balance compensation over the entire face—regardless of how we test for the chromaticity of “the illumination”. Our dilemma is far more fundamental than the matter of the orientation of our test target card.

We must come to grips with this, by answering the question, “what part of the face do I wish to be best white balance corrected?” Once we decide this, then the answer as to the orientation of the target card comes automatically: it should face the same direction as the “chosen” part of the surface of the face. (And, in a practical sense, we might typically just “choose” the portion of the face that “faces” the camera.)

Now, suppose that the subject surface, while planar, is substantially non-Lambertian. (This is very true, for example, of “pearlescent” automotive finishes, which are in fact intended to exhibit a different “lightness” when viewed from different directions). Again, in this case, no simple measurement of incident

light will give us an “ideal” basis for color correction, so we need not complain that the premise for our theoretical “orientation” doesn’t hold up here—the whole process doesn’t work in any simple way.

Application to diffuser measurement

The same chain of reasoning applies to the use of a measurement diffuser.

First, we must recognize that, for the diffuser measurement to be equivalent to the reflective target technique, the diffuser must have the property that the luminance created on the back of the diffuser by a particular arriving light “stream” must be proportional to the luminous flux density of the stream and the cosine of the angle of its incidence on the diffuser face. We can speak of this as the diffuser having a “cosine acceptance pattern”. It is the transmissive equivalent of a “Lambertian” reflective surface.

Now, with that out on the table, we should be able to recognize that the guidelines for the orientation of the face of the diffuser are exactly the same as for the orientation of a reflective target. Simplistically, the diffuser face should be placed parallel to the “principal” surface of the subject.

What about “at the camera location” diffuser measurement?

We might then ask about guidelines for orienting the camera and diffuser when measurement is made from the camera location to be used for the actual shot.

Recall that this technique only works if the illumination on the camera location is essentially the same as that upon the subject.

But we can hardly expect to answer the question. “what orientation of the diffuser here will cause the ambient illumination on the diffuser to be precisely the same as the ambient illumination on some particular surface of the subject **over there?**”

There is no answer.

The common misconception about this technique (that it is based on measuring the light reflected by the subject) would suggest that the measurement be made with the lens axis pointed at the subject. Of course, there is no theoretical or conceptual basis for this outlook.

But we might as well do it that way as any other—it eliminates the need to re-aim the camera between measurement and actual shots.

ACKNOWLEDGEMENT

Special thanks to Carla Kerr for her careful and insightful copy editing of this tedious manuscript.

APPENDIX B

Color Temperature

We often hear the chromaticity of a light source described in terms of its *color temperature*. In this appendix, we briefly discuss the concept of color temperature.

The blackbody radiator

In physics, a *black body* surface is one that absorbs all the radiation landing on it, regardless of wavelength. Since all is absorbed, none is reflected, hence the apt moniker “**black body**”.

Any surface whose temperature is greater than absolute zero (0 K^{11}) emits radiation. The spectrum of the radiation depends on the temperature (as well as on the nature of the surface). For most surfaces, from a temperature of perhaps 2000 K to perhaps 30,000 K, part of the spectrum is in the visible range, and thus the radiation is visible. At the lower end of the temperature range, the chromaticity of the visible radiation tends toward the red, and at the higher end, toward the blue.

For a black body surface, the spectrum at any temperature—and thus the chromaticity of the light emitted—can be precisely predicted on theoretical grounds. If we plot that chromaticity on the CIE chromaticity diagram (with coordinates x and y), over the range of temperatures, we get the curve shown in figure 2. This curve is called the *blackbody radiation locus* (or Planckian¹² locus).

Color temperature

Any chromaticity lying on the Planckian locus (a “blackbody chromaticity”) can be exactly specified by giving the temperature at which blackbody radiation has that chromaticity. That temperature is called the *color temperature* for that chromaticity. Thus a color temperature precisely specifies a certain chromaticity, **which will always be a blackbody chromaticity**. (Conversely, no chromaticity not lying on the Planckian locus has a color temperature.)

¹¹ The scale of absolute temperature is the Kelvin scale. Its unit is the Kelvin (symbol K). The unit was formerly called the degree Kelvin ($^{\circ}\text{K}$). A temperature of 0°C is equivalent to about 273 K. The unit Kelvin (K) is the same size as the unit degree Celsius ($^{\circ}\text{C}$).

¹² Named in honor of noted physicist Max Planck, who gave us the first complete theoretical analysis of blackbody radiation.

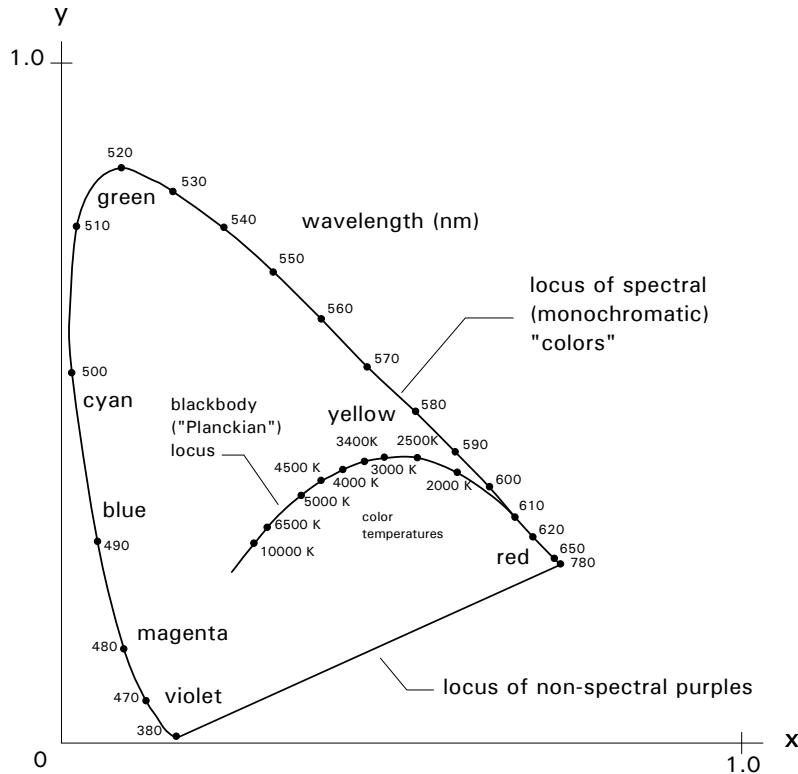


Figure 2. The blackbody locus

Correlated color temperature

However, for chromaticities not lying too far from the Planckian locus, it is handy to give a general idea of the chromaticity by giving the color temperature of that blackbody chromaticity that is “visually most similar” to the chromaticity of interest. That color temperature is called the *correlated color temperature* (CCT) of the chromaticity of interest.

On the CIE u - v chromaticity diagram (different from the CIE x - y chromaticity diagram, which we saw in figure 2), distances between points representing different chromaticities correspond roughly to the degree of visual difference between the chromaticities. Thus, the point on the blackbody locus whose chromaticity is “visually nearest” the chromaticity of interest is the point on the locus that is geometrically closest to the point representing the chromaticity of interest (see figure 3).

When we hear of a chromaticity having a “color temperature” of $xxxx$ K (such as the “color temperature” of a certain photographic light source), rarely is it a true color temperature that is actually meant, because in fact few such sources have a blackbody chromaticity—one lying on the Planckian locus—and no other kind of chromaticity has a color temperature. Rather, what is usually meant is that the chromaticity has a *correlated color temperature* of $xxxx$ K.

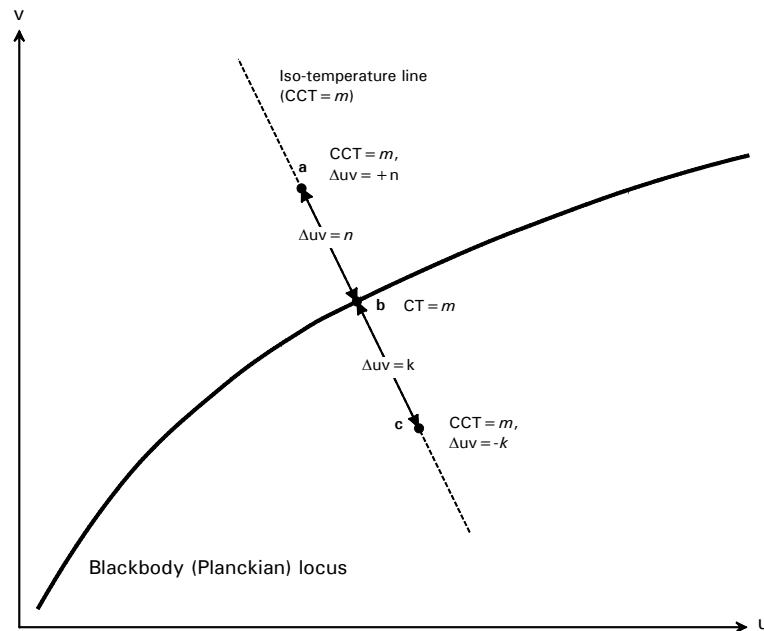


Figure 3. Correlated color temperature (CCT) on the CIE u-v plane

In the figure, point **a** represents the chromaticity of interest. Point **b** is the point on the blackbody locus closest (geometrically) to point **a**. The color temperature of point **b** is m . Thus we say that the correlated color temperature of the chromaticity at point **a** is m .

The other value we must state

But of course there are many chromaticities that have the same correlated color temperature. On the figure, for $CCT = m$, all points that lie on the dotted line¹³ will represent chromaticities having that same CCT. The line is called the "iso-temperature" line ("equal temperature", *temperature* of course meaning "correlated color temperature"). Thus we cannot specify a non-blackbody chromaticity by stating its correlated color temperature alone. To complete the specification of chromaticity, we must specify in some way how far the chromaticity of interest is from its blackbody locus "idol", and in which direction. This can be done in several ways.

A common way (the best, actually) is to specify, again working on the CIE u-v chromaticity diagram, the actual geometric distance between the point representing the chromaticity of interest and the nearest point on the blackbody locus.

Owing to the properties of the u-v coordinate system, that distance in fact is basically proportional to the visual difference between the chromaticities of points **a** and **b**.

¹³ It is perpendicular to the tangent to the locus at point **b**.

This distance, called *delta uv* (or Δuv), is in the same units as the scales on the u-v diagram. By convention, a positive value means that the chromaticity of interest is “above” the Planckian locus; a negative value means it is below.

Thus, in the example, for the chromaticity of point **a**, that distance is n , and so $\Delta uv = +n$. Thus, for the chromaticity of point **a**, by stating its CCT (m) and its value of Δuv (specifically, $+n$), we completely specify the chromaticity.

Note that point **c**, also located on the same iso-temperature line, and representing a different chromaticity than point **a**, also has $CCT = m$. It however has a different value of Δuv (specifically, $-k$).

Many image development software packages, in connection with white balance color correction, allow the user to specify the assumed chromaticity of the incident Illuminance in terms of correlated color temperature and Planckian offset. But they are rarely labeled as such.

Often the setting for correlated color temperature is labeled “color temperature”, or just “temperature”. A typical label for the Planckian offset is “tint”.

“Temperature” is of course expressed in Kelvin. In Photoshop (which, with an accessory, can provide for image development), the value of “tint” is 3000 times the amount of the Planckian offset in the units of the CIE u-v chromaticity diagram. A positive sign refers to a chromaticity “above” the Planckian locus (the “greener” direction), while a negative sign refers to a chromaticity “below” the Planckian locus (the “more-purple” direction).

#