

Color Reproduction for Digital Cameras

Michael Kriss

Adjunct Professor of Physics at the Portland State University, Oregon, USA;

Previously with Eastman Kodak; University of Rochester, NY; and Sharp Laboratories, USA

1 Introduction

Figure 1 shows one of several scenarios for the color reproduction in a digital still camera (DSC). The scene objects are illuminated by a single or multiple light sources. The resulting image collected by the lens has a continuous, point-by-point, spectrum that is focused and filtered by the lens. For most cases, one can ignore the filtration of the lens because it has little impact on the color of the scene. Normally there is a combination of an array of micro-lenses, used to collect most of the light, and a color filter array, CFA, used to encode the color on either a CCD (charge-coupled device) or CMOS (complementary metal oxide semiconductor) sensor. The CCD or CMOS sensors normally have a broad response to light ranging from the ultraviolet (UV) to the near infrared (IR). To ensure that the sensor is responsive to only light seen by the human visual system, from roughly 400–720 nm, two filters are used to block light below 400 and above 720 nm. Owing to the CFA, three sparse images are created on the sensor. The most common CFA is due to Bayer [1] and contains 25% red, 50% green, and 25% blue pixels and can be seen in Figure 2 along with two other CFAs. The next step is to correct for the white balance of the scene and this is done after the “raw” data has been digitized. The white balance can be obtained in several ways. In the automatic mode, the image is analyzed during the focus and exposure level operation before the camera “captures” the scene. During this step some algorithm is used to determine the optimum balance between the red, green, and blue signals (encoded by the CFA) to give an optimum white or gray value for any neutral portion of the scene. The output from the algorithm drives the relative gains of the red, green, and blue pixels in the raw data. The green gain is also used to adjust the camera ISO (International Organization for Standardization) speed of the sensor

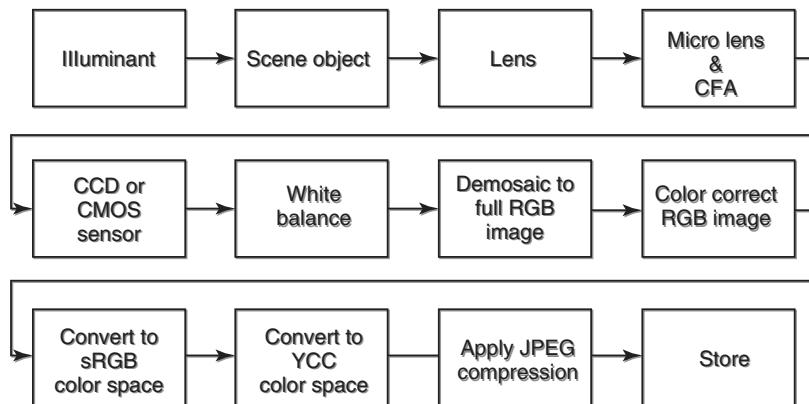


Figure 1 The color reproduction chain for a digital camera

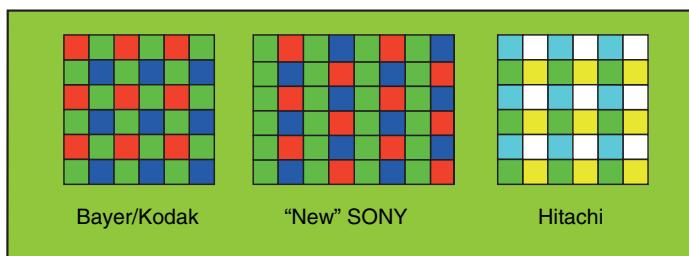


Figure 2 Three CFAs. The Bayer CFA is found in most digital cameras. The SONY CFA was introduced in the late 1980s to replace the Bayer CFA. The Hitachi CFA was used first in video cameras and in some digital cameras; it has better sharpness, but more color artifacts

(not the native speed of the sensor, but the camera speed). Other methods for white balance include preset gains for well-determined illuminants such as blue-sky, cloudy-sky, tungsten illumination, and fluorescent illumination. Another method allowed by most advanced cameras is to let the user photograph a gray card under the prevailing illuminant and the image processing hardware and software calculate the relative red, green, and blue gains and apply them to the subsequent image sequence. In the case of other than the Bayer CFA, the camera image processing algorithms use different algorithms to achieve the same results. The next step is to use some sophisticated demosaicing and interpolation algorithms to create full red, green, and blue images. These algorithms are designed to preserve any true neutrals in the image and to eliminate any color artifacts because of the sparse sampling introduced by the CFA. Once an optimum red-green-blue image is obtained, free of artifacts, several types of color processing is used to “correct” the color image so that the color errors introduced by the less than optimum spectral sensitivities of the sensor and CFA elements. The spectral sensitivities of the human eye are very complex and subject to adaption to the surrounding lighting conditions. Consider walking indoors from a bright blue sky to a room illuminated by a tungsten or fluorescent light fixture. Initially the room will appear dark and yellow or green depending on the illuminant. However, in a short time, the room will appear at a normal level of illumination and neutrals will appear white or gray. A camera cannot do this type of adaptation, hence the need for the

white balance discussed earlier. Furthermore, owing to practical needs for camera speed, the color cross talk between the red, green, and blue channels (or other channels for non-Bayer CFAs) the colors may appear muted and drab. Better color separation and reproduction can be achieved by both linear and nonlinear manipulation of the data. Linear methods include simple linear transforms using a 3×3 matrix that tends to compensate for the channel cross talk while preserving neutrals. Nonlinear methods can be applied to the red-green-blue data of each pixel to compensate for the color cross talk in a more detailed manner and even increase the color saturation of each channel while preserving the neutrals. Additional color management can be done once the data is converted to luminance and chrominance channels used for compression; see **Image Compression and File Formats**. The next step is to convert the enhanced color channels to a standard color space. The standard color space is used to drive color monitors and is based on the nature of the color monitors. Before liquid crystal displays, LCDs, took over the market the phosphors in CRTs (cathode ray tubes) determined the color spectrum (gamut) of a color monitor. On the basis of the phosphor emissions, color primaries, a color space called *sRGB* was developed that would optimize the color form a CRT-based monitor. So, ideally, each camera manufacturer would produce a color conversion algorithm to store *sRGB* values (before and after compression) so that the color images would look more or less the same on any monitor calibrated for *sRGB*. There are other color standards that provide a wider (larger) color gamut based on LCDs primaries and these will be discussed in detail the following sections. Before JPEG (Joint Photographic Experts Group) compression takes place, the red-green-blue image is converted to YC_rC_b color space where Y is the luminance channel and there are two chrominance channels denoted by C_r and C_b . Figure 3 shows a red-green-blue image converted to an YC_rC_b image. Note that only the luminance channel, Y , carries detailed information while the C_r and C_b chrominance channels carry very little detail, but all the color information. Before compression it is possible to enhance the detail of the image by operation on the Y luminance channel and enhance the color by operation on the two chrominance channels. The operations on the luminance channel will be covered in the chapter on Digital Imaging, **Digital Imaging: An Introduction to Image Processing**. One can easily enhance the color saturation by simply operating on the two chrominance channels before the JPEG compression. The chapter on Compression and File Formats, **Image Compression and File Formats**, will cover the compression processes.

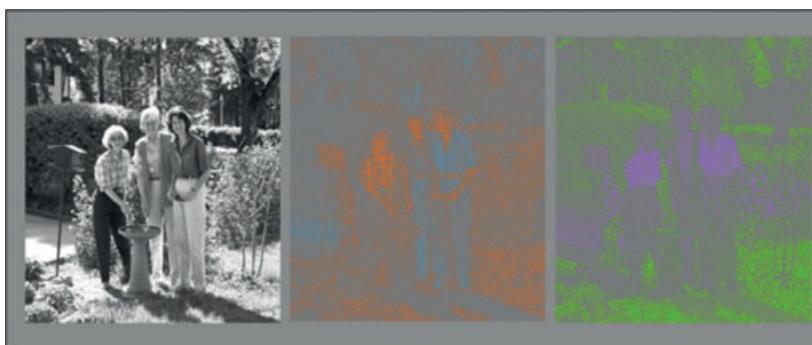


Figure 3 An YCC image where the two chrominance channels have a digital value of 128 added to all pixels to eliminate the negative values. It is clear that the chrominance channels carry little detail information

With this short introduction to color reproduction in mind, it is necessary to consider how one defines color within the human visual system, and then move to the details of the color reproduction in DSCs.

2 CIE Color Matching System

It is necessary to review how one defines a color and establishes the methodology of matching colors based on different sets of primary wavelengths, dyes, and inks. The heart of such a processes is defined by the CIE (International Commission on Illumination) [2–6] color system: the Commission de l'Eclairage. The first step is to understand that the human visual system seems to respond to three color signals often called the *short*, *middle*, and *long wave cone receptors* in the eye's retina (along with the noncolor response of the rods). In simpler language, the human visual system seems to respond to blue, green, and red lights. Newton [7] was first to demonstrate that white light is made up of a full spectrum of colors (violet, blue, green, yellow, orange, red, or the full visible set of colors), and Maxwell [8] demonstrated that three color separations, red, green, and blue could produce a full color image with all the colors of the rainbow represented. The first definitive experiments to define color vision made two very important sets of measurements. The first was to define the photopic and scoptic luminance efficiency [9] curves as shown in Figure 4. It was noted that the brightest appearing color was a green at 555 nm. At the same radiant energy, all other colors were less bright. So an experiment was designed to get the relative and exact difference in brightness between all visible colors. Holding the brightness of the green signal at 555 nm fixed, a calibrated white light source was used in conjunction with a prism or grating to generate all other colors in nanometer increments. The subject was asked to adjust the power (radiant energy) of the white light source for each wavelength (from 380 to 780 nm) until the brightness of the test color and that of the green signal at 555 nm appeared to be the same. The ratio of the radiant energy then defined the luminance efficiency curves. The photopic curve represents adaptation to normal light levels, whereas the scoptic curve represents adaptation to low light levels (moon and star light). Figure 4 shows the results. Once the photopic curve was established, the second experiment was undertaken.

The second experiment [10, 11] is outlined in Figure 5. Three primary colors of red, 700 nm, green, 546.1 nm, and blue, 425.8 nm, are chosen such that they will stimulate one and only one of the three cones on the retina of a human observer. The idea is simple, given the right amount of each of these primaries can, in theory, create any color visible to a human observer. Consider using three modulated lasers at these wavelengths in a DLP (digital light processing) [12] monitor, which uses an array of very small mirrors in an on-or-off node to display an image. There can be three mirrors, one for each primary, or a single mirror that uses time-multiplexing combination of the three primaries. In principle, the image on the monitor should be able to recreate all possible color images. Given this concept, one now needs to measure what amount of light of the three primaries will be needed to reproduce each color (say 1–5-nm wide) in the spectrum. Figure 5 shows how this can be done. Each of the three light sources can be varied in its radiant energy (resulting in some perceived brightness). A white light source is used to create a test color using a prism or grating and an aperture mask that limits the wavelength from 1 to 5 nm in width. The observer is then asked to adjust the three light sources until the test color is matched. Note that these three primaries have very different luminance efficiencies as shown by the photopic curve in Figure 4. The relative amounts of radiant energy to give an arbitrary

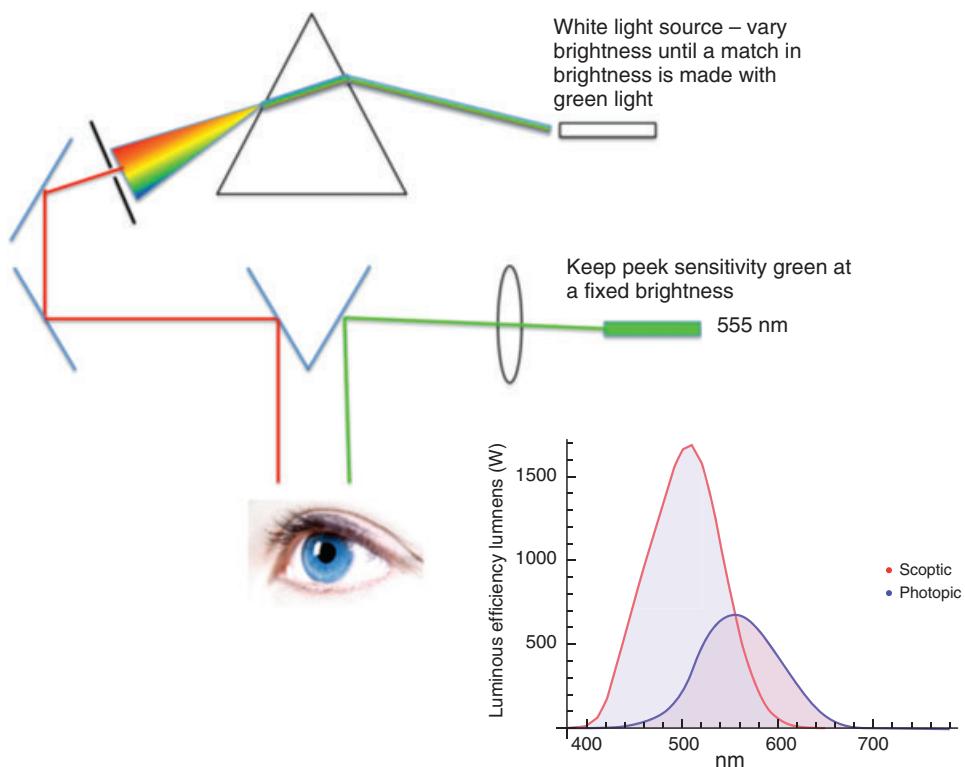


Figure 4 The photopic and scoptic curves are measured by comparing other wavelength of light with the brightest color of green at 555 nm. The observer adjusts the white light source until the selected color is as bright as the fixed 555 nm green so the ratio of the radiant energy at the same brightness gives the luminance efficiency

unit of white (measured as brightness) light is 1 unit of radiant energy of blue light at 425.8 nm, 1.4 units green light at 555 nm, and 72.1 units of radiant energy of red light at 700 nm. These power ratios then form the units that need to be adjusted together to maintain a “white” of various intensities. Calibrated in this way, the three sources are adjusted to get the red, green, and blue brightness values to match the equal brightness values of the spectral colors from the white light source. These responses are called the *CIE standard observer color matching functions*, $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, and $\bar{b}(\lambda)$. Figure 6 shows the CIE standard observer color matching functions.

The first thing one notes about the color matching functions is that $\bar{r}(\lambda)$ has a strong negative lobe in the blue-green (cyan) region. As light cannot have negative values, this clearly indicates that just three color primaries will not produce all visible colors. In the experiment outlined earlier, it became clear that some colors could not be matched by the chosen (or any) three primary colors. However, a valid mathematical construct could be achieved by adding some red light to the test color and then making the match with the green and blue primaries; this led to the negative red lobe in the cyan region. The r-g-b color matching functions can now be used to define a color based on its reflection spectrum, $Ref(\lambda)$, and the illuminant, $I(\lambda)$. Let R

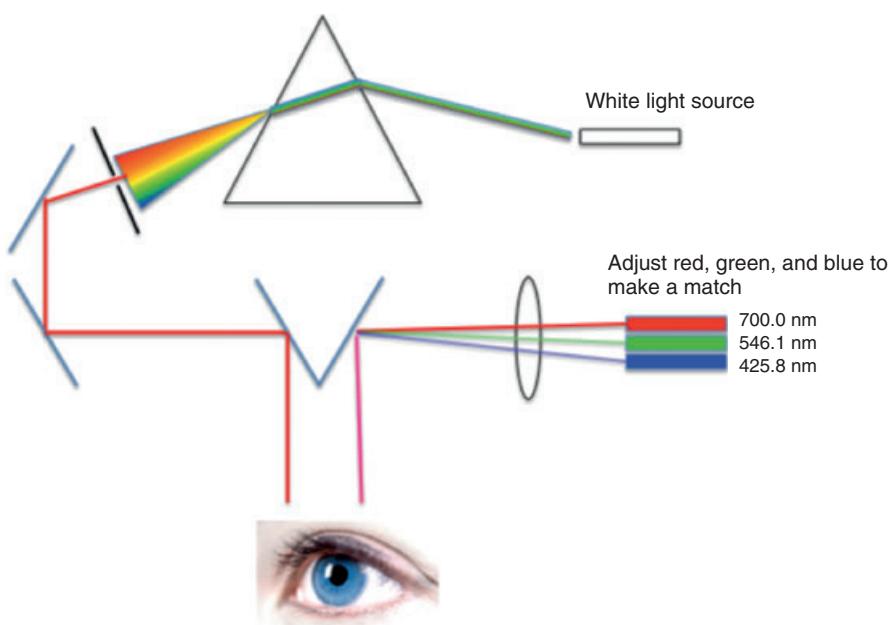


Figure 5 The color matching concept. Three adjustable sources are used to match a test color defined by a white light source of equal visual brightness source and a prism (or spectral grating) and slit (5 nm or less) to define the test color. Each primary source is adjusted and calibrated to form a white or neutral shade when set to equal values

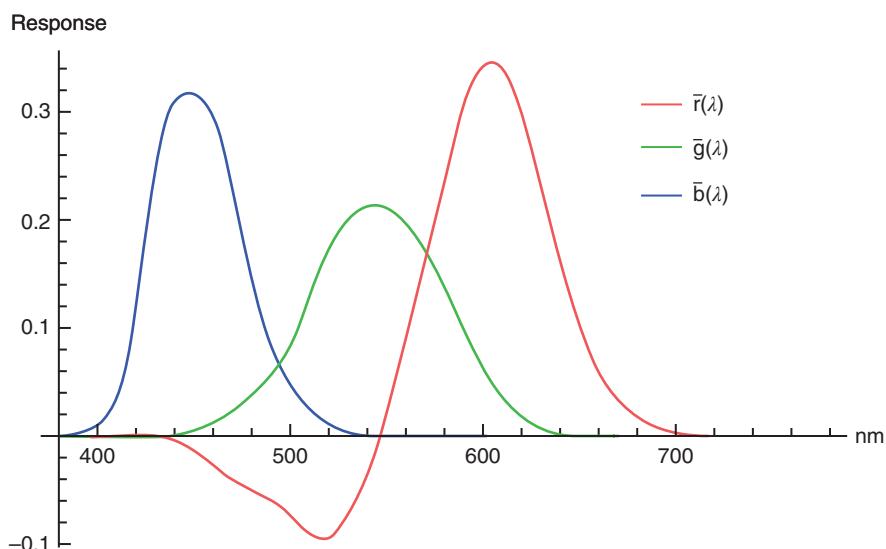


Figure 6 The CIE standard observer color matching functions $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$

stand for the red response to a color and it is given by

$$R = \frac{\int_0^{\infty} \text{Ref}(\lambda)I(\lambda)\bar{r}(\lambda)d\lambda}{\int_0^{\infty}\bar{r}(\lambda)d\lambda}. \quad (1)$$

In a similar manner, the green, G, and blue, B, values of a color are given by

$$G = \frac{\int_0^{\infty} \text{Ref}(\lambda)I(\lambda)\bar{g}(\lambda)d\lambda}{\int_0^{\infty}\bar{g}(\lambda)d\lambda} \quad (2)$$

and

$$B = \frac{\int_0^{\infty} \text{Ref}(\lambda)I(\lambda)\bar{b}(\lambda)d\lambda}{\int_0^{\infty}\bar{b}(\lambda)d\lambda}. \quad (3)$$

The above triad of R, G, B values for all combinations of reflection spectra and illuminants will fill a three-dimensional space. However, if one lets the rays from the origin to any {R, G, B} point in this color space and note where they intercept the {1, 1, 1} plane in the R-G-B space, one will get a well-defined measure of the color, but not of its intensity or brightness. In the {1, 1, 1} plane, let r, g, and b be defined as

$$r = \frac{R}{R + B + G}, \quad (4)$$

$$g = \frac{G}{R + B + G}, \quad (5)$$

$$b = \frac{B}{R + B + G}. \quad (6)$$

Further note that in the {1, 1, 1} plane,

$$b = 1 - r - g. \quad (7)$$

Using Equations 1 through 6, one can calculate the r, g, b values for all visible wavelengths, thus defining the volume of all possible colors in the R, G, B space or the boundary of all possible colors in the {1, 1, 1} plane as two-dimensional plot of r versus g, where b can be calculated by Equation 7. Figure 7 shows what is called the *r-g chromaticity plot*. Note that there are large areas of negative values for r.

Although completely adequate to define a color, the R-G-B and r-g-b system was modified (by means of a linear transform) to the CIE X-Y-Z system including the x-y-z chromaticity plot. The CIE X-Y-Z color system was defined to make the resulting color matching functions,

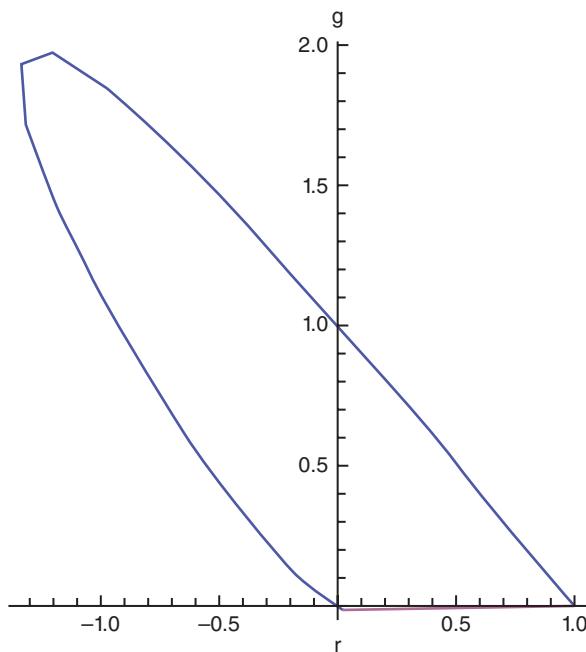


Figure 7 The r-g chromaticity plot of spectral colors in the RGB {1, 1, 1} plane. All colors lie inside this plot and the red line at the bottom indicates where all colors made up of the red and blue primary lie

$\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)$, positive for all wavelength and that the $\bar{y}(\lambda)$, color matching function be equal to the photopic luminance efficiency curve, $V(\lambda)$. To do this requires the calculation of a new set of color “nonreal” primaries, which are a linear transform of the “real” primaries at 700, 546.1, and 425.8 nm. These new primaries can be obtained by finding three primaries in the r-g chromaticity plot, Figure 7, that when connected by straight lines will just totally encompass all the spectral colors. Along with the criteria for $V(\lambda)$ noted earlier, the new nonreal primaries is given by

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{bmatrix} = \begin{bmatrix} 1.275 & -0.278 & 0.002 \\ -1.739 & 2.767 & -0.028 \\ -0.743 & 0.141 & 1.602 \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{G} \\ \mathbf{B} \end{bmatrix} \quad (8)$$

where $\{\mathbf{R}, \mathbf{G}, \mathbf{B}\}$ represent the original primaries at 700, 546.1, and 425.8 nm and $\{\mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$ are the new, nonreal primaries. The new primaries are a linear combination of the original primaries, but with strong negative values. As light cannot have negative values, while these new primaries are mathematically valid (one can match colors), they are not physically realizable. On the basis of Equation 8 and the constraint on $V(\lambda)$, one has

$$\begin{bmatrix} \bar{x}(\lambda) \\ \bar{y}(\lambda) \\ \bar{z}(\lambda) \end{bmatrix} = \begin{bmatrix} 2.769 & 1.752 & 1.130 \\ 1 & 4.591 & 0.060 \\ 0 & 0.057 & 5.594 \end{bmatrix} \begin{bmatrix} \bar{r}(\lambda) \\ \bar{g}(\lambda) \\ \bar{b}(\lambda) \end{bmatrix}. \quad (9)$$

The x-y-z color matching functions are shown in Figure 8.

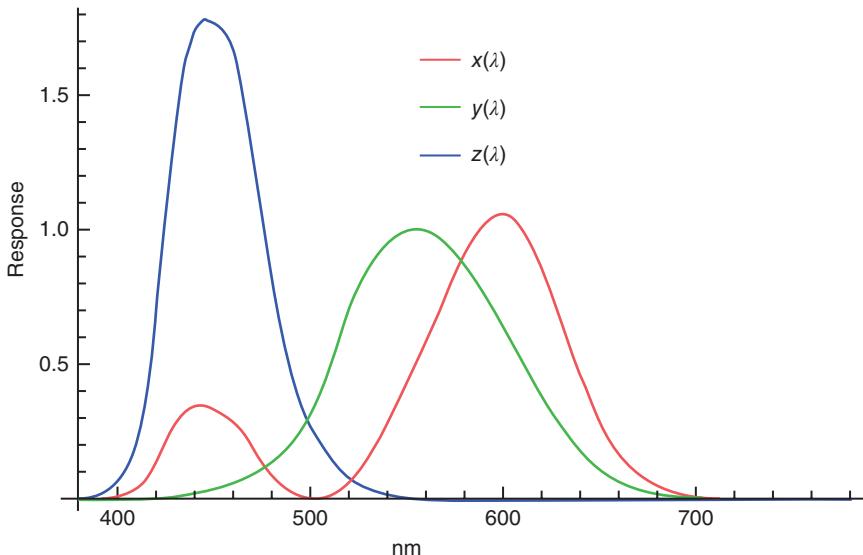


Figure 8 The x - y - z CIE color matching functions based on the r-g-b CIE color matching functions shown in Figure 6

The color specification in the X - Y - Z space is similar to that in the R-G-B color space and is given by

$$X = 100 \frac{\int_0^{\infty} \text{Ref}(\lambda) I(\lambda) \bar{x}(\lambda) d\lambda}{\int_0^{\infty} \bar{x}(\lambda) d\lambda}, \quad (10)$$

and in a similar manner,

$$Y = 100 \frac{\int_0^{\infty} \text{Ref}(\lambda) I(\lambda) \bar{y}(\lambda) d\lambda}{\int_0^{\infty} \bar{y}(\lambda) d\lambda}, \quad (11)$$

and

$$Z = 100 \frac{\int_0^{\infty} \text{Ref}(\lambda) I(\lambda) \bar{z}(\lambda) d\lambda}{\int_0^{\infty} \bar{z}(\lambda) d\lambda}. \quad (12)$$

The x - y chromaticity diagram is generated in the same manner as the r-g chromaticity diagram. Any given color $\{X, Y, Z\}$ is projected on to the $\{1, 1, 1\}$ plane of the three-dimensional X - Y - Z color space with the following relationships holding:

$$x = \frac{X}{X + Y + Z}, \quad (13)$$

$$y = \frac{Y}{X + Y + Z}, \quad (14)$$

$$z = \frac{Z}{X + Y + Z}, \quad (15)$$

and

$$z = 1 - x - y. \quad (16)$$

Figure 9 shows the x - y chromaticity plot for the spectral colors. The line joining the blue to the red points represents where a linear combination of the red and blue primaries will produce the violets in the X - Y - Z color space. Note that there are no negative values for x or y . All possible spectral colors fall inside the horseshoe-like curve in the CIE x - y chromaticity plot.

Figure 10 shows three primary additive primaries, red, green, and blue, in the form of Gaussian sources with centers at 450, 550, and 650 nm and a sigma of 40 nm, each. Figure 10 also shows four other combinations to give cyan, magenta, yellow, and white lights. Using Equations 10 through 15 and assuming an equal energy source (flat) illuminant, Figure 11 shows where the seven colors fall in the CIE x - y chromaticity plot. As these primary colorants are rather broad, they cannot give saturated colors. Note that the white is plotted at $\{0.333, 0.333\}$ as a true neutral white should plot. The red primary is a little on the orange side based on the plot in the chromaticity diagram.

Figure 12 shows a Gretag–Macbeth ColorChecker used in almost all color studies. Gretag provides the spectra for each of these patches (under several standard illuminants). Again using the Equations 10 through 15, the CIE x - y chromaticity values are calculated and plotted in Figure 13.

Note that the colors from the Gretag–Macbeth ColorChecker occupy only a central portion of all the possible colors. The cluster of dots in the center of Figure 13 represent the gray scale

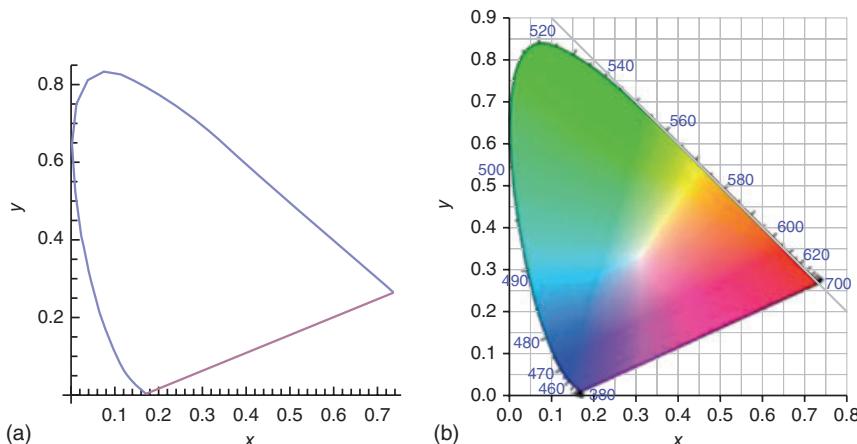


Figure 9 (a) The CIE x - y chromaticity plot. All possible colors fall within the envelope of the spectral colors that form the outer curve. (b) Representation of the colors associated with the CIE x - y chromaticity plot. These colors are not exact and are used only to show the relative placement of the colors. *Source:* Reproduced from Illustration of the CIE 1931 color space, as published in Wikimedia Commons, 2007, <http://commons.wikimedia.org/wiki/File:CIEExy1931.svg> © Sakurambo, 2007, distributed under the Creative Commons Attribution-Share Alike 3.0 Unported license

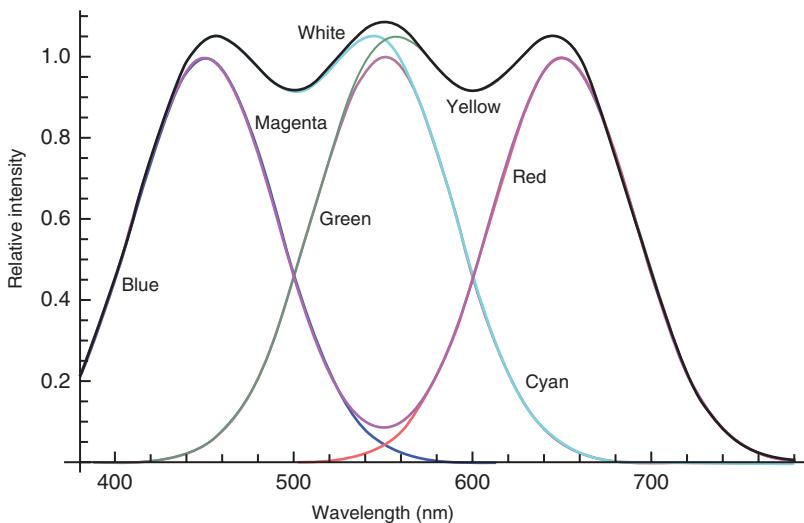


Figure 10 Three basic additive color primaries and several combinations to give a white along with a cyan, magenta, and yellow color

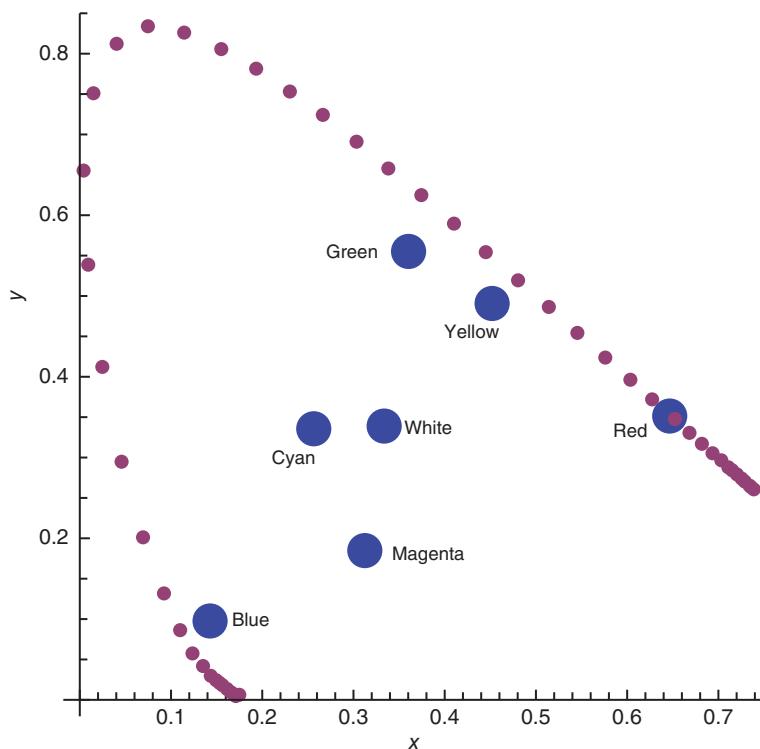


Figure 11 A plot in the CIE x - y chromaticity diagram of the spectrum found in Figure 10 along with the spectral colors forming the boundary of the allowed color space



Figure 12 The Gretag–Macbeth ColorChecker

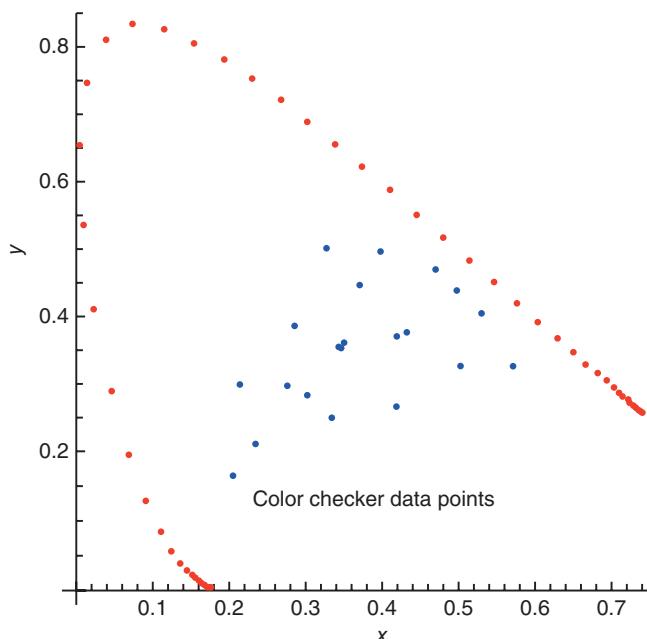


Figure 13 The plot of the Gretag–Macbeth ColorChecker in the CIE x - y chromaticity diagram

in the color checker; there is a slight variation in the neutral values, but they do center around $\{0.333, 0.333\}$.

The CIE X-Y-Z system provides an accurate definition of a color. However, different spectra (under a given illuminant) may give the same X-Y-Z values and will indeed look the same to the “standard CIE observer” for a given illuminant. However, under a different illuminant,

they will appear to be different colors. Such spectra are called *metamers* for a given illuminant [3–6, 13]. Figure 14 shows four different spectra, all of which have their X-Y-Z values (for a spectrally flat illuminant) of {0.309399, 0.272397, 0.66071}; a cyan-blue color. When these metamers for a flat spectral illuminant are illuminated with a more tungsten-like light source, see Figure 15, the resulting spectra will look to be different colors as shown in the CIE x-y chromaticity plot as shown in Figure 16. The lack of blue light in the tungsten source tends to discount the spectra with blue peaks shifted to the deep blue or UV (to lower than 450 nm).

There are several practical implications about metamers. It is clear that if a color is to be matched under a given illumination with a new set of dyes or pigments, it can be done.

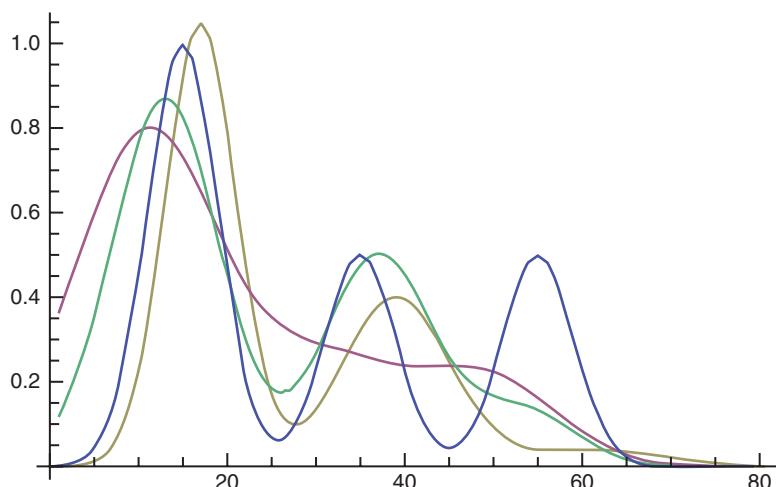


Figure 14 Four different spectra that have the same CIE X-Y-Z values

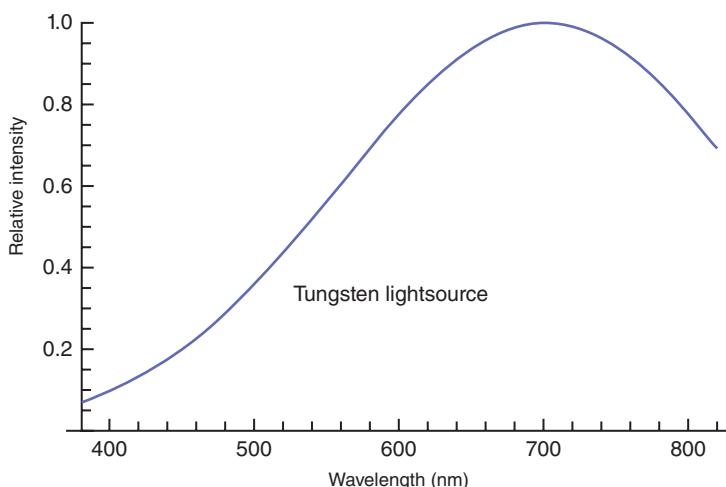


Figure 15 A tungsten light source that has low blue spectral content

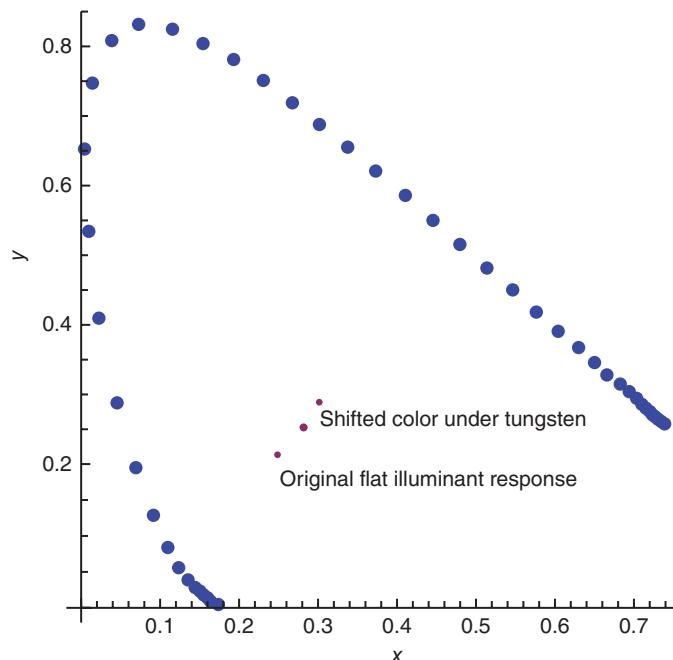


Figure 16 The original spectrum gives a CIE x - y chromaticity values of $\{0.249012, 0.219232\}$ under a flat spectral illuminant. When three other sets of dyes were used to generate the same x - y values (under the flat spectral illuminant) the x - y values were the same. But when a tungsten source such as in Figure 15 is used, the colors split up as shown, far from the original colors. Here a cyan color is shifted toward a neutral

However, if the new match is viewed with the “old” set of dyes or pigments under a different illuminant, they may not match. So a skirt and sweater that match under the store illumination may not match in daylight. Another complication is that some dyes fluoresce under UV illumination [14]. So matching slacks and jacket (like a tuxedo) under the rental store illumination may not match when wedding pictures are taken with flash equipment that gives off UV radiation. So when taking pictures, visual matching colors (under some illuminant) may not match in the image because the camera spectral sensitivities are not the same as the CIE standard observer’s color matching functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$.

The earlier discussions bring to light another question. What is the variation in X - Y - Z values (or x - y values) that are imperceptible to the human observer? This question was answered by a series of experiments by Dave Mac Adam in 1942 [15]. Figure 17 shows some of the results. At the center of each ellipse is the test color, and the boundary of the ellipse defines the colors that will be “seen” as the same color. The size and shape of the ellipses vary depending on where in the CIE x - y chromaticity diagram one is looking. There is far more tolerance in the green colors than in the blues or reds. Also, the orientation of the ellipses rotate as the test color goes from the greens to the reds or blues. Hence, it is clear that the CIE X - Y - Z color space is not uniform. It would be beneficial to create a mathematical representation (transform) of the CIE X - Y - Z color space that gave equal changes in color as circles rather than ellipses, which in turn might make matching colors easier.

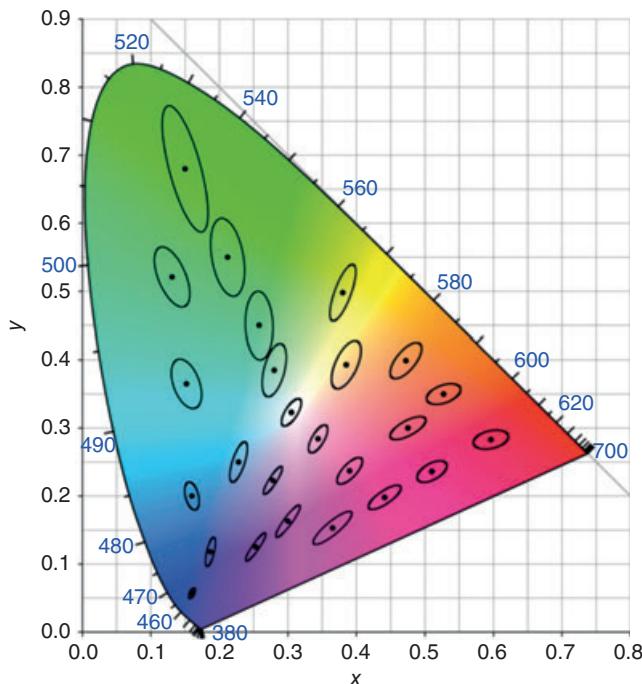


Figure 17 The CIE x - y chromaticity plot showing ellipses of equal color differences around the test colors marked by the black dots. *Source:* Reproduced from Ellipses plotted on the CIE xy 1931 chromaticity diagram as published in Wikimedia Commons, 2010, http://commons.wikimedia.org/wiki/File:CIExy1931_MacAdam.png © MacAdam, 1942, distributed under the Creative Commons Attribution-Share Alike 3.0 Unported license

Several other color spaces (transforms from the CIE X - Y - Z system) have been introduced over the years to make the color plots in these new color spaces more visually uniform (equal changes in color values map to equal visual changes in color). Two of the most popular (useful) are the $L^*u^*v^*$ and the $L^*a^*b^*$ or known as the *CIELuv* and *CIELab* systems, respectively [3]. The mathematical transforms are given later. Each of these transforms is “normalized,” denoted by the subscript “ n ,” to a given illuminant and thus provide some level adaptation, which corresponds to how the human visual system adapts to different illuminants. Also, unlike the rectilinear CIE X - Y - Z system, a constant luminance plane in the CIELuv or CIELab system does not have a fixed spectral boundary; the spectral boundary will change with each level of luminance.

CIELuv (for levels of $Y/Y_n > 0.01$)

$$\begin{aligned} L^* &= 116 \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16 \\ u^* &= 13L^*(u' - u'_n) \\ v^* &= 13L^*(v' - v'_n), \end{aligned} \quad (17)$$

where

$$\begin{aligned} u' &= \frac{4X}{X + 15Y + 3Z} \text{ and } v' = \frac{9Y}{X + 15Y + 3Z} \\ u'_n &= \frac{4X_n}{X_n + 15Y_n + 3Z_n} \text{ and } v'_n = \frac{9Y_n}{X_n + 15Y_n + 3Z_n}. \end{aligned} \quad (18)$$

CIELab(Y/Y_n , X/X_n , $Z/Z_n > 0.01$)

$$\begin{aligned} L^* &= 116 \left(\frac{Y}{Y_n} \right)^{1/3} - 16 \\ a^* &= 500 \left[\left(\frac{X}{X_n} \right)^{\frac{1}{3}} - \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} \right] \\ b^* &= 500 \left[\left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - \left(\frac{Z}{Z_n} \right)^{\frac{1}{3}} \right]. \end{aligned} \quad (19)$$

The CIELuv and CIELab color spaces were designed to be more uniform in terms of visual differences (ellipses turn into circles), so “geometric” distances between the desired color and the produced color is an accurate measure of color errors. The so-called ΔE color errors [3–5] are given by

$$\Delta E_{uv}^* = [(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2]^{1/2}$$

and

$$\Delta E_{ab}^* = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]^{1/2}. \quad (20)$$

These color difference errors can be used to fine tune the color reproduction algorithms that are used within DSCs (or associated software) to produce, on average, a “true” (that best match the scene X - Y - Z values) color reproduction, a vivid (more saturated colors) color reproduction, and a memory color reproduction (green grass and blue skies). Although often used to compare one color algorithm to another, in the authors opinion this is not very useful, as color errors are only obvious if the observer knows what a color is supposed to be. For example, the wife of a household will always note when the sofa or table cloth did not come out “right.” The one area where such color differences is very import is the flesh-to-neutral balance in an images. The color reproduction algorithms need to be tuned so that the errors in neutrals and flesh tone are very small, as any errors are always detected by an observer. For more recent measures of color differences based on opponent color theory see work by Wandell and Fairchild [16, 17].

Figure 18 [18] shows the R-G-B (X - Y - Z) color space along with the CIELab and CIELuv color spaces. One can easily see that the CIELab and CIELuv color spaces are highly distorted transforms of the R-G-B space, thus demonstrating the complex nature of human color vision.

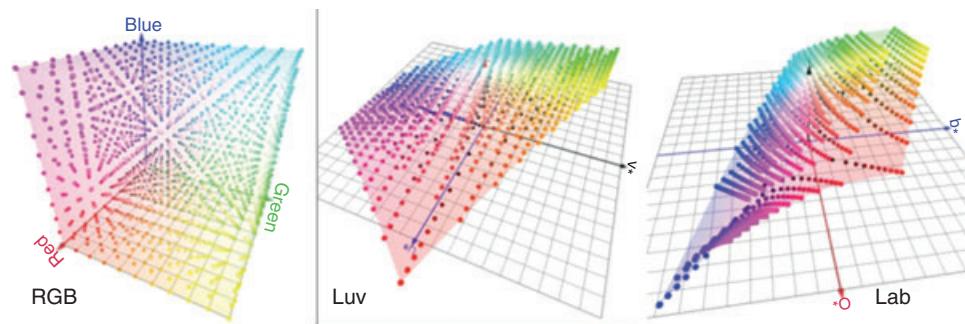


Figure 18 Three color spaces based on the CIE color system. Source: Reproduced with permission from Ref. 18. © Couleur, 2012

3 The Spectral Sensitivity of a Digital Camera

The net spectral sensitivity of a camera is determined by the spectral transmittance of the lenses (including micro-lenses), the color filters spectral transmittance, and thin film coatings on the sensors [19] leading to selective interference, and finally on the native quantum efficiency (QE) [20, 21] of the sensor used to convert light into stored electrons. For the purpose of this discussion, the impact of the lens transmittances will be ignored as well as any interference effects. This leaves the intrinsic sensor QE and the various color filters used in the CFA. Here it will be assumed that the camera uses a typical R-G-B filter array. In all cases an inferred filter is assumed as all CCD and CMOS sensor have strong IR responses, thus an IR elimination filter must be used.

The basic response of CCD and CMOS sensors will be assumed to be the same in what follows. There are some basic issues for QE for both sensors. Front and back illuminated [22, 23] sensors have different QE spectra. The more conventional front illuminated sensor has the light passing through some insulating layers, silicon dioxide, and poly-silicon electrodes all of which absorb blue light. Thus, the blue sensitivity suffers. Red and green light must pass deeper into the silicon substrate before they can be absorbed and producing photoelectrons, which are in turn captured by the potential well, depletion layer, near the top of the sensor. The thicker the substrates, the better the red and green sensitivity, but the less sharp the image is due to more lateral diffusion before the electrons are captured by the potential wells (located under the designated color element of the CFA). If illuminated from the back, the blue response improves, but a lower sharpness may occur because of the increased diffusion distance to the depletion region where the potential wells are found. Illuminating from the back improves the red and green sharpness because of the lower lateral diffusion on the way to the depletion region. Thinning the backside of the sensor can improve the blue sharpness without harming the blue response. Coating the backside of the sensor with anti-reflection thin film further improves the overall sensitivity, but may introduce some small peaks and valleys because of the wavelength dependence of the thin film interference. The use of micro-lenses [24] can greatly increase the QE as they collect light that would otherwise fall outside the active area of the pixel. Although all these characteristics are important for the overall quality of the sensor, one can ignore some of them when focusing on color reproduction. Any factor that affects the overall QE or that either creates or eliminates cross talk between pixels designated to “capture” red, green or blue light are very important. Figure 19 shows a range of effects that alter the QE

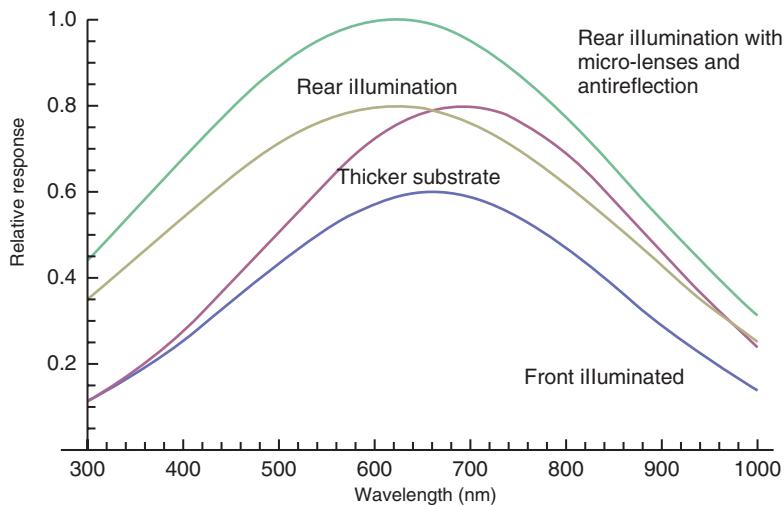


Figure 19 A series of QE curves for sensors shows variations due to physical differences

of a sensor; these plots are only meant to show relative changes and do not represent any given class or brand of sensors. The bottom curve represents a front illuminated sensor, which has low blue sensitivity owing to light absorption due to the silicon dioxide and poly-silicon above the active area (depletion region) and low red sensitivity due to the relatively thin substrate. The next curve shows the impact of a thicker substrate that allows greater red sensitivity. The third curve shows the first curve with rear illumination and the final curve represents the improved response to the rear illuminated sensor when micro-lenses and an anti-reflection coating is used. Again, these curves are just representative of the effects and do not correspond to any given sensor or manufacturer.

Figure 20 shows the impact of an IR filter, on the long wavelength response of a sensor. It is necessary to remove the IR response if one wishes to have good color reproduction. On the other hand, if one wished to look for temperature (heat) related imagery, one would remove the IR filter and maybe the CFA as well.

The issues surrounding the spectral sensitivities of a sensor and the subsequent color reproduction will now be considered. In this exposition, simple models will be used to define the transmittance of the color filters (in the CFA) as well as target color to be reproduced. In all cases, Gaussian functions will be used to represent the color filters and the colors to be examined. The sensor will be defined as the one in Figure 20 with the IR filter as shown. The QE is given as

$$\text{QE}(\lambda) = \frac{\text{QE}_{\text{MAX}}}{e^{\frac{\lambda - \lambda_{\text{IR}}}{h}} + 1} \cdot e^{-\frac{((\lambda - \lambda_0)^2)}{2\sigma^2}} \quad (21)$$

where for Figure 20, $\text{QE}_{\text{MAX}} = 0.8$, $\lambda_{\text{IR}} = 740 \text{ nm}$, $\lambda_0 = 730 \text{ nm}$, $h = 50 \text{ nm}$ and $\sigma = 300 \text{ nm}$. The value of h controls how fast the IR filter cuts off at the 50% frequency of λ_{IR} . Next three filters, red, green, and blue for the Bayer CFA, are chosen. They are shown in Figure 21 along with the effective spectral sensitivities based on the color filters multiplying the QE in Figure 20. The spectral sensitivities in Figure 21 will provide the exposure values for the digital camera for any

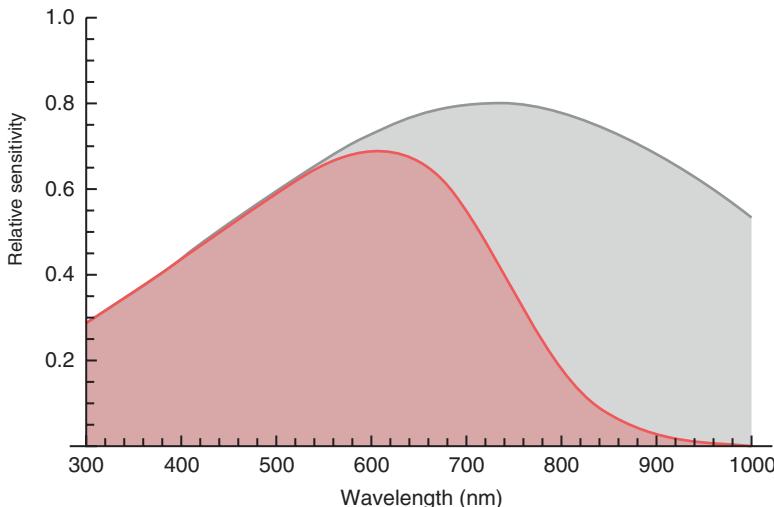


Figure 20 The impact of an IR filter on the quantum efficiency of a sensor. The red shaded area is the reduced QE curve, while the gray area is the native QE curve

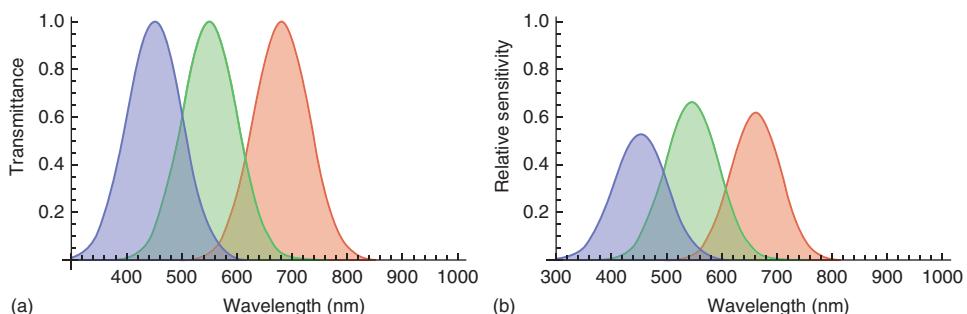


Figure 21 (a) Transmittance of the red, green, and blue filters used in the Bayer CFA. (b) Effective (relative) spectral sensitivity of the red, green, and blue channels when the QE curve in Figure 20 is used to represent the sensor

specific color spectrum. Assume a flat spectral input (neutral), using the spectral sensitivities in Figure 21, the relative red, green, and blue exposures is $\{0.8875, 1.0, 0.804\}$; the exposure is proportional to the area under the spectral sensitivity curves. Thus, the output will have a green tint to the color instead of a neutral or white.

The color imbalance (relative to a neutral) can be corrected by changing the color filters or later in the imaging chain using white balance methods in the camera (adjusting the relative red, green, and blue gains). White balance for different illuminants will be covered in its own section. For the sake of discussion, here the color filters will be altered so that the spectral sensitivities will produce a neutral for a flat spectral illuminant. The basic sensor QE will not be adjusted because that is a far more difficult process. Figure 22 shows the equivalent to

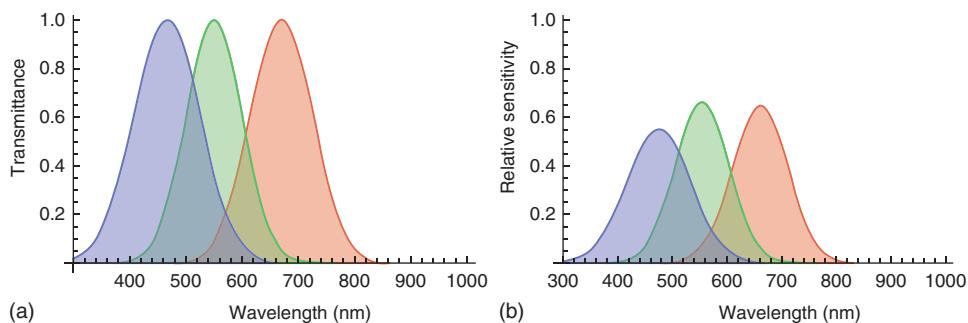


Figure 22 (a,b) Adjusted color filters and resulting spectral sensitivities to provide a neutral response to a flat spectral illuminant

Figure 21 with different color filters, all other factors being constant. Both red and blue filters were made wider and each shifted toward the green slightly to get the enhanced sensitivity.

Figure 23 shows two levels of gray (flat spectral illuminant at two lower levels). The original filters, Figure 21, give a clear green tint, whereas those in Figure 22 demonstrate a good neutral reproduction.

The neutral patch correction of the net spectral sensitivities ensures a good neutral (which is requirement for any good color reproduction scheme). The next step is to demonstrate the overall color correction over a wide range of colors. A test chart made up of 16 overlapping spectra is created using simple Gaussian spectra ranging from 400 to 700 nm, 20 nm apart and each Gaussian has a sigma of 40 nm. There is very little energy in each of these spectra and

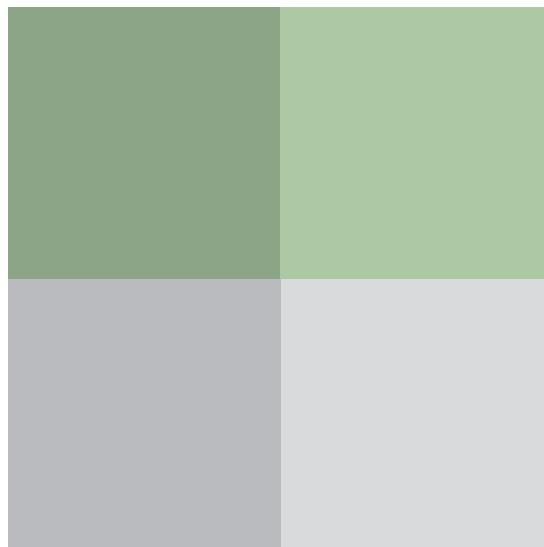


Figure 23 A neutral exposure produces green patches with the filters and spectral sensitivities in Figure 21, while the adjusted filters and sensitivities shown in Figure 22 shows good neutrals

the resulting image will be dark unless enough exposure time is assumed for each spectrum; the images shown below have exposure times adjusted to give a “bright” image. Figure 24 shows the spectra in the center panel. Each of these spectra is then captured by the spectral sensitivities shown in Figures 21 and 22. The resulting red, green, and blue exposures are then normalized to the maximum value of red, green or blue in each exposure output. They are then displayed as a 4×4 color chart in Figure 24; the original spectral sensitivities on the left and the adjusted (for a neutral) spectral sensitivities on the right. It is easy to see the color shifts. The saturated blues and reds look to be the same (even if the exposure values differ a small amount), but the cyan to orange colors clearly reproduced differently. The previous work indicated that the original spectral sensitivities produced a green shift from a neutral. This would imply that the colors would all move to the green with the original spectral sensitivities. Note that the cyan patch, second from the left in the second row, occurs earlier than before with the adjusted spectral sensitivities. Also note that in the left color chart, there are more greens than on the right, and that the yellows and oranges are very different. As the neutrals are better (exact) with the adjusted spectral sensitivities, it is safe to assume that the overall color reproduction is better. However, one cannot assume that it is accurate, but does it really have to be accurate, as long as the neutrals (and flesh tones) are “good?” The question will be addressed later. Note that the color patches in Figure 24 have been adjusted for maximum brightness at constant hue to clearly show the color differences.

The above color reproduction does not consider if the color rendering (how the colors look) is accurate in terms how the CIE standard observer would see the same spectra. Before this question is addressed, take into consideration that the rendering of the colors depends on how the tools used to create these color patches (Mathematica), the monitor calibration (Macintosh 27 inch Cinema monitor using sRGB profile), and finally the printing process effect the final rendering. So what one sees here is a function of many steps and while color differences can be seen between different spectral sensitivities, there is no absolute color calibration and there is no assurance that the specified spectral looks like the color a CIE standard observer would see. One theoretical measure of how well a digital camera’s spectral sensitivities record color relative to the CIE color matching functions is the Luther Criteria [25]. The Luther Criteria states that if the spectral sensitivities of a digital camera (film) are a linear combination of the CIE color matching functions, then the camera will record color the way the CIE standard observer would “see” the color. However, recall that if there was an exact match to the CIE color matching functions, the color reproduction would have to be done with the three primary illuminants that define the \bar{x} , \bar{y} , \bar{z} functions and that they are a linear combination of the original primaries, 700, 546.1, and 435.8 nm, used to obtain the \bar{r} , \bar{g} , \bar{b} CIE color matching functions. Further recall that the linear combination of the original primaries includes negative terms, meaning that there is no physical way one can reproduce all colors with either the original primaries or the linear combination of them to form the x - y - z color matching system. Hence, it is not possible to have a perfect CIE colorimetric match to all spectra even when using spectral sensitivities that meet the Luther Criteria.

There are several ways to test the Luther Criteria. One is to do a linear regression (least square fit) between the three camera spectral sensitivities and the x - y - z CIE color matching functions. This will result in a matrix that in the least square sense will transform the camera spectral sensitivities to the best fit x - y - z color matching functions. Figure 25 shows such an example. Here, three Gaussian camera spectral sensitivities have been chosen for the camera response curves and they are shown with x - y - z CIE color matching functions. By means of least square

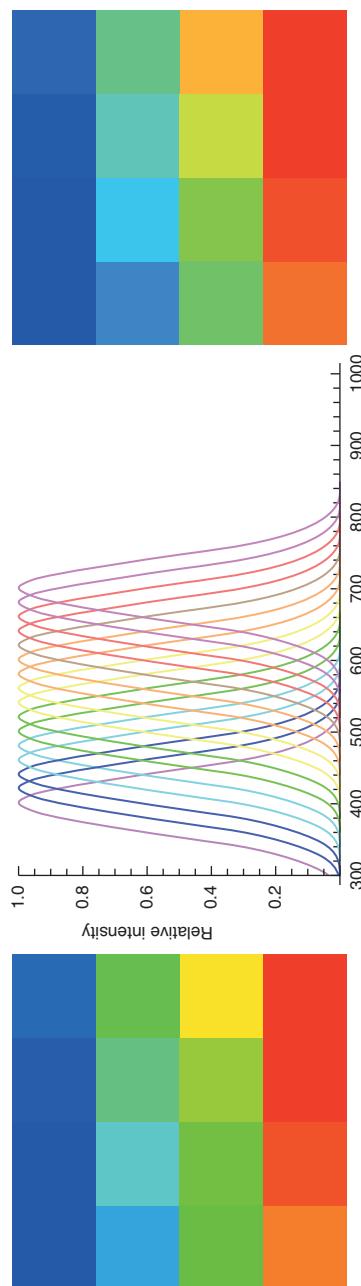


Figure 24 The color spectra used to define a color chart is shown in the center. The color reproduction for the spectral sensitivities in Figure 21 is shown on the left and that for the spectral sensitivities in Figure 22 is shown on the right. The x -axis is measured in nanometers

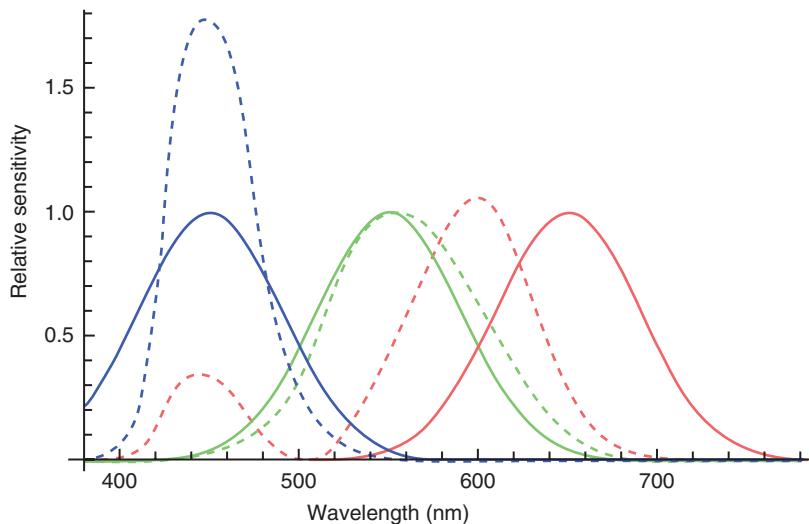


Figure 25 An example of camera spectral sensitivities (solid curves) as compared to the x - y - z CIE color matching functions (dashed curves)

linear regression a transform matrix was calculated to convert from the camera sensitivities to the best fit of the x - y - z CIE color matching functions. This conversion is shown in Figure 26. It is clear, that these particular camera sensitivities are a poor match for the x - y - z CIE color matching functions and thus a poor example of the Luther Criteria. One can easily shift the red camera spectral sensitivity to the left (peak at 600 nm) to better the red color matching curve and repeat the least square linear regression. The results are shown in Figure 27. Although

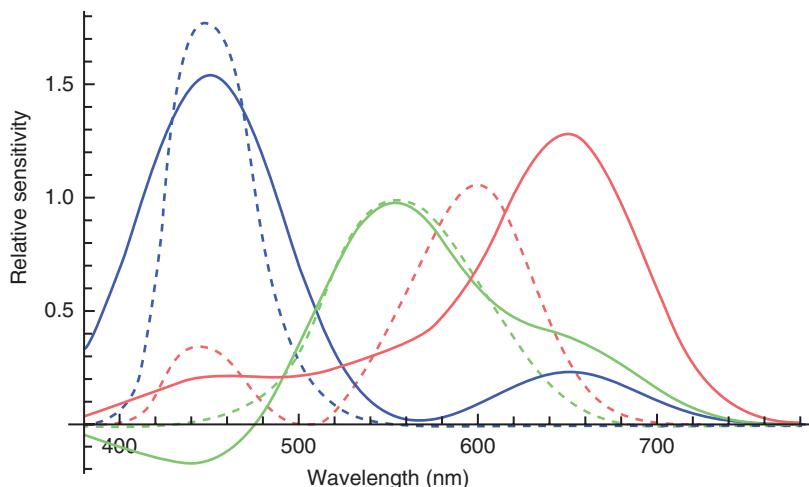


Figure 26 A comparison of the best fit (least squares regression) between the x - y - z CIE color matching functions and the set of camera spectral sensitivities shown in Figure 25 (solid curves) and the actual x - y - z color matching functions (dashed curves)

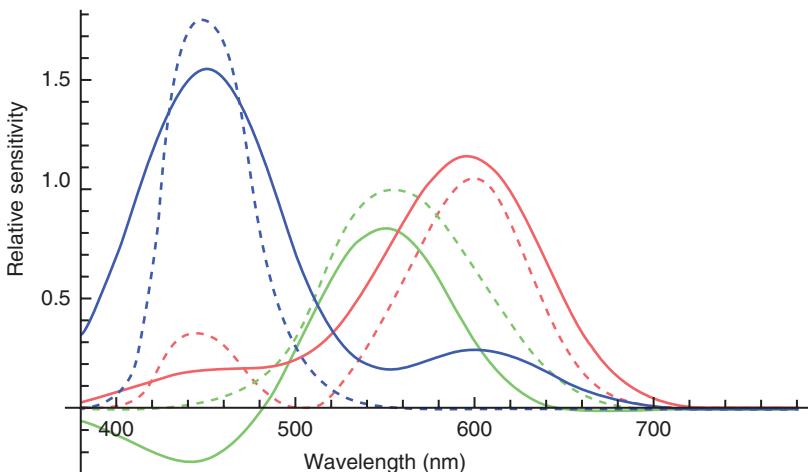


Figure 27 Like Figure 26 but the camera red spectral sensitivity in Figure 24 was moved so that its peak was at 600 nm. This results in a better match to the x - y - z CIE color matching functions, but far from matching the Luther Criteria

there is a better match, there are still significant problems with negative green values. This process can be repeated to find some optimum results, but in the end, it is the combination of color filters in the CFA and the native spectral sensitivity of the sensor that will determine how a given digital cameras spectral sensitivities matches the Luther Criteria.

Now consider the case where the x - y - z CIE color matching functions are used as the spectral sensitivities for a digital camera and one exposes the colors shown in Figure 24; now the Luther Criteria is met exactly. The resulting exposures are first transformed to the x - y - z values by means of Equations 13 through 15. This expresses the red, green, and blue color values between {0, 1} and then expressed in Mathematica. The results are shown in Figure 28 along

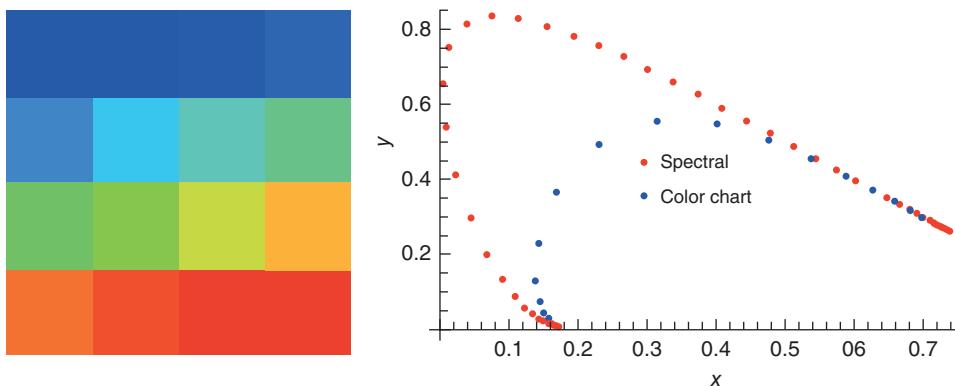


Figure 28 The color chart defined by the spectra in Figure 24 captured by a digital camera with spectral sensitivities equal to the x - y - z CIE color matching functions. The image on the left is a plot of the color chart in the x - y chromaticity chart

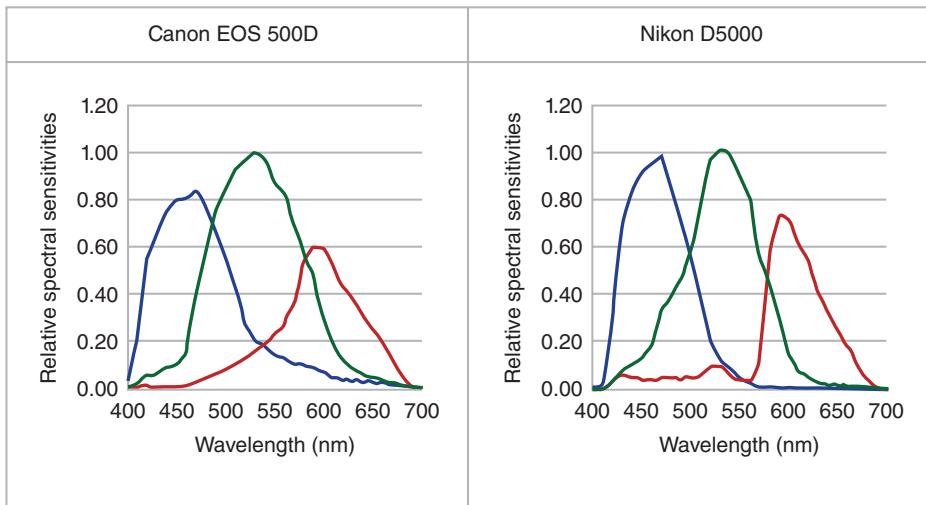


Figure 29 Relative spectral sensitivity for two consumer digital cameras

with their plot in the CIE x-y chromaticity plot; the color chart is based on sufficient exposure to give a bright image. The color chart colors follow the chromaticity envelope in the red region of the visible spectra, but as the colors move to the blue-green region they fall far short of the spectral envelope and will appear less saturated. If the color spectral shown in Figure 24 broaden, the color chart colors will move further from the spectral envelope. In the same manner, if the colors are defined by a narrower spectral, the chart colors will move closer to the spectral envelope.

Figure 29 shows the published relative spectral sensitivity functions for two popular consumer cameras, a Nikon D5000 and A Canon EOS 500D [26]. Both cameras have peak sensitivities near 450, 530, and 600 nm, but their respective spectral sensitivity overlap (red to green to blue) is very different, so to get a good neutral exposure for a flat spectral illuminant will require different amounts of relative red, green, and blue amplifications (after the analog-to-digital conversion). The literature shows a wide range of spectral sensitivities for consumer, professional, and scientific digital cameras [27–29]. In each case, once the sensor design and manufacturing process is set, the camera's overall exposure and ISO speed settings are adjusted by the embedded hardware to calibrated for a good white (under some assumed illuminant) and this becomes the default operating point for the camera.

4 White Balance

The human visual system has the ability to adapt to any illuminant and to preserve a white and the visual appearance of colors [5]. It is common to enter a building from the outside and note that the interior illumination is yellow (incandescent lights) or even a little green (fluorescent lights) for a short time, but eventually the visual system adjusts to the artificial illumination and the light sources appear white and colors seen to be correct. Film and digital imaging systems

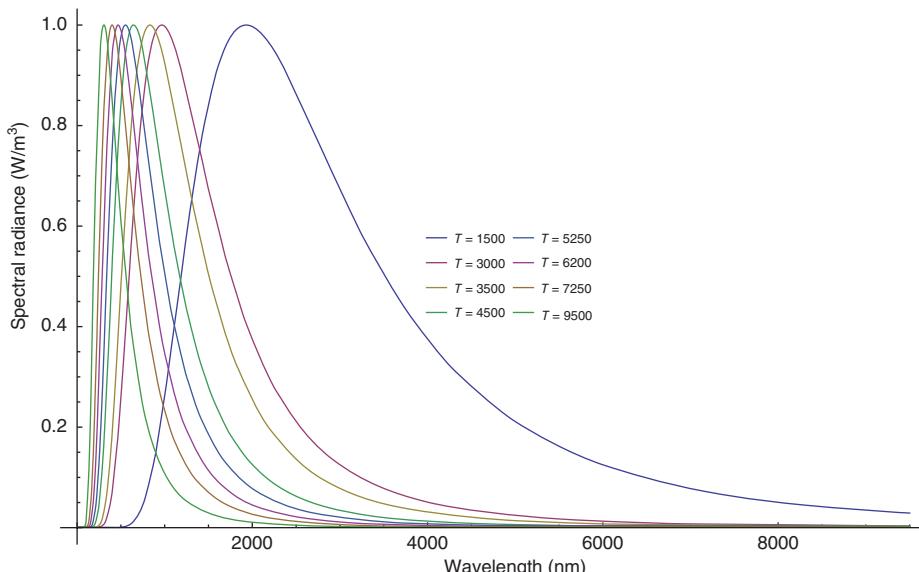
do not, in general, adjust to illuminant changes. Color film was provided, in most cases, in two color balances; a daylight balance for a color temperature of about 6200 K and a tungsten balance with a color temperature of about 2200 K. The daylight balance had equivalent ISO (ASA, American Standard Association) speed ratings for the three color channels, but the tungsten balanced film had a much faster blue layer to compensate for the lack of blue light in tungsten illumination. Color filters were used to balance a film, but always at a loss of speed. So if you were using an indoor movie film balanced for tungsten light, when one photographed outside in the bright sunlight, one would use a yellow filter to slow down the blue layer. On the other hand, if one used daylight balanced film to “shoot” indoors, one would use a blue filter for tungsten light or a magenta filter for green looking fluorescent lighting. Digital cameras could use the same set of filters as film cameras if the default exposure setting was for daylight exposure. However, digital cameras have built in abilities to apply different electrical gain to each of the red, green, and blue channels. In most digital cameras, the green gain is fixed for a given ISO speed setting. The red and blue channels have their respective gains adjusted (by some white balance algorithm or fixed settings for known illuminants) to give a good neutral for a white or gray. Digital cameras have automatic adjustments based on some form of analysis of the scene content, which will be discussed later. They also have a series for fixed adjustments that can be set by picking to scene illuminant from a table in the menu (such as bright day, overcast day, tungsten, fluorescent, sunrise, and deep shadows) or by picking a color temperature that defines the illuminant (from 1500 K for a candle to 9500 K for deep shadows). Some cameras will allow the user to take a picture of a gray target (neutral target) under the illumination in question, and the embedded software in the camera will calculate the differential gains for the blue and red relative to the green for a good neutral exposure. In what follows a simple color exposure model will be used to show how the fixed adjustments can be calculated and then the question of automatic white balance algorithms will be discussed. Figure 29 shows the results from a digital camera of a portion of the Macbeth ColorChecker under different illuminants (mostly compact fluorescent, CF, lights of specified color temperatures) taken under automatic white balance, except for the far right segment that was taken with the camera set to a tungsten taking illuminant. Although the visual differences are apparent it is best to look at the red-green-blue values (0–255) of the central neutral patch on the bottom. The flash image, on the right, gives values of {168, 164, 159}, which is slightly warm. The CF-warm, second from the right, has values {169, 161, 151}, which is still warm. The center patch was taken under CF-daylight and has values {176, 177, 176}, which is a nearly perfect neutral. The fourth from the right is taken with a CF-tungsten and has values {177, 177, 163}, which is slightly yellow. The auto white balance has done a fairly good job but the color differences are visible. The far left image shows the impact of picking the wrong white balance illuminant. Here the illuminant was CF-daylight and the camera was set to tungsten, which means a lot of blue gain was given to the image. The values for the gray patch are {123, 185, 232}, a very strong blue to blue-cyan. With this example in mind, consider the following simple model to establish the fixed differential gains for different color temperatures.

It is possible to establish a reasonable relationship between black body radiators at a given color temperature, degrees Kelvin, to most light sources and their spectra [3]. Table 1 shows such a relationship. Figure 30 shows the normalized spectrum for each of these blackbody illuminants; each is normalized to its peak value.

A simple set of balanced Gaussians (to a flat spectra illuminant) is chosen for the camera spectral sensitivities and are shown in Figure 31. When exposed to a unit flat spectra, the each

Table 1 Color temperatures for different illuminants

Type of Illuminant	Color Temperature (K)
Candle light	1500
Tungsten lights	3000
Sunrise or sunset	3500
Fluorescent lights	4500
Flash	5250
Daylight	6200
Moderate overcast	7250
Shade	9500

**Figure 30** Blackbody illuminants normalized to their peak values

give an exposure value of 100 (arbitrary units) and form a perfect neutral. The next step is to expose these camera spectral sensitivities to each of the normalized blackbody spectra shown in Figure 30 and calculate the exposure values. As most digital cameras adjust the red and blue gains relative to the green, the red, and the blue spectral sensitivities will be adjusted to ensure a neutral exposure. The relative gains are found in Table 2. Figure 32 shows the impact on the neutral if no exposure compensation is made for a white balance. It is clear that the flat spectra balanced spectral sensitivities reproduce a “neutral” that has the color of the illuminant. The first color patch is “white” and represents a flat spectrum, and as the color temperatures increase, the reproduced color moves from red to blue. Table 1 corresponds to the colors as one moves from left to right. The high blue gain required for the first three illuminants can introduce large amount of noise into the blue channel and will require more noise reduction

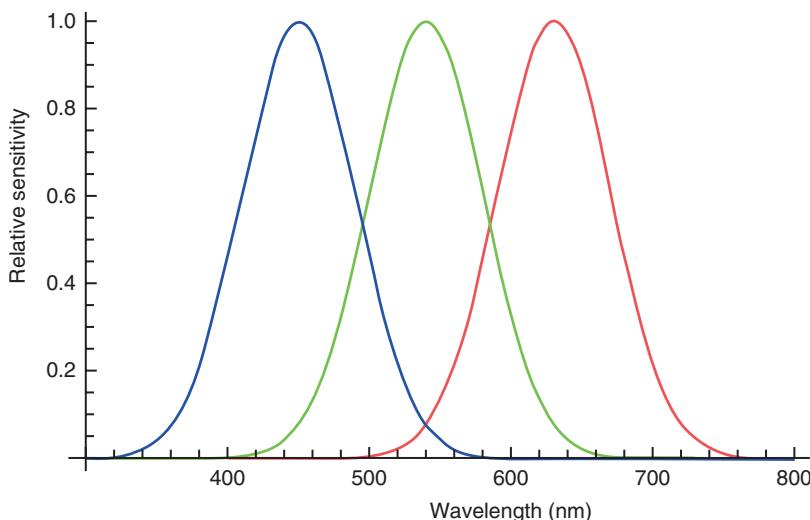


Figure 31 Model spectral sensitivities chose for a digital camera to study white balance

Table 2 Relative gains required for the red and blue channels to maintain a neutral when the illuminants shown in Figure 30 are used

Illuminant	Red Relative Gain	Fixed Green Gain	Relative Blue Gain
Candle light	0.20	1.0	9.3
Tungsten	0.62	1.0	2.26
Sunrise/sunset	0.73	1.0	1.80
Fluorescent light	0.92	1.0	1.32
Flash	1.03	1.0	1.12
Daylight	1.14	1.0	0.97
Modest overcast	1.25	1.0	0.86
Shade	1.40	1.0	0.73

by means of image processing and might introduce additional loss of detail information in the blue channel. The flash and the daylight blackbody illuminants come closest to a pure neutral, one being a little yellow and one being slightly blue.

Using the red and blue relative gain settings, the internal embedded software can pick these differential gains when either the illuminant name or color temperature is selected in the menu of the camera. However, if one neglects to change the setting and tries to take a picture under a different illuminant, then results like that shown in the last patch of Figure 33 can result. Figure 34 shows a three of the adjusted spectral sensitivities required to maintain a neutral with specified illuminant. It should be re-stated that the camera spectral sensitivities remain fixed and that the relative gains, Table 2, are used in the camera to achieve the effect shown in Figure 34.



Figure 32 The neutral color shift before compensation when different blackbody illuminants are image with the camera spectral sensitivities shown in Figure 31



Figure 33 Portion of the Macbeth ColorChecker under different illuminants taken with a Nikon Coolpix P7700 digital camera. From left to right: flash, compact fluorescent balanced for warm, compact fluorescent balanced for daylight, compact fluorescent balanced for tungsten, and compact fluorescent balanced for daylight but camera set to tungsten

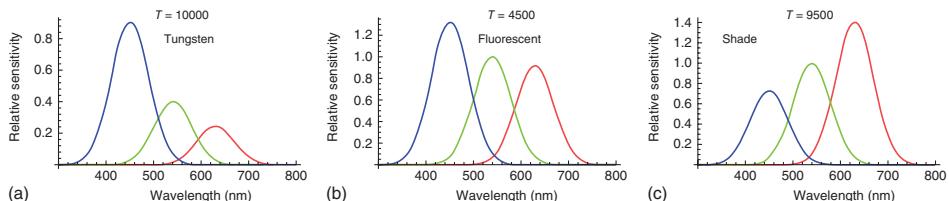


Figure 34 Spectral sensitivities in Figure 31 adjusted to ensure a neutral with the illuminants listed in the figure. Moving from tungsten to shade (a), the relative amplification of the red (b), and blue change as the green gain stays constant (c)

Most digital cameras use some form(s) of automatic white balance [30–32]. The sequence is as follows. When the user pushes the “shutter release button” (SLR) on the camera half-way-down, the image focus and exposure operations are actuated. In non-SLR digital cameras the image sensor (or some small part of it) is used to find the right focus, overall exposure and white balance. In an SLR digital camera a smaller, but separate sensor in the camera accomplishes this. By the time the picture is taken the best focus is obtained, the overall exposure has been translated into exposure time and aperture setting (f-number) and the required gain factors have been sent to the white balance module and the overall gain (ISO setting) amplifier. Figure 35 shows a conceptual drawing of this operation. The image sensor records a mosaicked color image based on the CFA array used to encode the color. Here the Bayer CFA is assumed. The data stream from the sensor has a well-defined arrangement of red, green, and blue pixel values, which are first “scrubbed” of some noise, associated with the readout electronics of the sensor. This noise removal is called *correlated double*

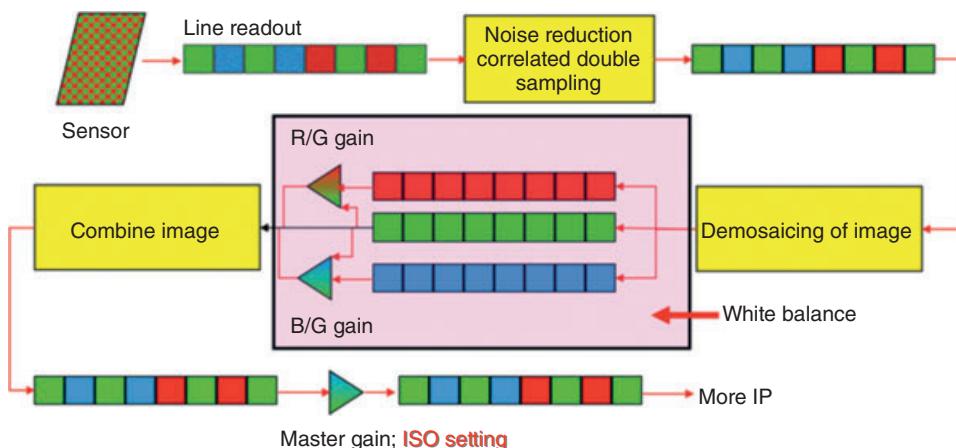


Figure 35 A conceptual flow diagram for white balance

sampling and will be described in the chapter on Digital Imaging, **Digital Imaging: An Introduction to Image Processing**. The noise-reduced data stream is then broken up into separate red, green, and blue channels by a demosaicing module that uses the master clock of the sensor to pick out the correct colors. The white balance module then uses the differential red-to-green and blue-to-green gains to adjust for a white balance. These differential gains either come from fixed setting as described earlier or from some algorithms to be described later. Once the three signal channels have been adjusted for a good white balance they are then combined to a single signal channel and amplified by the master gain that has been set by the ISO value chosen by the user or by some other algorithm in an automatic mode. The signal is now operated on other embedded modules to create the full red, green, and blue image layers (more on that below), more noise reduction, more color enhancement (more below), sharpness enhancement, and then compression and storage.

Automatic white balance is designed such that the user need not worry about the predominant taking illuminant. The embedded software and hardware in the digital camera will analyze a sampled portion of the scene while the automatic focus and overall exposure operations are taking place. There are several approaches to automatic white balance. The first, and easiest to understand and implement is the gray world concept [33]. The basic assumption is that on the average, an image will average to a gray. If one calculates the average red, green, and blue values of an image they should have the same value (for a given overall exposure). If they do not, then one only adjusts the red and blue average values so they match the green value, and then a good white balance should be obtained. The overall exposure can then be adjusted to get the best image as shown in Figure 35. There will always be exceptions to the gray world assumption, particularly when a large object of a given color is imaged or when one is taking pictures at sunrise or sunset where there is always a red-yellow-pink cast to the image. The other methods are designed to determine what the color temperature of the scene illuminant “might” be from the image data itself. One method looks for the highest luminance value in the image (but not saturated) and looks at the relative red-green-blue values, and by means of calibration of the camera in question, it can identify which illuminant defined by the list

given in Table 1 is illuminating the scene and then the appropriate set of red and blue gains are used as defined in Table 2 [34, 35]. Specular highlights or images of the illuminant itself are good sources of this type of data. One can also try to correlate the colors within the image and determine if the correlation is representative of a given illuminant and thus establish the relative red and blue gains. The same can be done with the actual gamut of colors in any given scene. An overall shift in some color direction can be used to determine the apparent illuminant. In each case, except for the gray world concept, one is trying to find the apparent illuminant and then finding the most likely illuminant color temperature and then altering the relative red and blue gains in the camera to get a good white balance and a natural looking image.

Figure 36 shows five images of the Macbeth ColorChecker taken by a Nikon Coolpix 7700 digital camera under a fixed lighting condition, but at five different white balance settings.



Figure 36 The Macbeth ColorChecker taken under CF tungsten illumination using, from top to bottom and left to right, auto white balance, flash, tungsten white balance, fluorescent white balance, and daylight white balance

The illumination is a CF fixture that gives off a warm tungsten light. It was photographed in the auto white balance mode, by the camera flash unit, the tungsten white balance setting, the fluorescent white balance setting, and the daylight white balance setting. Using Photoshop, the average red, green, and blue values for each image were obtained and they are given in Table 3. It is clear from Figure 36 and Table 3 that the flash mode provides the best exposure for an overall neutral image; it is darker than the others, but this is easily corrected when making a print or viewing on a monitor. The automatic setting resulted in an image that is warm relative to a neutral. The tungsten setting results in some improvement, but it is still too warm. The fluorescent setting, which tries to correct for excess green light, adds more red light and is thus very warm. The daylight setting expects strong blue light so it increased the green and red to the point where the image is very warm and yellow. Other than the flash, none of the white balances produced a good neutral looking image.

It is instructive to apply the gray world criteria, in two ways, to these incorrectly adjusted images to see if we can correct for the biased exposures. Two Mathematica programs were written to test this concept. First, a 1000×750 version of the images in Figure 36 was used to calculate the needed offsets in the red and blue to match the green average value. First the image is separated into its red, green, and blue layers. From the data and histograms of each layer, the average values were determined for each color and the offsets calculated. The offsets were then added to the red and blue images and the layers were then recombined to form the color image. When this is done to all the images in Table 3 and Figure 36, using the Mathematica program, the resulting red and blue offsets almost completely agree with the offsets obtained from Photoshop. Figure 37 shows the changes for the automatic white balance, the tungsten white balance, the fluorescent white balance, and daylight white balance images in Figure 36. The flash image is not shown as there was very little change.

Only the daylight white balance for the CF tungsten image does not correct (all most) fully for the exposure errors. The reason for this is that the offsets were so large relative to the mean image value (in the red) that when the offset was applied the red values went negative and had to be clipped. Hence, it is clear that there are limits for this version of the gray world correction even when the scene (in this case the Macbeth ColorChecker) does have an average neutral scene content. As can be seen in Figure 37, the gray world correction is very reasonable.

The second gray world algorithm follows the one above, but instead of a set of red and blue offsets it calculates the differential gains between the red and blue to equal the fixed gain in the green channels. These gain values are shown in Table 3 within the square brackets, [].

Table 3 The average red, green, and blue values (0–255) for the Macbeth ColorChecker under a tungsten-like CF lamp at various white balance settings

White Balance	Red	Green	Blue
Automatic	124 (-20) [0.836]	104	75 (28) [1.374]
Flash	75 (-4) [0.945]	71	72 (-1) [0.985]
Tungsten	114 (-8) [0.928]	105	82 (23) [1.280]
Fluorescent	150 (-52) [0.649]	97	76 (21) [1.284]
Daylight	149 (-55) [0.659]	98	43 (51) [2.284]

The values in parentheses are the offset calculated to make the image, on average, neutral. The terms in the square brackets, [], are differential gains for the red and blue channels to create a neutral.



Figure 37 The Macbeth ColorChecker corrected using a gray world algorithm when illuminated by a CF lamp with tungsten coloration using, from left to right and upper to lower, an automatic white balance, white balance set to tungsten, white balance set to fluorescent, and white balance set to daylight

Figure 38 shows the results. The use of differential gains in the red and blue channels based on a gray world concept does a better job than using the red and blue offsets as shown in Figure 36. The results in Figure 38 are a little dark, but this can be automatically adjusted using the master gains as indicated in Figure 35.

Most advanced cameras also allow the user to calibrate to a given illuminant by first imaging a neutral patch, which then provides the red, green, and blue gray world offsets directly. These offsets are then used to set the red and blue differential gains as shown in Figure 35. Also, advanced digital cameras provide RAW files, which just record the raw data with no processing. Advanced software packages allow the user use many image enhancement programs including ones that allow one to correct for the color temperature of the taking illuminant.

In an automatic white balance, one attempts to calculate from the image (during the auto-focus and exposure control step) what the most likely illuminant that is present in the scene. The details will be considered later. In general some measure of the color temperature must be obtained from the red, green, and blue statistics of the image data or from some conversion of the data to some other color space such as $Y\text{C}_b\text{C}_r$ used for compression in JPEG or JPEG2000. Figure 39 shows three plots of where various areas of three color spaces correspond to various illuminants. The first color space is the CIE x - y plot and indicates where the color of a neutral might fall if under some illuminant. This information is then used to set the differential gains in the red and blue channels to obtain a good neutral. In any given digital camera this would not be possible, as the spectral sensitivities of the camera are not, in

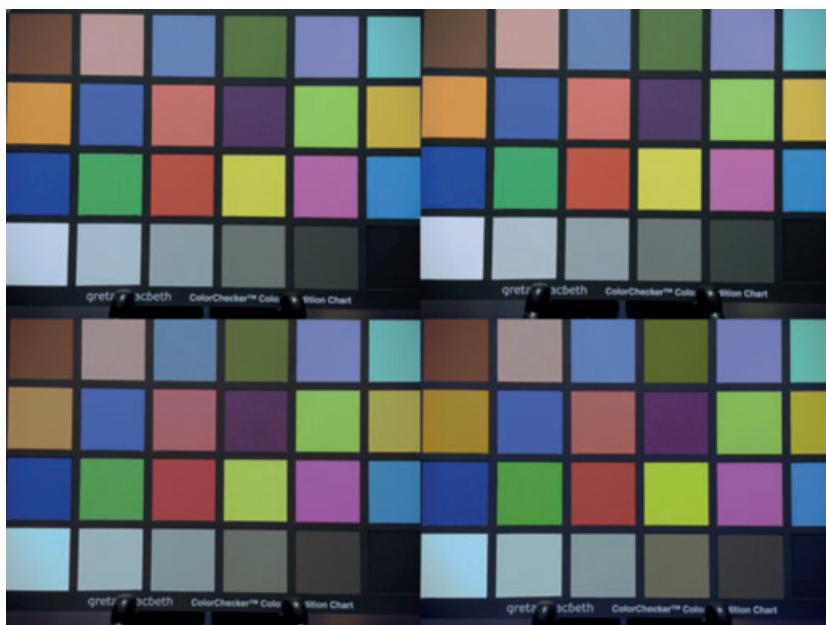


Figure 38 The Macbeth ColorChecker corrected using differential gains in the red and blue channels. From left to right and top to bottom are the corrected images for the automatic white balance, the tungsten setting for white balance, the fluorescent setting for white balance, and the daylight setting for white balance

general, close to the CIE $\bar{x}\bar{y}\bar{z}$ color matching functions. Instead the spectral sensitivities like those shown in Figure 31 must be used to find the equivalent camera red-green-blue chromaticity values and the YC_bC_r values. The second plot in Figure 39 shows the red-green-blue chromaticity plots of the illuminants shown in Figure 30 projected on to the [1, 1, 1] of the red, green, and blue responses of the digital camera. The third plot in Figure 39 shows the C_b and C_r chromaticity values of the illuminants where the offset value of 128 has not been used, resulting in some negative values. Table 4 shows the values. If one is able to extract either the red-green camera chromaticity values or the $C_b - C_r$ from the image data and compared then to the data in either chromaticity plot, then the closest illuminant can be found and the differential red and blue gains can be set.

In what follows the YC_bC_r transform and its inverse are used.

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ -0.169 & -0.331 & 0.5 \\ 0.5 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (22)$$

and

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.402 \\ 1 & -0.344 & -0.714 \\ 1 & 1.772 & 0 \end{bmatrix} \begin{bmatrix} Y \\ C_b - 128 \\ C_r - 128 \end{bmatrix}. \quad (23)$$

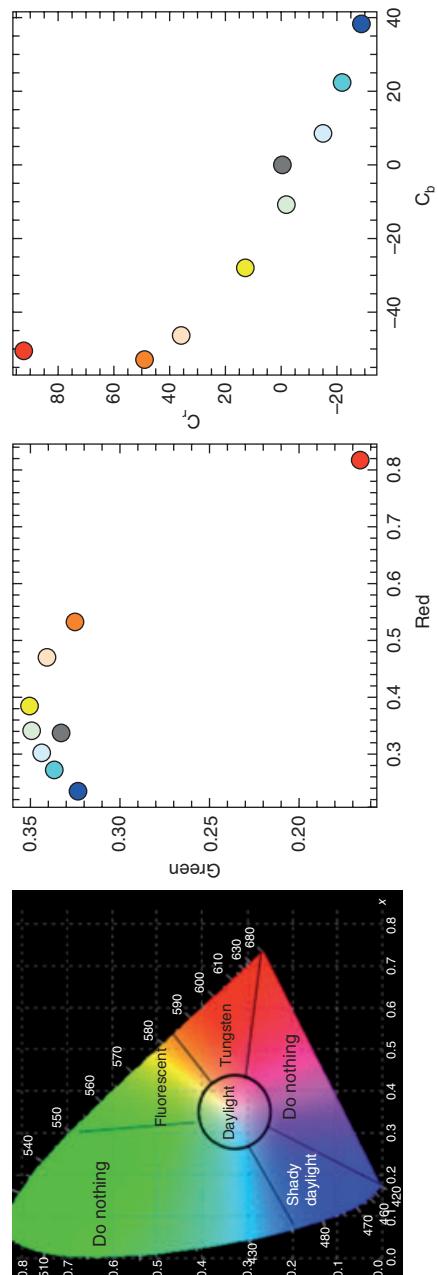


Figure 39 From left to right plots of the spectral region of various illuminants in the CIE x - y chromaticity plot, the camera r-g chromaticity plots of the blackbody illuminants in Figure 30 and the same data in a YCC chromaticity plot. The color of the dot approximates the color temperature of the illuminant. *Source:* Reproduced from Illustration of the CIE 1931 color space, as published in Wikimedia Commons, 2007, <http://commons.wikimedia.org/wiki/File:CIEExy1931.svg> © Sakurambo, 2007, distributed under the Creative Commons Attribution-Share Alike 3.0 Unported license

Table 4 Chromaticity values for various blackbody illuminants

Color Temperature (K)	Camera Red-Green Chromaticity	Camera $C_r - C_b$ Chromaticity
1500	0.816	0.166
3000	0.530	0.326
3500	0.468	0.342
4500	0.382	0.351
5250	0.339	0.350
6200	0.300	0.344
7250	0.270	0.337
9500	0.232	0.324

In some cases, the 128 values can be dropped for ease of exposition as they have been added to ensure that there are no negative values for YC_rC_b . Equation 22 was used to transform the camera r-g chromaticity plot to the C_bC_r chromaticity plot in Figure 39.

Earlier we have considered two versions of gray world white balance and found that the gray world method that calculates the differential gain in the red and blue channel to be better than using the offset values in the red and blue channels based on obtaining an overall gray for the image. For automatic white balance it is necessary to analyze the image (or some smaller version or sample of the image) before the final exposure. Consider the following methods. Assume that an algorithm can be developed that finds the most likely neutral area in the image and uses that local red, green, and blue data to calculate the differential red and blue gains and uses it for the complete image. The same local data could be used to calculate the C_bC_r values for the illuminant and then use that data to calculate the nearest color temperature adjustment stored in the camera and apply it to the whole image. Equation 22 is used for this purpose. Another method would be converting the entire small or subsampled image in YC_rC_b values and create a C_bC_r histogram of all the pixels in the small or subsampled image. One then searches the histogram for peaks (not including a neutral peak) and weights each of them by their count in the histogram and then takes a weighed average to determine the apparent illuminant [36]. The threshold for counting a peak can be used to limit the “noise” in the calculation, however, the threshold cannot be too high that one misses significant data. Once the weighted average for C_bC_r has been found, one can use these values to match a known color temperature adjustment or use Equation 23 to calculate the equivalent red, green, and blue values of most likely illuminant; to do this the Y value used is the average Y value for the entire subsampled image. Once the red, green, and blue values are obtained they are used to calculate the differential gains in the red and blue channels and applied as shown in Figure 35. When exercising these different algorithms it is important to move from a Macbeth ColorChart to real, complex images. However, it is also prudent not to “image” the light sources directly, for they will bias the results and can introduce some unexpected color shifts.

Figure 40 shows an original image, upper left, taken with a Nikon D700 in the RAW mode under a room illuminated by indirect LED (light emitting diode) lighting rope, which is a little pink to the eye. This image is then processed in three different ways. The lower left shows the gray world white balance using differential red and blue gains. The middle lower image represents measuring a part of the wall in the original and then calculating the red and blue differential gains and applying them to the original. The lower right image shows the

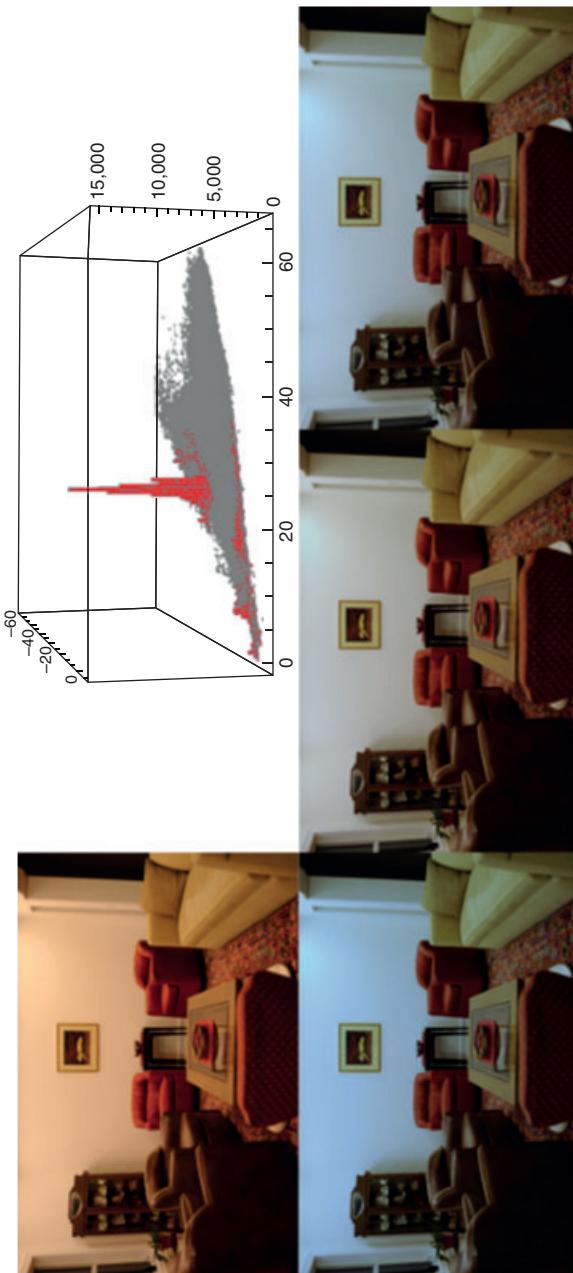


Figure 40 The upper left original is taken under a LED rope illumination and the histogram to the right shows the $C_b C_r$ distribution of chrominance. The lower left image is the effect of gray world white balance. The lower center image is white balance by taking a measure of the wall red, green, and blue values. The image on the lower right shows the effect of using the histogram to determine the apparent illuminant

original after the histogram in the upper right has been analyzed in C_bC_r space to find the most probable illuminant and then using Equation 23 to calculate the corresponding red, green, and blue values of the illuminant, finding the red and blue differential gains and applying them to the original. The original is clearly too warm. All three automatic white balance methods push the original toward a more neutral. In this particular case the white balance based on measuring the red, green, and blue values of the wall seems to do the best job. Although this may produce the best color balance, the algorithm to find the best place to measure in the image can be very difficult, while the other two methods are very simple to implement.

Figure 41 shows the results for the same scene but under mixed illumination. In addition to the LED rope lighting in the ceiling, several compact fluorescent lights were used, but none visible in the scene. The original image is much lighter and the histogram indicates that there are many more peaks, making it more difficult to pick a single illuminant. The lower left image is the simple gray world white balance, the middle lower image is the adjustment based on a measure of the red, green, and blue values on the wall (where no clipping had taken place) and the lower right image is based on the weighted average of the most significant peaks in the histogram shown on the upper right. Although there is clear improvement in the color balance in all three cases, the wall is clearly too cyan with the wall measurement-based algorithm being slightly better. Table 5 shows the red and blue differential gains for the images in Figures 40 and 41. Although there is some difference in the gains when one finds an appropriate area on the wall for the calculation, the results from using the gray world concepts or the histogram method are almost identical, thus giving very similar results. The issue of mixed illuminants is very real, even in outdoor scenes where light and dark shadows will cause any automatic white balance algorithm to give some undesirable results. If any one of the red, green, or blue channels is severely clipped owing to a low exposure, then the ability to correct for the illuminant will be limited.

To calibrate the above “models” for white balance, the Nikon D700 digital camera was used to image the same scene under the same LED rope light illumination. The settings for white balance used in the full resolution, fine image mode (rather than the RAW mode) were “automatic,” “incandescent,” “fluorescence,” and “full daylight.” Figure 42 shows the unedited results along with alterations using Photoshop to bring the images back to an optimum state (based on the author’s perceptions). It should be clear that the automatic white balance mode gives the best overall exposure for lightness, but is a little too warm. In Photoshop, a “cooling” filter was used to shift from the warm to neutral. The second image represents a white balance for an incandescent illuminant and has good to cool (blue) color balance. Using Photoshop the apparent exposure was increased to form the image in the lower set. The third image is based on a white balance for a fluorescent illuminant and has moved strongly to the magenta (absence of the green light in the LED illuminant). A green filter was used in Photoshop to bring the image back to neutral plus an increased exposure. The fourth image is due to a white balance for daylight images. As the LED illuminant has less blue light than daylight, the image is very warm or yellow. Again a cooling filter was used in Photoshop to bring the image back to a neutral balance. The use of the automatic mode gives an image that is too warm, while the methods used in above algorithms, as shown in Figure 40, tend to be too cool. Presetting the white balance by taking an image of a neutral surface, such as the wall, and calculating the differential red and blue gains may be the best method (as seen in the lower middle image of Figure 40) to obtain white balance, but it is not a consumer friendly solution for taking pictures with a digital camera.

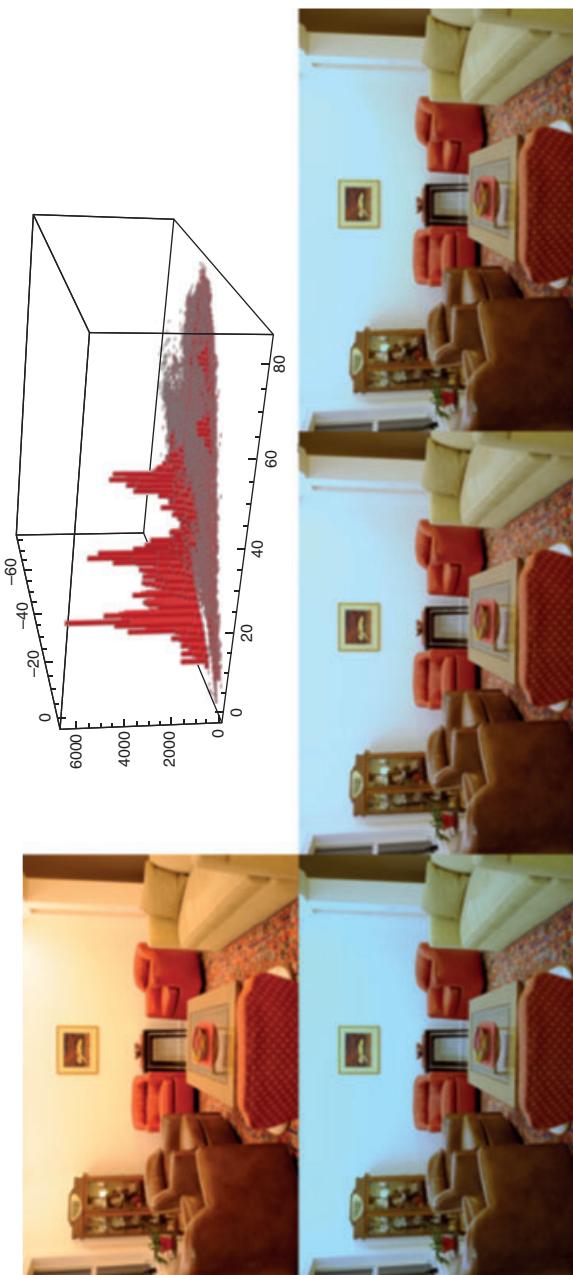


Figure 41 The upper left original is taken under a LED rope illumination with additional CF lighting. The histogram to the right shows the C_b , C_r distribution of chrominance. The lower left image is the effect of gray world white balance. The lower center image is white balance by taking a measure of the wall's red, green, and blue values. The image on the lower right shows the effect of using the histogram to determine the apparent illuminant

Table 5 The differential red and blue gains for the images in Figures 39 and 40, respectively

Lighting	Red-Blue Gains from Wall		Red-Blue Gains Gray World		Red-Blue Gains from Histogram	
	Red	Blue	Red	Blue	Red	Blue
LED light	0.79	1.31	0.66	1.58	0.72	1.44
LED and CF	0.77	1.47	0.68	1.58	0.76	1.43

With the onset of smart phones as well as DSCs, some form of automatic white balance is required. The resulting images may vary a great deal owing to how the user picks to focus the camera (not being aware of shadows, brightly colored backgrounds, location of lamps of all kind, etc.). The issue of mixed illuminants both outside and inside will continue to create a challenge for digital camera makers to incorporate the sophisticated segmentation and color constancy research that will enable them to correct for local color variations due to illuminants as well as the easier global changes demonstrated here.

5 Color Reproduction

The exact nature of color reproduction depends on the desired usage. The section will focus on color reproduction of digital camera images to be viewed by both professional and consumer users. The dilemma faced by all digital camera owners is how do I want my color rendered. For the sake of this discussion, it will be assumed that the images will be viewed on a well-calibrated LCD monitor so the monitor accurately displays the desired results. Some users might wish to have “exact color reproduction” such that the colors in the image looks exactly like the colors in the scene. Other might just want a very pleasant image that brings back the memories of the original scene. Yet others might wish to produce a vivid or overly saturated image, one that is more vibrant than the original scene, but still easily recognized by the user. Others might wish to mute the image and create a more pastel look. No matter the final rendering desired by the user, it is best to first accurately “capture” the image so that the colors are as close to the original colors as possible and then to use color manipulation techniques (in software) to achieve the desired goal.

To capture the most accurate colors possible, one might wish to satisfy the Luther condition discussed earlier using a combination of color filters and native spectral sensitivity of the sensor to form the color matching functions shown in Figure 8. To test this concept we use the Macbeth ColorChart shown in Figure 12. The spectral reflectance of each patch is well known and can be used to simulate the imaging of the Macbeth ColorChart. The resulting images will be, first, displayed on a monitor and then copied to a Tiff file and then place on hardcopy using some sort of inks or pigments. The final result will not show exact color reproduction, but will provide a basis for comparison for other spectral sensitivities. Keep in mind, the red, green, and blue exposure values obtained from the color matching functions are used to “drive” three color primaries, which in fact do not exist in nature because of their negative spectral values. Instead the conventional sRGB primaries will be used to display and print the images. The use of sRGB primaries and other primaries will be discussed later.

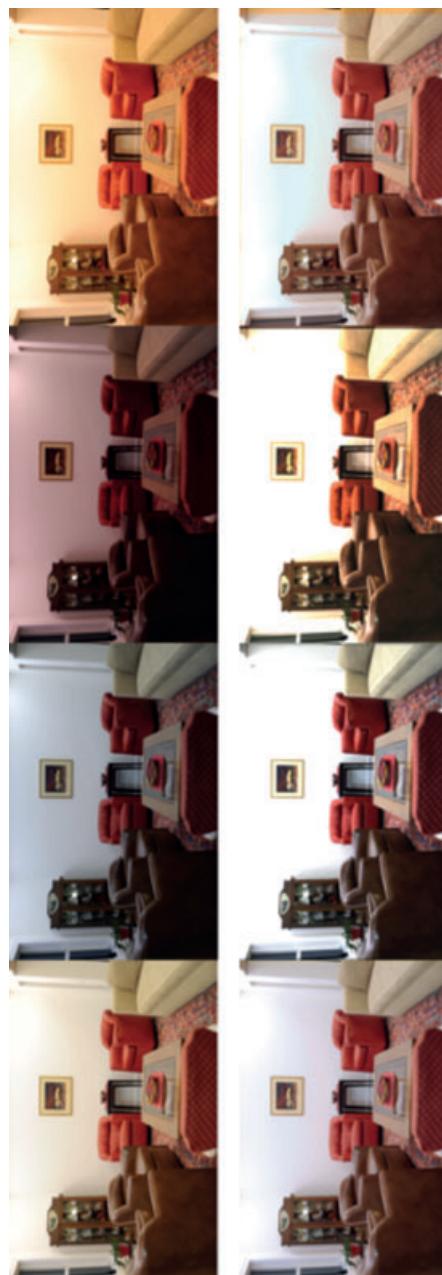


Figure 42 A series of images taken with a Nikon D700 digital camera. The white balance, from left to right, was set to automatic, incandescent, fluorescent, and daylight, respectively. The lower row shows the same images but corrected using Photoshop to give a more neutral rendering by the use of photographic filters



Figure 43 The Macbeth ColorChart imaged with the CIE color matching functions used as the spectral sensitivities of a digital camera

Figure 43 shows the actual colors generated by the spectra of the Macbeth ColorChart when using the CIE color matching functions as spectral sensitivities. The results are very poor in terms of color reproduction owing to the very large unwanted exposures in the red channel because of blue and green light, see Figure 8, along with the fact that the sRGB primaries used to render the image are far from the imaginary color primaries used to calculate the CIE color matching functions. The image in Figure 43 shows colors that are greatly muted and desaturated. The gray scale has been produced very well. Without a great amount of additional color processing, the CIE color matching functions do not make a suitable set of spectral sensitivities. Careful study of Figure 43 shows that while the color reproduction is poor, the CIE color matching functions clearly separate each color patch from the others. Hence, there is good color differentiation and with the complex processing of the human visual system.

A word is in order on how Figure 43 and those that follow are made in Mathematica. The spectral reflection of each patch is multiplied by the spectral sensitivities chosen for the red, green, and blue channels. As the chart was measured using a D50 standard illuminant (close to a black body spectrum of color temperature of 5250K), the spectral sensitivities are adjusted (electronic gain as seen in Figure 35) to create a neutral when illuminated by a D50 source. As it is required that the full chart be imaged within the dynamic range of the monitor (printer) the exposure of the chart is adjusted such that only the most saturated colors (the yellow and the white) will be clipped and not distort any other colors.

The systematic study of how spectral sensitivities effect color reproduction starts with defining a standard set spectral sensitivities. Figure 29 shows two sets of real spectral sensitivities and Figure 31 shows a set that will be used as the standard set for what follows. The Gaussian spectral sensitivities shown in Figure 31 will produce a clean neutral for a flat spectrum. However, the Macbeth ColorChart spectral reflectance data was obtained with a D50 standard illuminant. Hence, the first step is to calculate (measure) the red, green, and blue responses to the D50 illuminant and then adjust the red and blue differential gains (within the camera) to ensure a good neutral under the D50 illuminant. Figure 44 shows the original and adjusted spectral sensitivities along with the possible spectrum of colors in a camera r-g chromaticity

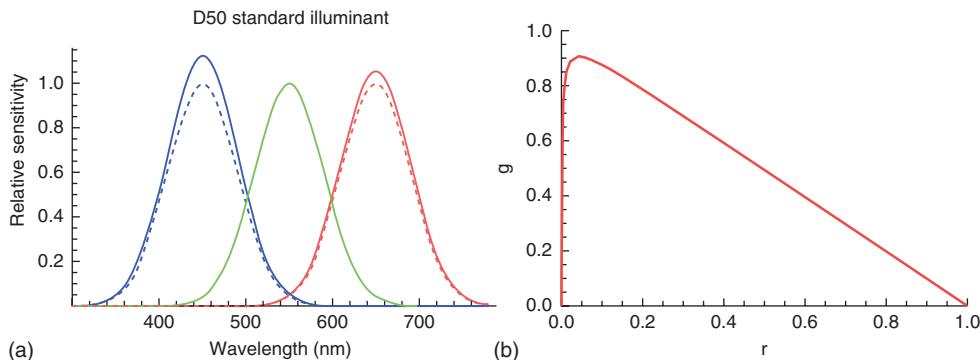


Figure 44 (a) The original spectral sensitivities for a flat spectral neutral (solid curves) and those adjusted for a neutral for a D50 standard illuminant. (b) The r-g chromaticity plot for the adjusted spectral sensitivities

plot such as the one used for the CIE x - y chromaticity plot. This is essentially the same operation defined by Equations 10 through 16, but with the CIE color matching functions replaced by the red, green, and blue spectral sensitivity functions. All colors within the red boundary are possible with the shown spectral sensitivities. The next step is imaging the spectral reflectance curves for the Macbeth ColorChart using the spectral sensitivities shown in Figure 44. The exposure is adjusted to ensure all colors are captured with minimum clipping. The image of the Macbeth ColorChart is now the starting point for future color reproduction operations.

Figure 45 shows the results when the spectral sensitivities shown in Figure 44 are used to image the Macbeth ColorChecker. Note that a good neutral is maintained. The plot on the right is the r-g chromaticity plot for Macbeth ColorChecker on the left. For future reference, this data will be called the *raw data* in which no further image processing has been done. Note the small cluster of dots near the $\{0.333, 0.333\}$ point in the r-g chromaticity plot; these are the neutral points and they have less scatter. The image in Figure 45 looks to be quite good for color, but there are some small variances from the real chart. The magenta (patch five in row three) is too light and too red and the dark cyan patch (next to the magenta patch) is too blue. Also, the dark

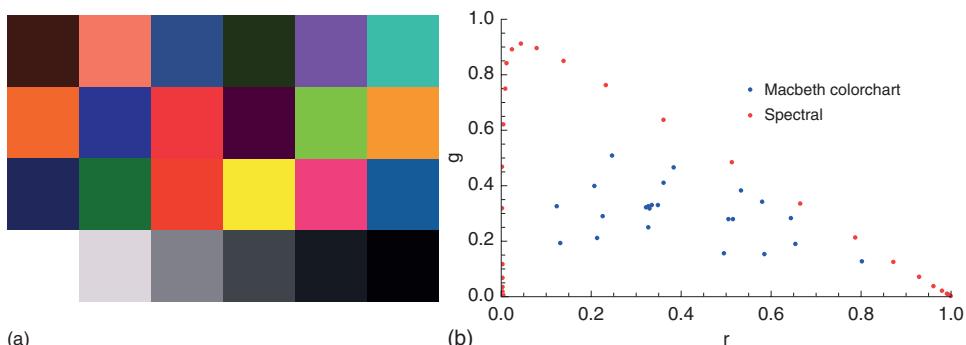


Figure 45 (a) The results of using the spectral sensitivities shown in Figure 43 to image the Macbeth ColorChecker under D50 illuminant. (b) The r-g chromaticity plot of the ColorChecker

green patch (fourth on the top row) is too dark as is the other greens in the chart. Some of these variances from the true colors are due to the overlapping of the spectral sensitivities. The green channel “sees” both red and blue light, the blue channel “sees” green light and the red channel “sees” some green light. These unwanted exposures can be approximated by calculating the overlapping areas of the spectral sensitivities by simply integrating the overlapping area (by taking the integral of the product of the any two adjacent spectral sensitivities) and dividing by the net exposure of the channel in question. Thus, one can easily adjust the exposures using a compensation 3×3 matrix that preserves a neutral and adds or subtracts the unwanted exposures. For the spectral sensitivities shown in Figure 44 the following compensation matrix has been determined;

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 1.15 & -0.15 & 0 \\ -0.16 & 1.33 & -0.17 \\ 0 & -0.15 & 1.15 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (24)$$

where $[R \ G \ B]$ represent the red, green, and blue exposures based on the spectral sensitivities and the chart spectral reflectance values. $[R' \ G' \ B']$ are the adjusted exposures. Note that the sum of each row in Equation 24 is unity; this ensures that neutrals are not altered (as the spectral sensitivities have been designed to produce a neutral for D50 illumination). If there is greater overlap in the spectral sensitivities, the terms in Equation 24 will be larger and if there is less overlap the terms will become smaller. If we apply the above exposure correction to the “raw” data a “cleaner” color image is obtained as shown in Figure 46.

As the actual image of the Macbeth ColorChart and the r-g chromaticity plot shows, the operation of the color exposure adjustment matrix given by Equation 24 has greatly improved the color reproduction of the patches. Note that the magenta patch has greatly improved as has the separation between the red and orange patches (third in on the second and third rows). The dark cyan has improved, but the hue is still off and the greens are a little lighter as they are on the actual Macbeth ColorChart. The r-g chromaticity charts show how the adjusted image, green dots, have moved outward from the raw data (blue dots), all moving closer to the spectral limit (red dots) (Figure 46).

Another simple way to improve color contrast (and saturation) is to apply a nonlinear color operation that preserves neutral and can partially make up for some of the unwanted exposures.

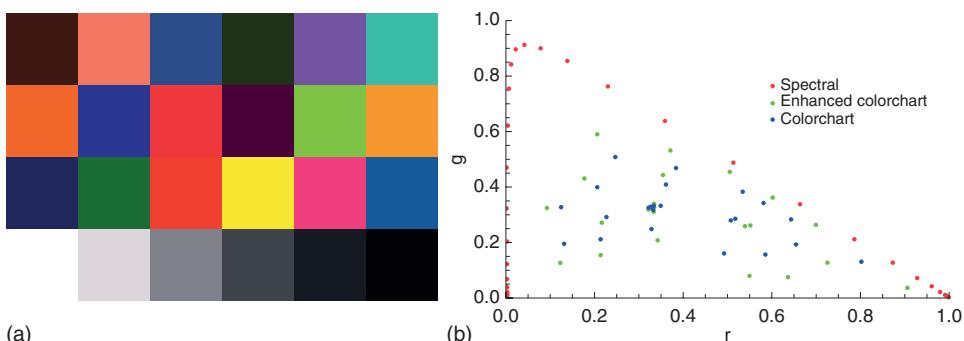


Figure 46 (a) The color chart show in Figure 45 after the exposure adjustment given by Equation 24 is applied. (b) The r-g chromaticity plot clearly shows the expanded color gamut and increased “purity” of the colors. Note that the neutrals were not affected by the operation

The operation is defined by:

$$R' = R[1 + a(R - G) + b(R - B)], \quad (25)$$

$$G' = G[1 + c(G - R) + d(G - B)], \quad (26)$$

$$B' = B[1 + e(B - R) + f(B - G)], \quad (27)$$

where the red, green, and blue values have the same meaning as above and the set of coefficients $\{a, b, c, d, e, f\}$ are picked to bring about a desired change in the color. The operations set forth by Equations 25–27 will not change a neutral and based on the differences between color channel values in any given patch will increase the color saturation in a quadratic manner for the dominant color while lowering the contrast of the less dominant color. If there is a small amount of unwanted color (say too much blue in a red patch), the above operation will reduce the unwanted color and increase the desired color. Figure 47 shows the impact of this operation on the raw camera image. In this case, the following values were used for the coefficients: $a = 0.024$, $b = 0.025$, $c = 0.05$, $d = 0.1$, $e = 0.05$, and $f = 0.05$. These are rather small values, but they do have a large impact as can be seen later.

Although the nonlinear color enhancement improves the overall color quality of the Macbeth ColorChart, its impact is a little less than that of the exposure adjustment method. The obvious next step is first to apply the exposure adjustment method and then apply the nonlinear color operation. The results are shown in Figure 48. The color patches in the chart are now very saturated and for the most part closer to the original Macbeth ColorChart. Some of the patches may be too saturated and have moved closer together in appearance. Hence, care must be taken to optimize this process. The r-g chromaticity plot in Figure 48 clearly shows the expanding gamut with each of the above operations. The second neutral patch from the left shows a little coloration, which indicates there were some small differences in the original red, green, and blue values, and that the significant processing has slightly affected the neutral.

Another way to increase the color saturation that is consistent with the compression algorithm, JPEG, used in most digital cameras is to use the conversion of the RGB image into an

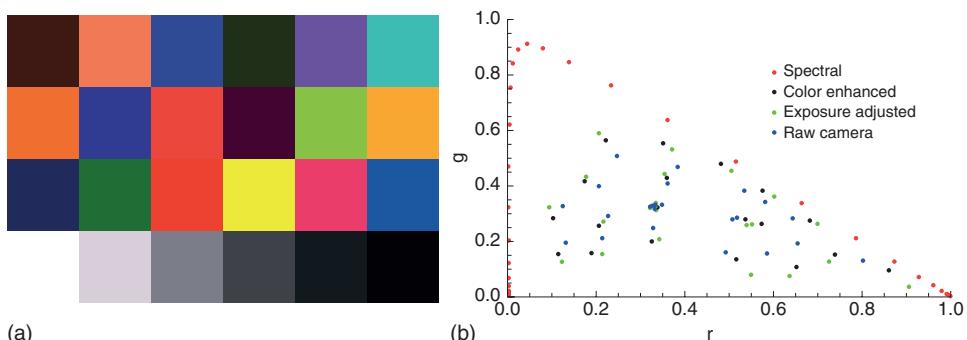


Figure 47 (a) The raw camera image operated on by the nonlinear color operation defined by Equations 25 through 27. (b) As shown by the r-g chromaticity plot, the neutrals remain fixed and the colors become more saturated, close to that of the exposure adjustment method

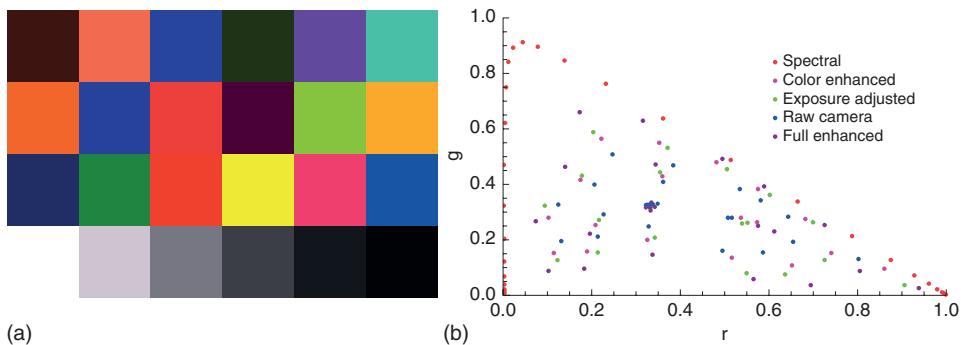


Figure 48 (a) Applying the exposure adjustment first, followed by the nonlinear color processing step. (b) The chromaticity shows how each stage color processing has increased the color gamut

YC_rC_b image. Once in the YC_rC_b format it is a simple matter to increase the chrominance layers of the image, while keeping the luminance channel, Y, fixed. Equations 22 and 23 show the transform to YC_rC_b and back to RGB. With the image in the YC_rC_b format a simple diagonal matrix is used to multiply each term to increase or decrease the chrominance of the image before it is converted back to an RGB image. Figure 49 shows such an operation with the following matrix operation;

$$\begin{bmatrix} Y' \\ C'_b \\ C'_r \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1.4 & 0 \\ 0 & 0 & 1.4 \end{bmatrix} \begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix}. \quad (28)$$

The off-diagonal terms can also be used to create some “artistic” effects, but they will not be considered here. The above increase in chrominance is significant and the impact can be seen in Figure 49 as compared to the raw camera image shown in Figure 45.

Figure 49 shows a Macbeth ColorChart that is much more saturated than the original raw image in Figure 45. The r-g chromaticity plot shows that in some cases the saturation exceeds the spectral limits of the spectral sensitivities. This means that in the image those patches that

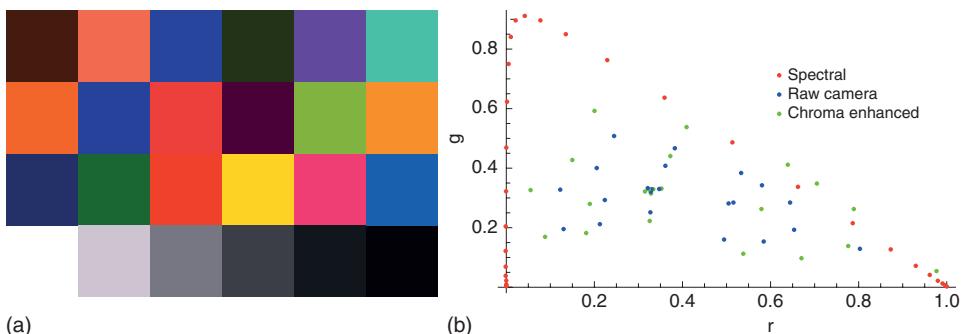


Figure 49 (a,b) Color enhancement using the YC_bC_a format

exceed the spectral limits will be clipped to some value on the spectral limit. In this case, those patches that exceed the spectral limit are the yellow, red, orange, and magenta patches. Note that the neutral scale is not affected by these operations.

The color gamut is controlled by a combination of the spectral sensitivities and the ability to compensate for the overlapping spectral sensitivities. For the sake of this discussion (and the above work) the spectral sensitivity, S_i , of a given color channel is defined by

$$S_i(\lambda) = S_{0i} e^{-\frac{-(\lambda-\lambda_i)^2}{2\sigma_i^2}} \quad (29)$$

where $i = R, G$, or B for the three color channels and S_{0i} is the amplitude, λ_i is the peak wavelength and σ controls the spread in the spectral sensitivity. The smaller σ is, the more narrow will be the spectral sensitivity, which improves the basic color gamut but at the expense of ISO camera speed. The more the spectral sensitivities overlap, the smaller will be the color gamut before employing exposure compensation as defined by Equation 24. Figure 50 shows the color gamut for three sets of spectral sensitivities where the red, green, and blue channels are centered at 450, 550, and 650 nm, respectively. Each channel in a given set has the same σ value and are 30, 40, and 50 nm, respectively. The narrow spectral sensitivities have a wider color gamut and as the spectral sensitivity overlap increases the color gamut decreases.

Figure 51 shows the result of imaging the Macbeth ColorChart with the above set of spectral sensitivities. Each set of spectral sensitivities is adjusted to give a neutral under a D50 illuminant (solid lines) as compared to the nonadjusted set (dashed curves). The top set of color charts represent the “raw” color reproduction based on the spectral sensitivities. One notes that as the overlap in the spectral sensitivities increases the colors become less saturated and simply “duller.” The bottom row of color charts represents the adjusted color chart when exposure compensation is applied as outlined earlier. Equation 24 shows exposure compensation matrix

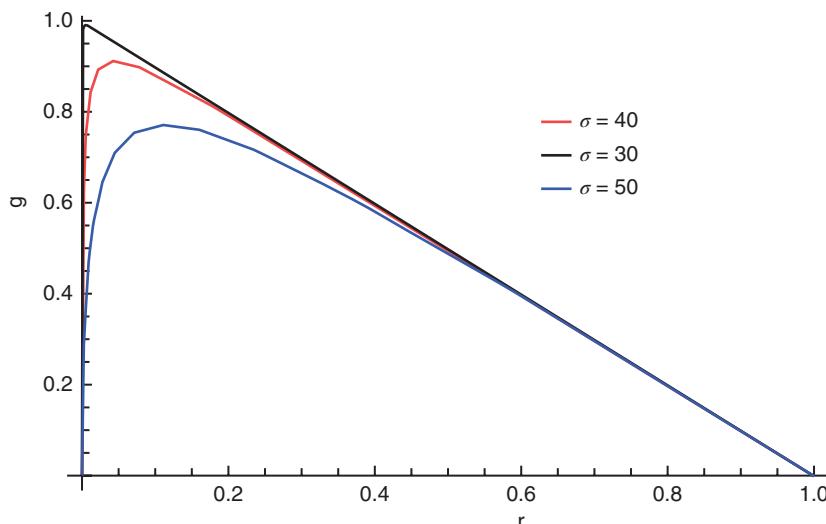


Figure 50 The possible color gamut in the r-g chromaticity diagram for three sets of spectral sensitivities with varying spectral widths defined by σ in Equation 29

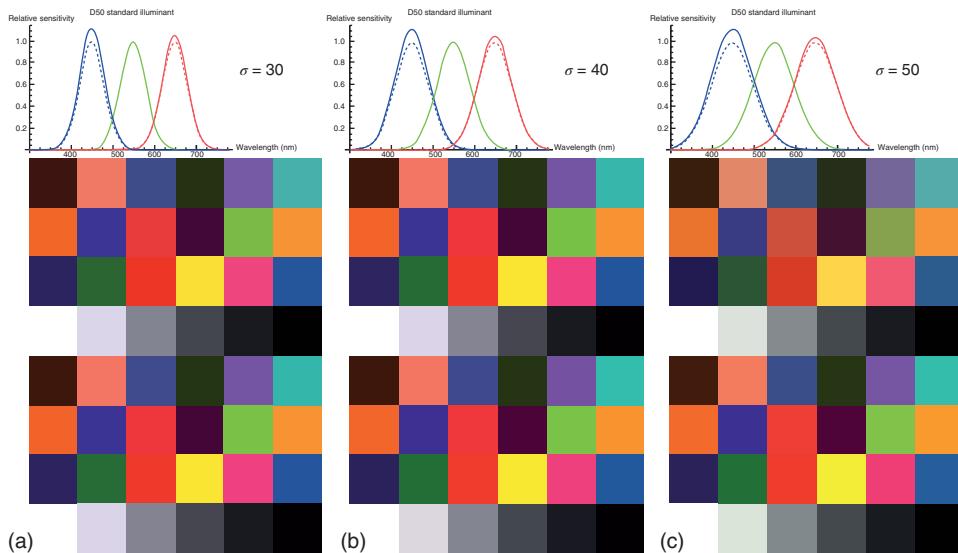


Figure 51 (a,b,c) Spectral sensitivity variations resulting in “raw” color chart results, top row, and after exposure adjustment as defined by Equations 24, 30, and 31

for the case of $\sigma = 40$ nm. The exposure compensation matrix for the cases of $\sigma = 30$ nm and $\sigma = 50$ nm are given below.

$$\sigma = 30, \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 1.04 & -0.04 & 0 \\ -0.05 & 1.1 & -0.05 \\ 0 & -0.04 & 1.04 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (30)$$

$$\sigma = 50, \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 1.28 & -0.26 & -0.02 \\ -0.27 & 1.57 & -0.3 \\ -0.01 & -0.26 & 1.27 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (31)$$

On the basis of the results shown in Figure 51 it is clear that the broader spectral sensitivities the less saturated the colors. But the exposure adjustment as defined by Equations 24, 30, and 31 make it visually easy to see the difference. Figure 52 shows the r-g chromaticity plots of the three color charts after the exposure adjustments have been made. The neutrals are very well clustered with only a small visual shift from neutral. For most of the color patches in the color chart, the saturation increases as the overlap increases, pointing to the usefulness of employing the exposure adjustment matrix. Greater color enhancement can be obtained by employing the techniques discussed earlier.

The above work indicates that having wider, overlapping spectral sensitivities is good for both ISO speed and color reproduction when the exposure adjustment is used. However, there is still room for optimization by moving the peaks of the red, green, and blue spectral sensitivities to their respective optimum positions. In the above example, the peaks are at 450, 550, and 650 nm, respectively. When additional color enhancement is applied it is clear that the neutrals tend to go yellow, which implies more blue signal (exposure) is required. This can be

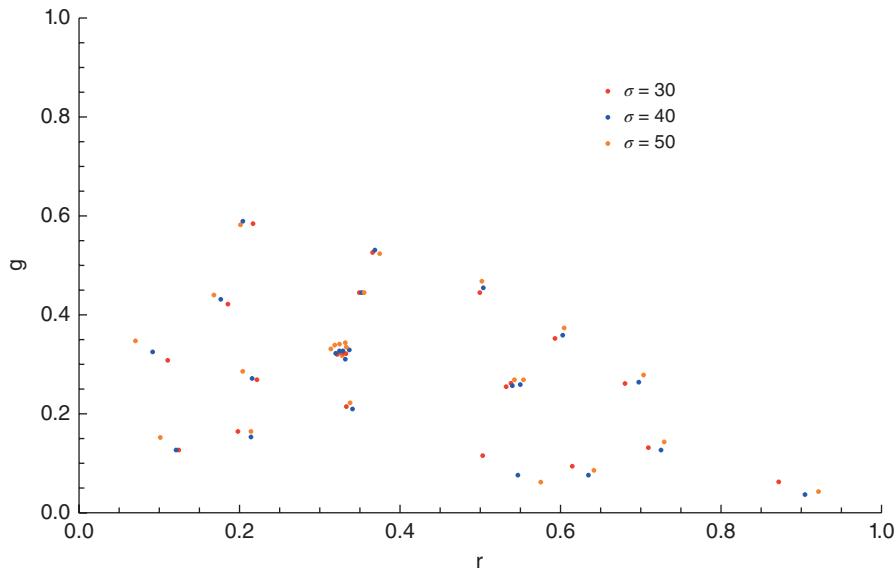


Figure 52 The r-g chromaticity plot of the color chart data of the color charts shown in Figure 51

obtained by σ values of 50 nm and peak sensitivities of 450, 560, and 650 nm, respectively. The base spectral sensitivities are still adjusted for a D50 illuminant and the exposure adjustment matrix is calculated and applied. Figure 53 shows the results. Note that the color saturation is very good and that the neutrals remain “neutral” when added color enhancement takes place. Although these models provide a lot of flexibility in adjusting the spectral sensitivities, this optimization will have to be accomplished in embedded software in the digital camera and may not be as robust as presented here.

6 Color Spaces Used in Digital Imaging

In the discussions earlier, all the color data is based on the values generated by the various camera models and is not related to the CIE XYZ or CIELab color systems [2–4]. The color reproduction is based on one of many color space definitions available for Mathematica and the LCD monitor used. In most cases the color space used is called sRGB, but others such as Adobe RGB are available [37, 38]. All the color images to date have been generated by calculations that give data based on the models used, which in turn were scaled between {0, 1} to drive the graphics in Mathematica. Without any user intervention this data was then converted to sRGB for printing and viewing. However, it is desirable how to define the various color spaces and how they are all centered on the CIE XYZ system. Figure 54 shows the range of sRGB and Adobe RGB in a CIE x-y chromaticity plot. Note that the Adobe RGB space includes more of the “blue-green” range of possible visible colors. Each color space is predicated on three “real” primaries from which all other colors can be constructed by a simple linear combination of the primaries. Recall that the primaries for the CIE XYZ system are not real and contain negative and positive combination of the original RGB primaries.

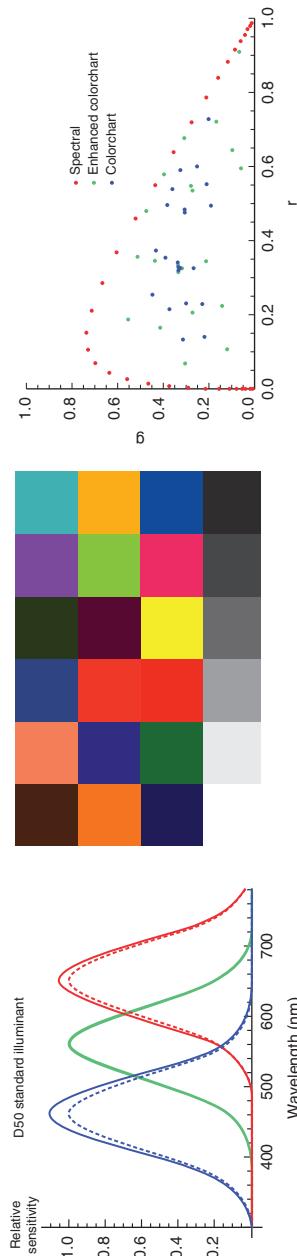


Figure 53 Optimized spectral sensitivities and the resulting color chart after exposure adjustment has been applied. The r-g chromaticity plots show good color correction and saturation while keeping good neutrals

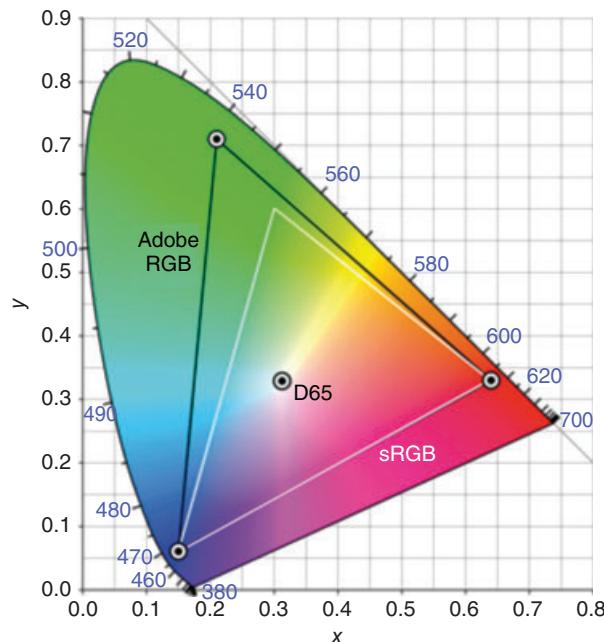


Figure 54 sRGB and Adobe RGB color spaces within the x - y CIE chromaticity plot. *Source:* Reproduced from A comparison of the Adobe RGB and sRGB color spaces within the CIE 1931 xy chromaticity diagram as published in Wikimedia Commons, 2013, http://commons.wikimedia.org/wiki/File:CIEExy1931_AdobeRGB_vs_sRGB.png © Mbearnsstein, 2013, distributed under the Creative Commons Attribution-Share Alike 3.0 Unported license

The camera calibration that transforms the camera RGB values to, say, sRGB values is done as follows using the Macbeth ColorChart. Take the spectral reflectance curve from each patch of the color chart and calculate the corresponding XYZ values using the CIE xyz color matching functions with the same “exposure” setting used by the camera; this is just an overall multiplying constant used to get a good exposure of the color chart and is 5.6 in most of the work above. Now by means of multivariate linear regression one finds the best least squares fit matrix, \mathbf{C} , that transforms the camera RGB values to the measured (and calculated) XYZ values. Next, using the defined sRGB primaries find a matrix, \mathbf{M} , that transforms XYZ values to sRGB values. Then one can transform camera RGB values to sRGB values by means of:

$$\begin{bmatrix} R_{sRGB} \\ G_{sRGB} \\ B_{sRGB} \end{bmatrix} = \mathbf{MC} \begin{bmatrix} R_{Camera} \\ G_{Camera} \\ B_{Camera} \end{bmatrix}. \quad (32)$$

This simple operation then provides accurate, reproducible sRGB values for storage in the camera (removable storage media) in 8-bit to 14-bits depending on the mode used by the camera. 8-bit data is used in most cases and the image is compressed using JPEG (or JPEG2000 in some cases). When the image is stored in RAW mode, 14-bits or more can be used, but software provided by the company that makes the camera must be used to generate a full color image. Some cameras can store images in the TIFF format as well and up to 16-bits

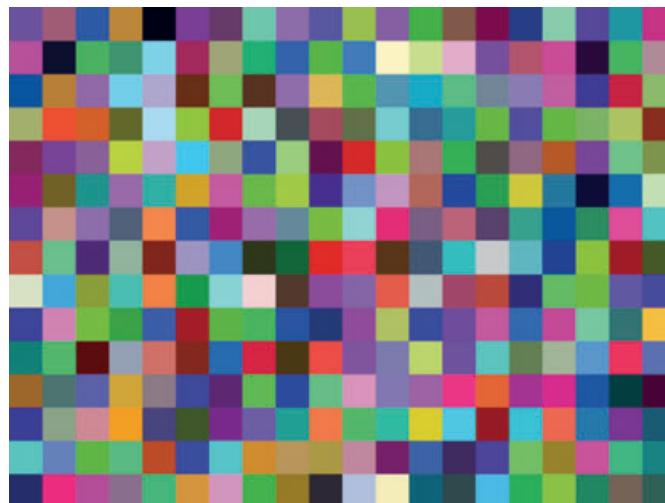


Figure 55 Large color chart used for color calibration

per color can be used. In practice a much larger samplings of colors is used to make these calculations such as the one shown in Figure 55.

To understand the process, consider two sets of primaries. Let \mathbf{R} , \mathbf{G} , and \mathbf{B} represent the unit vectors of the first primary, and let the second set of primaries be some linear combination of the first set and denoted by \mathbf{R}' , \mathbf{G}' , and \mathbf{B}' . Either set of primaries can describe the visible color space and any color in it. Assume that some color Q is defined using both sets of primaries;

$$Q \equiv r\mathbf{R} + g\mathbf{G} + b\mathbf{B} = r'\mathbf{R}' + g'\mathbf{G}' + b'\mathbf{B}', \quad (33)$$

where the lower case letters define the amount of each unit vector primary need to form the color Q . The linear transform relating the two sets of primaries is defined as

$$\begin{bmatrix} \mathbf{R}' \\ \mathbf{G}' \\ \mathbf{B}' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{G} \\ \mathbf{B} \end{bmatrix} = A \begin{bmatrix} \mathbf{R} \\ \mathbf{G} \\ \mathbf{B} \end{bmatrix}. \quad (34)$$

The matrix A can easily be found if the two sets primaries are known.

Using the relation defined by Equation 34 one can express \mathbf{R}' , \mathbf{G}' , and \mathbf{B}' in terms of \mathbf{R} , \mathbf{G} , and \mathbf{B} . This leads directly to the following relationship:

$$\begin{bmatrix} r' \\ g' \\ b' \end{bmatrix} = [A^T]^{-1} \begin{bmatrix} r \\ g \\ b \end{bmatrix}. \quad (35)$$

Using Equation 35 one can create the required transform from the CIE XYZ system to any other color system defined by a set of primaries. The actual transform does depend on the on the illuminant used, so here the transforms for sRGB and Adobe RGB will be given for the D50 illuminant to be consistent with the work above with the Macbeth ColorChart. For Adobe

RGB the transform is (up to three decimal places):

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.962 & -0.611 & -0.341 \\ -0.979 & 1.916 & 0.033 \\ 0.029 & -0.141 & 1.349 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (36)$$

For sRGB the conversion matrix (up to three decimal points) is given by:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.134 & -1.617 & -0.491 \\ -0.979 & 1.916 & 0.033 \\ 0.072 & -0.229 & 1.405 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (37)$$

In Equations 36 and 37, the XYZ values are those calculated from Equations 10 through 12 and R, G, and B represent the red, green, and blue values for the given space. For most graphic display programs these numbers are scaled between {0, 1}.

The above calculations all take place in a linear color space independent of any display device. The display of color was accomplished by the use of CRTs from about 1950 to around 2005 when a whole host of other display devices replaced the CRTs including LCDs, plasma screens, and DLPs micro mirrors, which can be found in most digital projectors today in motion picture theaters. Each of these devices has their own, unique physical characteristics. Most of the signal preparation is based on the CRT's characteristics. Each of the three electron beams are accelerated by a complex series of voltage plates and magnetic focusing devices to hit individual red, green, and blue emitting phosphors on the face of the CRT. To achieve a linear increase in the emission output of these phosphors, each channel had to increase the voltage by a power law with an exponent, or gamma, of

$$\text{Signal}_{R,G,B} = \text{Code Value}^{2.2}. \quad (38)$$

The Code Values indicated earlier normally range from 0 to 255, but could have as many as 12 bits and the gamma value is 2.2. The color calculations made on the hardware in the digital camera will have from 32 to 64 bits to prevent round off error. The final results will then be expressed in 8–12 bits per channel and then Equation 38 is applied and the values are renormalized to 8–12 bits. In this way, the data can be used to drive a CRT without further modification. Today, LED displays dominate the market and their light forming characteristics are very different from CRTs, but the electronics are designed to accept the gamma corrected signals and then modified to meet the particular light emission characteristics of the display. Hence, sRGB and Adobe RGB contain the gamma corrected values indicated by Equation 38.

The human visual system is far more complex than the above linear analysis and there is a substantial body of work on how color vision changes under different viewing conditions and how it strives to maintain color constancy even when the physical reflection or emission spectra are changing. This work is well beyond the scope of this chapter [5, 16, 17].

7 Color Filter Arrays

The ideal digital camera (or scanner) would have three sensors, optically aligned in such a way that three geometrically identical images would be collected for the red, green, and blue

channels. Although ideal, such an optical-sensor combination is very expensive and bulky. Almost all DSCs use a single sensor array with a CFA as shown in Figure 2 [1]. Figure 2 shows three such CFAs; the Bayer array is by far the most popular, the Sony [38] array was developed to replace the Bayer array but is no longer in use because the Bayer patent has lapsed, and the Hitachi array that was first developed for a color video system. The Hitachi CFA [39] strength is that all pixels have a green component associated with it and it has less light absorption than the simpler Bayer CFA. However, the resulting demosaiced image has higher noise (owing to the off-diagonal terms in the reconstruction matrix) and more artifacts where color transitions take place. The complementary Bayer CFA, using cyan, magenta, and yellow instead of red, green, and blue, respectively, allow for better use of light (a “faster” array), but it too introduces more noise and artifacts. The staggered red, green, and blue CFAs have alternating R-G-B pixels along a line, but each subsequent line is staggered one pixel to produce a CFA that looks like diagonal lines of red, green, and blue. The staggered R-G-B shows promise, but has never been adapted in DSCs.

Any sampled image introduces aliasing [40] or low frequency banding and artifacts. The sampling theory developed by Nyquist simply states that all frequency information beyond half the sampling rate of a signal (image) will be reproduced as lower frequency information (folded back as a mirror image at the Nyquist frequency). Consider a black and white sensor that has a “square” sampling pattern with a pitch of 10 µm between the centers of the pixels. This means that the sampling frequency is 100 cycles per millimeter. Thus, any information in the image beyond 50 c/mm will be folded back along the 50 c/mm lines to a lower frequency. So a 60 c/mm signal will become 40 and a 90 c/mm signal will become a 10 c/mm signal. Consider a simple Sony Mavica type CFA [41] that has red, green, and blue vertical stripes with 10 µm pixels. For each channel, the sampling rate is 33 c/mm and they are out of phase. So a 50 c/mm neutral signal will be aliased to a 37 c/mm color signals and each is shifted one pixel. The resulting image will be a series of 37 c/mm colored (rainbow like) bands instead of the original 50 c/mm signals. The challenge is how does one eliminate these color aliasing artifacts from the captured images using a sampling sensor with a CFA [41–52].

Before more mathematical detail is given, it is best to show some real images from early digital cameras and some simulations. Kodak built in the mid-1990s a full frame CCD camera in a Nikon body called the *DCS 460* [53]. It had a resolution of 2400×3000 pixels and used a Bayer CFA. There was no attempt to reduce the frequency content of the image and hence it had strong color aliasing. Figure 56 contains a down sampled version of the image of a man in a purple pinpoint shirt and a full resolution sample of part of the shirt. As the lens used on the Nikon camera was very good, information well beyond the Nyquist frequency, 42 c/mm for the green pattern, was imaged onto the sensor. The resulting aliasing is found in the complex color bands in the shirt, the distortions being due to the various folds in the cloth. Figure 57 shows a picture of the (younger) author in front of a test chart designed to show the problems of aliasing in DSCs. The various test patterns indicate a variety of aliased images, and as the chart was slightly askew, the aliased patterns are at an angle; aliased images are very phase dependent in that a slight change in relative position between the sensor and subject can result in very significant changes in the final aliased image.

Figure 57 shows a more controlled study of the aliasing for the DSC 460. The vertical and horizontal lines progress by a factor of two. The first line pattern shows little aliasing while the second shows the actual lines mixed with some orange and blue low frequency bands; the angle is due to the fact that the camera was not aligned perfectly to the test chart. The third

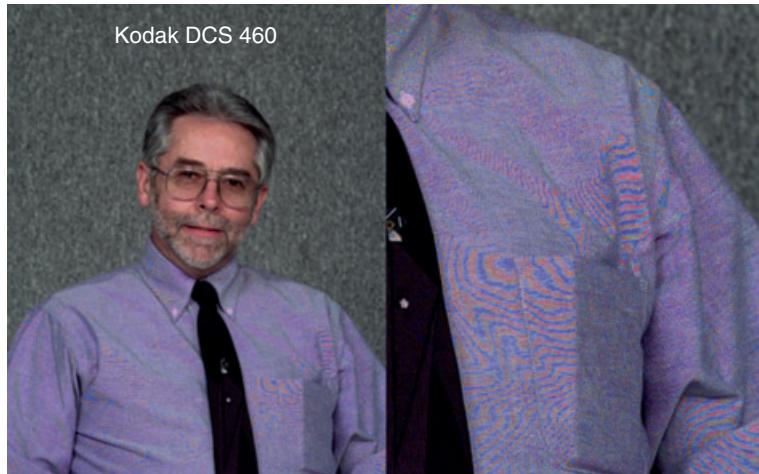


Figure 56 Image from a 2400×3000 pixel sensor with no anti-aliasing protection and a Bayer CFA. The full image on the right has been reduced in size, but the image on the right is full resolution

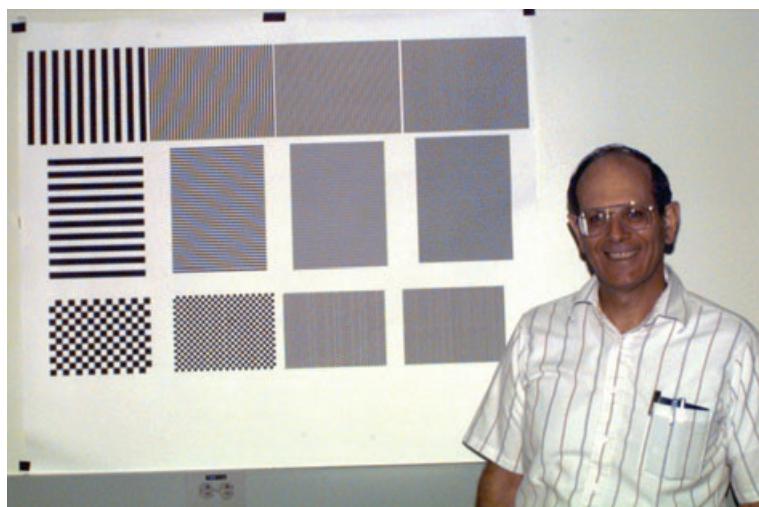


Figure 57 Image taken with DSC 460 of author and a test chart to show aliasing

and fourth line sets shows both color banding and the “neutral” lines are highly aliased to lower frequencies. The horizontal lines shows much the same results, but the checkerboard test patches shows a more complex response to the sampling.

Figures 56 and 57 define the problem of aliasing with a Bayer CFA. To study the impact of CFAs in general one needs to use image simulation. The first example will be to simulate a simple 500×500 camera using a Bayer CFA. To be able to show the impact of lens optics it is necessary to start with a much bigger array and then simulate the impact of the lens.

A 2000×2000 test target will be used to start the simulation. It will be assumed that sensor will have $5 \mu\text{m}^2$ pixel in a square array with a pixel count of 250,000 (500×500). The focal length of the lens will be 11 mm and its light gathering power will be an F/2 lens with a diameter of 5.5359 mm. Now such lenses do not normally have the MTF (modulation transfer function) of an F/2 diffraction limited lens. In the first simulation it is assumed that the MTF is of an F/5.6 diffraction limited lens, which is still very good. The optical spread function of the F/5.6 diffraction limited lens is then calculated and it is convolved with the 2000×2000 image to give an image that fall on the 500×500 sensor. To obtain the 500×500 resolution, four adjacent pixels in the 2000×2000 image are summed and averaged to produce on pixel of the 500×500 array. This sampling introduces aliasing and the MTF of a $5 \mu\text{m}^2$ pixel. Next the impact of the Bayer CFA is introduced by separating the image into three sparse channels, green, red, and blue, based on the Bayer CFA pattern. This will introduce more aliasing and the aliasing in the red and blue channels will be twice that of the green channel; later we will see this will have an impact on the ability to reconstruct the image with no color artifacts. The sparse images are then interpolated using the kernels given by Equation 39 in a convolution with the defined sparse layers. The red and blue convolution kernels are the same.

$$\begin{aligned}\text{Green Kernel} &= \begin{bmatrix} 0 & 0.25 & 0 \\ 0.25 & 1 & 0.25 \\ 0 & 0.25 & 0 \end{bmatrix}, \\ \text{Blue-Red Kernel} &= \begin{bmatrix} 0.25 & 0.5 & 0.25 \\ 0.5 & 1 & 0.5 \\ 0.25 & 0.5 & 0.25 \end{bmatrix}. \quad (39)\end{aligned}$$

The interpolated sparse image channels are next combined to form a final image. Figure 58 shows the interpolated test target with an F/5.6 and an F/16 diffraction limited lens MTFs.

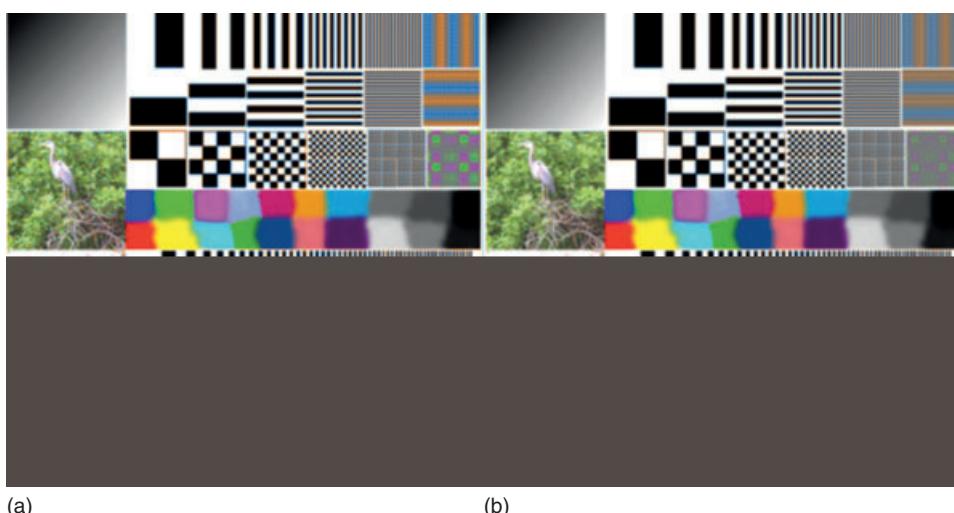


Figure 58 The simulation of a camera image taken with a 500×500 image sensor. (a) Diffraction limited F/5.6 lens and (b) diffraction limited F/16 lens

Note that the poorer F/16 diffraction limited lens produces a less sharp image and the aliasing, while still present, is less pronounced.

It is clear from Figure 58 that the sharper lens, F/5.6, produces a sharper looking image, but the color banding is stronger and more visible. The image on the right, F/16 diffraction limited lens, softens the aliasing but does not eliminate it. Most DSCs employ a low pass pre-filter to eliminate much of the aliasing in the sampled image [54]. The most common low pass pre-filter is made of up thin sheets of a birefringent material (quartz) and a quarter-wave plate that results in a single spot of light becoming four spots separated by a distance that is a function of the thickness of the two birefringent plates. Studies have shown that the best tradeoff between low aliasing and good sharpness is to move the displaced image (in both the vertical and horizontal directions) two pixel units should be when the Bayer CFA is used to encode color. Figure 59 shows the results of using such a low pass pre-filter within a camera that has an F/5.6 and F/4 diffraction limited lens MTF. It is clear the low pass pre-filter removes most of the aliasing when an F/5.6 diffraction limited lens MTF is present, but the sharper F/4 diffraction limited lens passes enough high frequency content that the aliasing starts to increase despite the low pass pre-filter.

The above examples are relatively low resolution in that today digital cameras often exceed 10 million pixels (2828×3535) on the low end cameras and 24 million pixels (3651×5477) on the high end. At these resolutions, aliasing is less because the sampling rate is so high and only a few images will alias (depending on content). That makes the reconstruction algorithms (demosaicing) much easier. Also, in many less expensive cameras with pixels around $2 \mu\text{m}^2$ or less, the optical spread function of the lens far exceeds the pixel size, so no low pass pre-filter is required. In the high end SLR DSCs the camera often has an option to remove the low pass pre-filter without any noticeable aliasing taking place in most images.

To provide a flavor for the type of demosaicing algorithms developed over the years, images ranging from 500×500 to 2000×2000 in resolution will be studied and the impact of the lens or a low pass pre-filter will be ignored so not to diminish their respective impact. The bilinear interpolation for the Bayer CFA given by Equation 39 will act as the base for comparison.

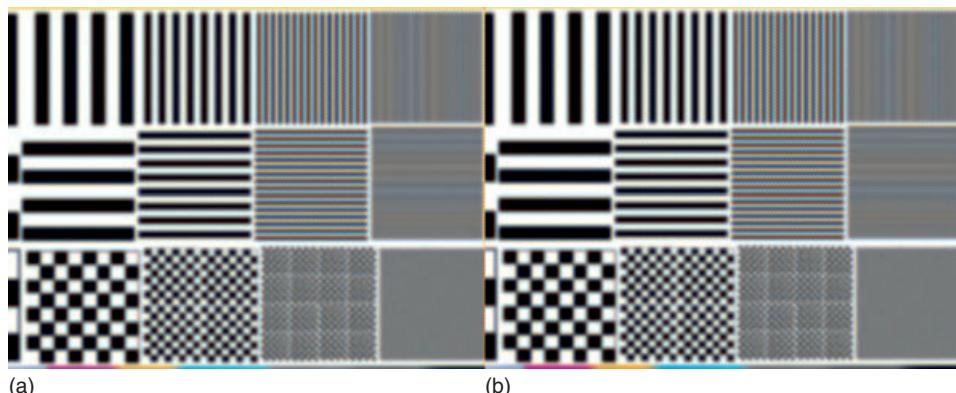


Figure 59 The low pass pre-filter used in conjunction with an F/5.6 (a) and F/4.0 (b) diffraction limited lens MTF. The F/4 diffraction limited lens passes more high frequency information and thus has more aliasing

A new demosaicing algorithm will employ logical evaluation of the sparse images in all three channels. For example, in the green channel one can ask three questions. There are four known green pixels around a given an empty green pixel. The normal thing is to take the average of these pixels. However, one can check to see if the center pixel is more likely to reside on a vertical or horizontal line. By taking the vertical and horizontal differences one can determine if the “true” value lies along one of these lines by rejecting the one with the largest difference and then just using two known pixels to calculate the missing value; a simple average. If the two difference values are the same or very large, then just take the average of the four pixels. The same concept can be applied to the red and blue channels, but one can include the two diagonal directions as well. Figures 60–62 show the difference between the standard bilinear interpolation and the above logical interpolation in the three channels for image resolutions of 500×500 , 1000×1000 , and 2048×2048 . The color aliasing in the 500×500 images is pronounced and for the most part is not corrected by either method, but the patterns that appear are more apparent in the simple bilinear interpolation. The bilinear interpolation tends to make vertical and horizontal images look like castle walls (green channel) and hence more visible. The logical algorithm tends to turn vertical and horizontal lines into smooth transitions from black to white with some uniform color banding (orange or blue), which is less visible. For the 1000×1000 image, the logical interpolation is far more effective, but does not fully eliminate the color artifacts. For the 2048×2048 image, both methods seem to do a good job, but there is still some aliasing in the test patterns, while the images show little impact. Hence, one can see why as the resolution increases, the issue of aliasing becomes less of a problem (unless one makes significant enlargement of the image). Using a Nikon D70 (2000 \times 3000 pixels) and a very good lens, one can find the typical bilinear artifacts caused by distant buildings; this is a case where the periodic low pass pre-filter is at a maximum transmittance at higher frequencies. Hence, it seems that for in-camera processing complex demosaicing may not be

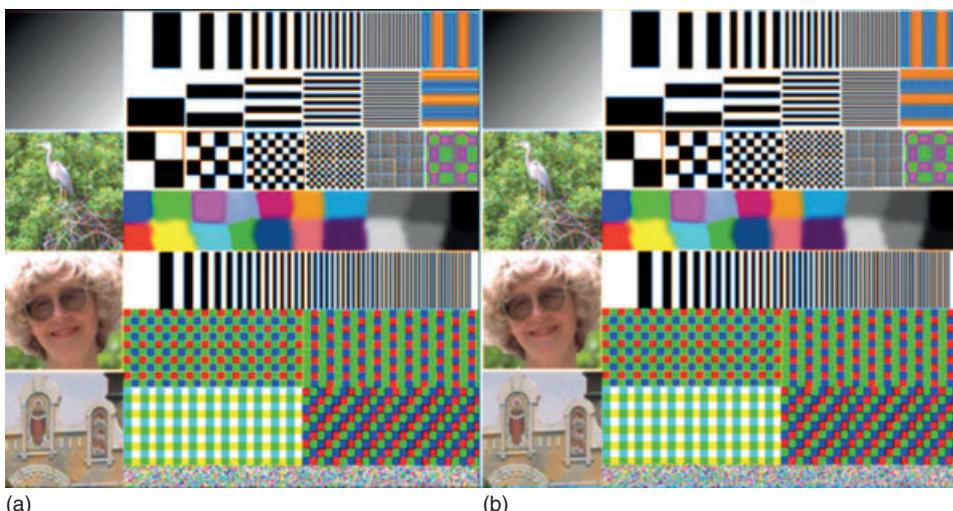


Figure 60 500×500 image interpolated by a logical algorithm (a) and the simple bilinear interpolation (b) for a Bayer CFA. Little overall improvement is apparent even though the exact natures of the artifacts are different

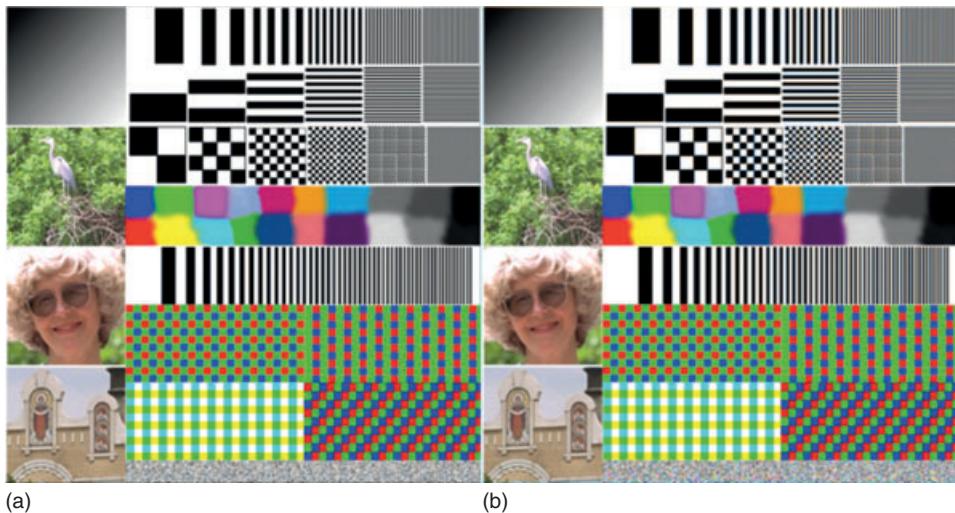


Figure 61 A 1000×1000 image interpolated by a logical algorithm (a) and simple bilinear interpolation (b) for a Bayer CFA. Here the logical interpolation does a better job of reducing the aliasing. Look at the vertical and horizontal lines as well as the random noise at the bottom of the image

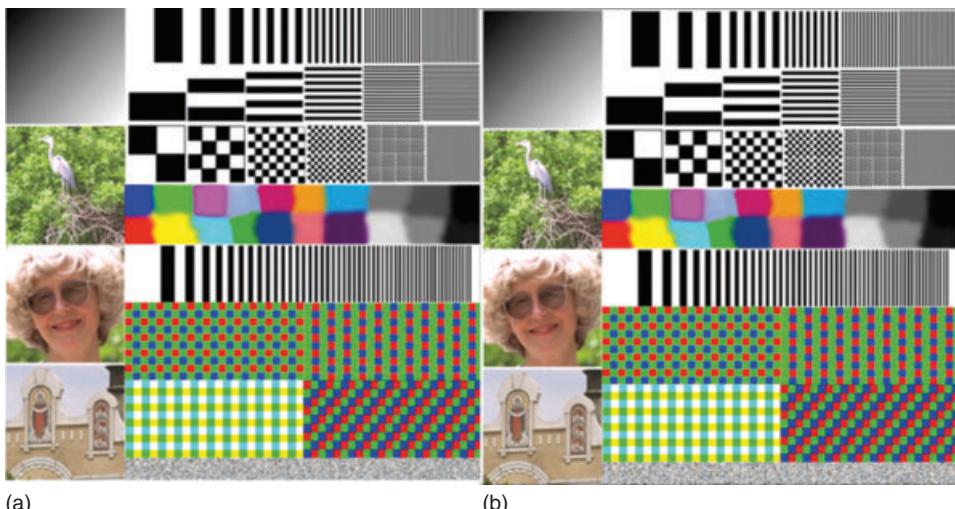


Figure 62 A 2048×2048 image interpolated by a logical algorithm (a) and a simple bilinear algorithm (b). Both do a good job at this resolution, but upon close inspection the logical interpolation is better

required. Using raw data, one is able to process the images with more sophisticated algorithms to meet the need of the image usage. Figure 63 shows the detailed aliasing in the bar patterns. Note that the logical interpolation has the greatest impact on the 1000×1000 image, greatly reducing the color artifacts.

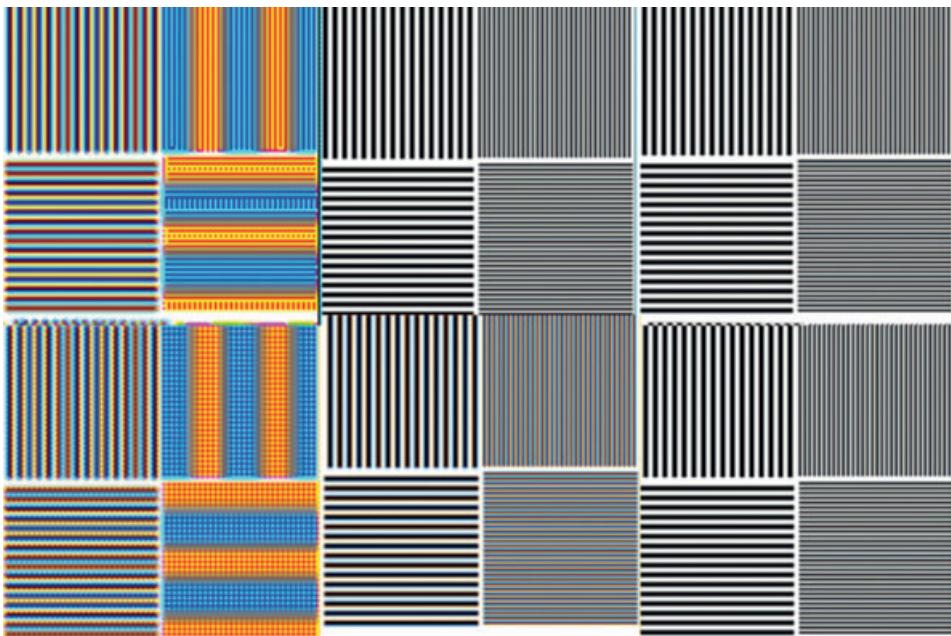


Figure 63 Details of the logical and bilinear interpolation. Top row is the logical interpolation and the bottom row is the bilinear interpolation. The resolution is, from left to right, 500×500 , 1000×1000 , and 2048×2048 . The biggest impact of the logical interpolation is on the 1000×1000 image

There are more complex logical interpolations schemes based on template matching (edges, diagonals, corners, etc.), but like the logical algorithm used here takes much longer to process and may not be suitable to in-camera processing. The kernel used in bilinear interpolation is a 3×3 matrix. It is possible to use larger interpolation kernels and get better approximations to the true value. One approach could be to use the Whittaker–Shannon interpolation [55] in two dimensions to get an ever better measure of the missing pixels, but once again the complexity of the calculations will require much more time and internal memory.

Another large class of demosaicing algorithms uses the natural correlation in images between the three channels [56, 57, 58]. In what follows a new CFA, the staggered R-G-B array will be used to demonstrate something besides the Bayer CFA. The staggered R-G-B array and the subsequent bilinear interpolation, see Equation 40, produces good rendering of vertical and horizontal lines, but does not do well with diagonal line structures as will be demonstrated later. Two types of images will be used, the test chart shown earlier and a more natural image of three women. The impact on a 1024×1024 image of the three women is shown in Figure 64. The resolution of Figure 64 may be such that it is hard to see any artifacts, just that the image is less sharp. Figure 65 shows some of the details.

$$\text{Staggered R-G-B Kernel} = \begin{bmatrix} 1/3 & 1/3 & 0 \\ 1/3 & 1 & 1/3 \\ 0 & 1/3 & 1/3 \end{bmatrix}. \quad (40)$$

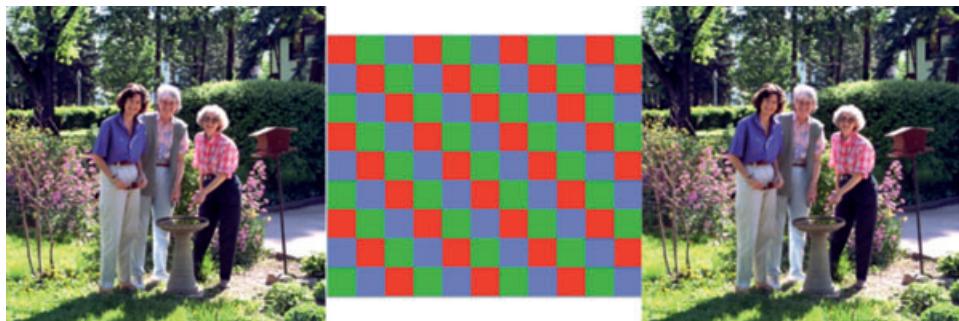


Figure 64 A natural image, 1024×1024 , sampled by the R-G-B staggered array as shown in the middle. The original is on the left and the demosaiced on the right

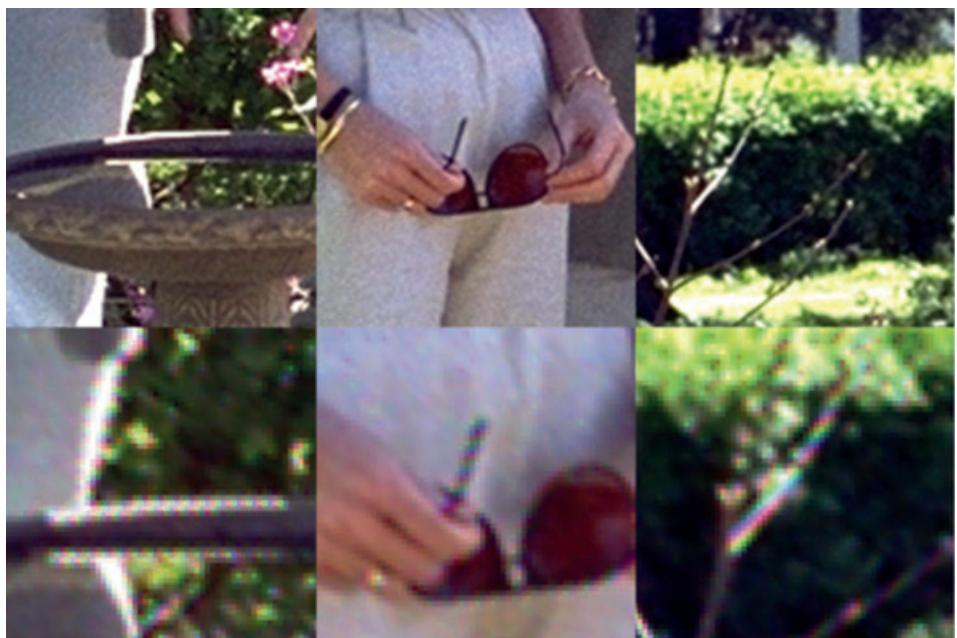


Figure 65 Enlargements from Figure 62. The top row represents the original 1024×1024 image (lower magnification) while the bottom row represents the image captured by the R-G-B staggered array and demosaiced by the kernel in Equation 40

From Figure 65 it is easy to see the aliasing artifacts introduced by the staggered R-G-B array. One method to remove these artifacts is to use the spatial correlation between the three channels. In this simplified algorithm the following steps take place. The original image is sampled according to the array pattern in Figure 64, creating three sparse images. These images are then demosaiced by applying the kernel given by Equation 40. The resulting three layers are then each smoothed (blurred) by a low pass filter, which looks like a pyramid. The blurred

image tends to smooth out the color aliasing. Now at each pixel site in the image ratio between the red and blue, red and green, and green and blue pixel values are taken. These ratios are used to fill in the missing pixel values in all three channels in the following way. Consider the sparse green layer first. At a missing green value site, one checks to see if a red or blue pixel was measured. Assume the red pixel was recorded. One then takes the ratio between the red and green pixels in the smoothed image and multiplies by the ratio of green to red (it may be the inverse of the value stored in the algorithm) to get the “new” green value. If it is a site where the blue image was recorded, then the new green value is the blue value times the green to blue ratio at that site. The identical process is repeated for the red and blue sparse layers. One now has a demosaiced image that is hopefully better than the simple demosaicing kernel given by Equation 40. Figure 66 shows the results when the 1024×1024 image of the three women is processed with the R-G-B staggered array and a blur kernel of size 5×5 pixels is used to smooth each demosaiced layer as described earlier. The results are clearly better and only a few areas of color aliasing remain, although less, as seen from Figure 65. In a 2048×2048 example of the same image, all sense of aliasing will vanish. Hence, reinforcing the idea that as resolution increases, the image requires a much less complicated demosaicing algorithm.

Although the combination of the R-G-B staggered CFA and the correlated demosaicing algorithm seems to do a very good job on a 1024×1024 natural image, it is instructive to look at the test chart seen earlier. Figure 67 will show four images, the original 1024×1024 image seen by the R-G-B staggered array and reconstructed by the use of the kernel in Equation 40. Then, three different blur filters, 5×5 , 9×9 , and 15×15 , will be used to blur the image used in the correlation algorithm. Each of the blur kernels looks like a pyramid, but the larger the kernel is, the more will be the blur and the less will be the “normal” aliasing artifacts. This can easily be seen in the fine checkerboard pattern at the lower far right of upper test patterns. As noted



Figure 66 On the left is the 1024×1024 image sampled by the R-G-B staggered array and processed with the correlating algorithm where a 5×5 blur kernel was used. To the right are the three detail images seen in Figure 63 along with one of the arms against the black trousers, which is a good place to notice aliasing

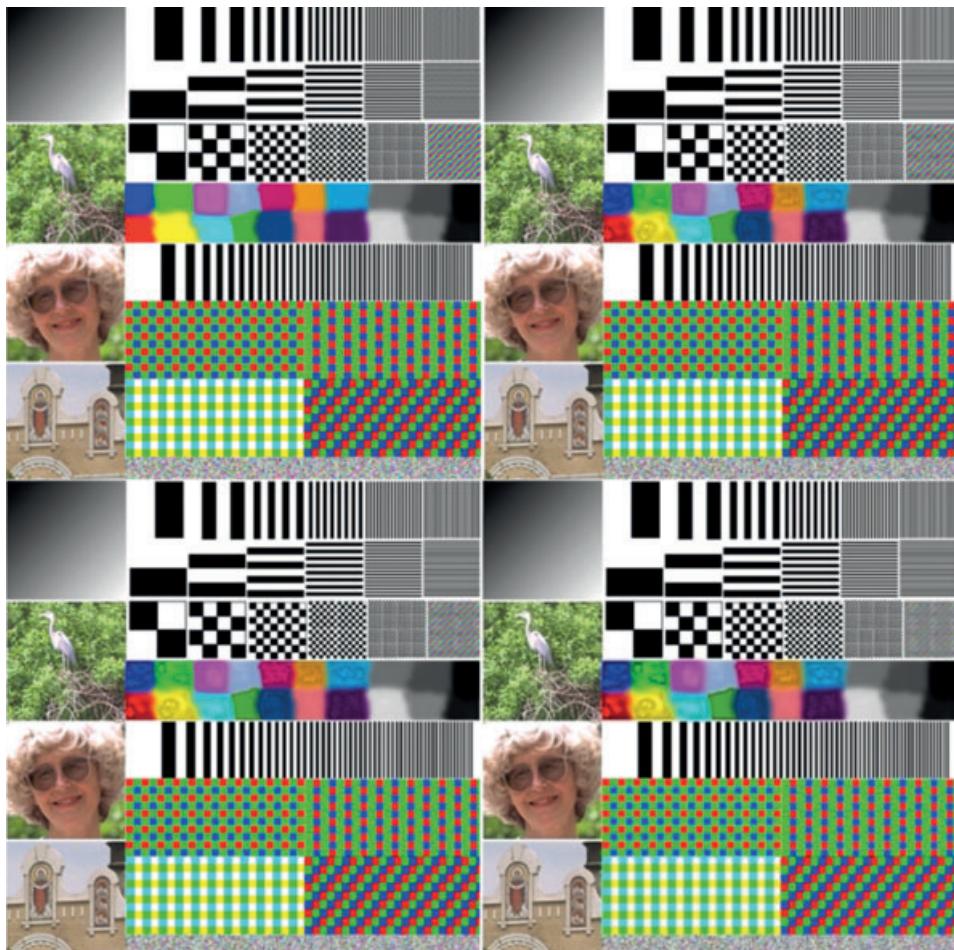


Figure 67 The upper left image is the conventional interpolation for a staggered R-G-B CFA using the kernel defined in Equation 40. The upper right is using the correlation demosaicing algorithm with a 5×5 blur. The lower left and right images use a 9×9 and 15×15 blur, respectively. Figure 68 shows some of the details

before, the staggered R-G-B CFA reproduces horizontal and vertical lines very well, but tends to strongly alias diagonal lines such as those seen in the checkerboard patterns. As the blur becomes wider the ratios generated seem to be closer to the true correlation between the original and hence the results are better. Keep in mind that we do not use the blurred values, only their respective ratios and that only actually recorded pixel values are used in the calculation of the missing pixels in each channel. However, the artifacts in the black and white test patterns are greatly reduced (these patterns are completely correlated) when a color transition takes place.

The algorithm introduces some black dots at the color change boundaries. The reason for this is that a color changing edges there is little correlation between the images and the ratios

calculated are not representative of the true values at one pixel site of the three channels, while in smoother varying regions the ratios are more representative of the true values. Look at the various color patterns near the lower part of the image to see these artifacts. Also, in the “painted” color patches, the correlation algorithm introduces a range of black dot clusters because of the varying nature of texture in the color patches. On the other hand, look at the tree branches in the picture of the blue heron on the river. Here one sees that as the blur increases to calculate the ratios, the image becomes free of the artifacts. In Figure 66, it is seen that the fine checkerboard pattern is highly aliased along the diagonal and as the blur kernel increases the color aliasing is greatly diminished on the diagonal and some vertical neutral aliasing become visible. Figure 68 also shows that the original image from the use of the staggered R-G-B CFA introduces some fine dots around the colored squares in the various colored checkerboards. When the correlated algorithm is used to diminish the color aliasing artifacts, the visible appearance of the dots is increased and is probably due to sudden changes at the edges where the ratios between layers introduce errors.

The “black dots” can be minimized (but not always eliminated) by not using the correlation between the blurred layers when the “known” color value is low or even zero, thus introducing a black dot. If one is in a uniform red area and the missing red value is located at a point where the blue layer has been sampled, that blue value may be zero or close to zero. By setting a threshold, one can ignore the correlated method, but instead use the value from the first demosaicing method (Equation 40 in the case of the staggered R-G-B array). The exact value of the threshold depends on the nature of the sampling array and the resolution. For most cases a threshold of 0.1 produces a good image in uniform colored areas where the pixel range is given by {0, 1}. Using the same method with the Bayer CFA also produces excellent images. Figures 69 and 70 show 2000×2000 images of the test chart and group of ladies using the correlated method and a threshold of 0.1 and a blur kernel set at 5×5 .

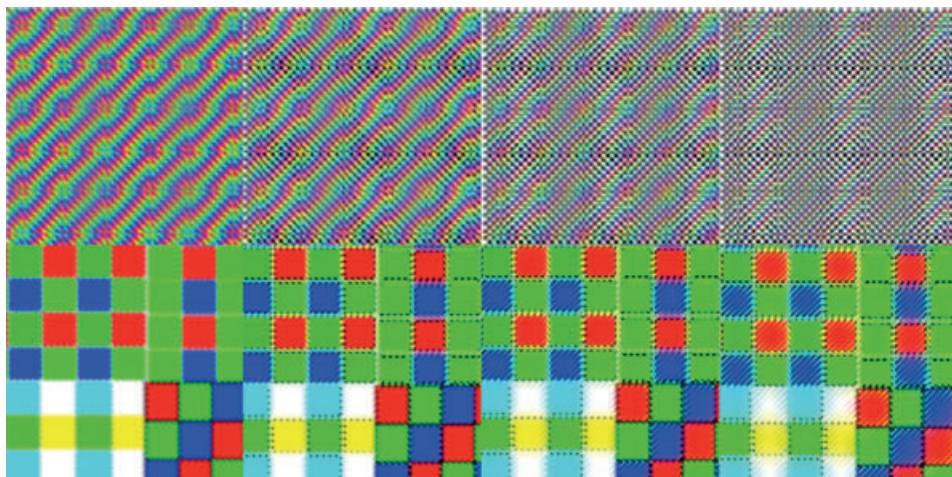


Figure 68 Strength and weaknesses of correlation algorithm on images captured by a staggered R-G-B CFA. The top row is the finest checker board patch and the lower row is the intersection of the different color checkerboards shown in the test chart image. In both cases, the original image is 1000×1000 and from left to right there is no blur kernel, a 5×5 blur kernel, a 9×9 blur kernel, and a 15×15 blur kernel



Figure 69 (a) 2000×2000 image sampled by Bayer CFA with a two-stage demosaicing algorithm; logical interpolation followed by correlation interpolation. (b) 2000×2000 image sampled by staggered R-G-B CFA with a two-stage interpolation; first stage is bilinear interpolation followed by correlation interpolation

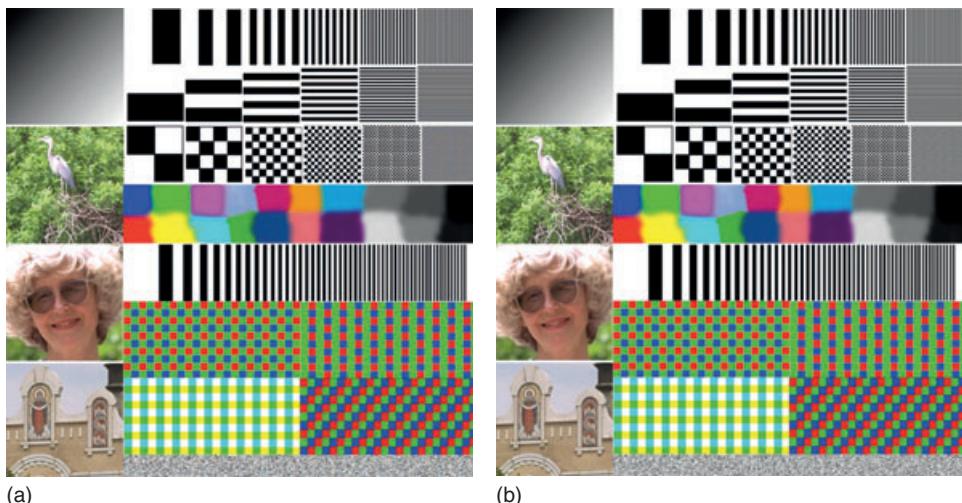


Figure 70 (a) 2000×2000 image sampled by Bayer CFA with a two-stage demosaicing algorithm; logical interpolation followed by correlation interpolation. (b) 2000×2000 image sampled by staggered R-G-B CFA with a two-stage interpolation; first stage is bilinear interpolation followed by correlation interpolation

The size of the smoothing or blurring kernel needs to be adjusted to the resolution of the image and the frequency content within the image. For low resolution images the blurring kernel needs to be large enough to “wash out” the color banding. This will result in the loss of detail in high frequency test target patches, turning them into a neutral gray, but this loss of detail is less noticeable than the strong color banding. As the resolution increases and color aliasing become less the blurring kernel can shrink for computational ease (with good effect on the highly correlated test patches), but keeping a fixed blur kernel will not greatly hurt the image except at sharp boundaries between colors, where some darkening may occur.

Another method of eliminating color aliasing is to transform the three layers into the spatial frequency domain by means of a fast Fourier transform (FFT). In the spatial frequency domain, each sparsely sampled layer will show the replicas of the image spectra away from the area around zero spatial frequency. This spectrum around zero frequency is within the two-dimensional Nyquist region and represents the un-aliased image. By eliminating the replicas outside the Nyquist region (set the Fourier coefficients to zero) and applying the inverse FFT to each layer and then combining them to form an image can eliminate a lot of the color artifacts. This is equivalent to applying the Whittaker–Shannon interpolation noted earlier.

As the resolution of high quality sensors increases (Digital SLRs) simple bilinear interpolation (ranging from kernels of 3×3 to as large as 9×9) are more than sufficient to produce high quality images, free of color aliasing. On the other hand, mobile phones with cameras and less expensive point-and-shoot cameras with high pixel counts (and small pixels) may have sufficient pre-filtering by the relatively (to pixel size) large optical spread function to prevent any serious color aliasing, so simple bilinear interpolation might be sufficient for an adequate image.

References

1. Bayer, B.E. (1976) Color imaging array. US Patent 3971065 A.
2. <http://www.cie.co.at>
3. Wyszecki, G. and Stiles, W.S. (1982) *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd edn, John Wiley & Sons, Inc, New York.
4. Hunt, R.W.G. (2004) *The Reproduction of Color*, 6th edn, John Wiley & Sons, Ltd, Chichester.
5. Fairchild, M. (2005) *Color Appearance Models*, 2nd edn, John Wiley & Sons, Ltd, Chichester.
6. Ohta, N. and Robertson, A.R. (2005) *Colorimetry Fundamentals and Applications*, John Wiley & Sons, Ltd, Chichester.
7. Newton, I. (1704) Opticks, <http://www.rarebookroom.org/Control/nwtopt/index.html>.
8. Longair, M.S. (2008) Maxwell and the science of colour. *Philosophical Transactions of the Royal Society A*, **366**, 1685–1696.
9. Le Grand, Y. (1968) *Light, Colour and Vision*, 2nd edn, London, Chapman and Hall.
10. Wright, W.D. (2007) Golden jubilee of colour in the CIE – the historical and experimental background to the 1931 CIE system of colourimetry, in *Colourimetry* (ed. J. Schanda), Wiley Interscience, Hoboken, NJ, pp. 9–24.
11. Stiles, W.S. and Burch, J.M. (1959) NPL colour-matching investigation: final report. *Optica Acta*, **6**, 1–26.
12. Hornbeck, L.J. (1993) Current status of the digital micromirror device (DMD) for projection television applications (invited paper). International Electron Devices Technical Digest, 381–384.
13. Thornton, W.A. (1973) Matching lights, metamers, and human visual response. *Journal of Color & Appearance*, **2** (1), 23–29.
14. Valeur, B. and Berberan-Santos, M. (2011) *Molecular Fluorescence*, 2nd edn, Wiley-VCH Verlag GmbH& Co. KGaA, Weinheim .
15. MacAdam, D.L. (1942) Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America*, **32**(5), 247–274.

16. Fairchild, M.D. (2001) A revision of CIECAM97s for practical applications. *Color Research & Application*, **26**, 418–427.
17. Moroney, N., Fairchild, M.D., Hunt, R.W.G., et al. (2002) *The CIECAM02 Color Appearance Model*, IS&T/SID CIC 10, Scottsdale, 23–27.
18. Colantoni, P., and Tréneau, A. (2003) *3d visualization of color data to analyze color images*, Proceedings of The PICS Conference, Rochester, NY, 500–505.
19. Hecht, E. and Zajac, A. (1974) *OPTICS*, Addison-Wesley, Menlo Park, pp. 311–323.
20. Theuwissen, A.J.P. (1995) *Solid-State Imaging with Charge-Coupled Devices*, Kluwer Academic Publishers, Dordrecht.
21. Janesick, J.R. (2001) *Scientific Charge-Coupled Devices*, SPIE Press, Bellingham.
22. Schroder, D.K. (1987) *Advanced MOS Devices*, Addison-Wesley Publishing Company, Reading.
23. <http://www.adorama.com/alc/0012961/article/FAQ-Whats-a-Backside-Illuminated-Sensor>
24. Huo, Y., Fesenmaier, C.C., and Catrysse, P.B. (2010) Microlens performance limits in sub-2 μm pixel CMOS image sensors. *Optics Express*, **18**(6), 5861.
25. Luther, R. (1927) Aus dem Gebiet der Farbreizmetrik. *Zeitschrift für Technische Physik*, **8**, 540–558 Translation: <http://iscc.org/pdf/LutherColorStimMetrics.pdf>.
26. Vora, P.L., Farrell, J.E., Tietz, J.D. and Brainard, D.H. (1998) Image capture: modelling and calibration of sensor responses and their synthesis from multispectral images, HPL-98-187, Hewlett Packard.
27. http://www.maxmax.com/spectral_response.htm
28. http://scien.stanford.edu/pages/labsite/2007/psych221/projects/07/camera_characterization/index.html
29. Sigernes, F., Dyrland, M., Peters, N., et al. (2009) The absolute sensitivity of digital colour cameras. *Optics Express*, **17**(22), 20211–20220.
30. Chikane, V. and Fuh, C.-S. (2006) Automatic white balance for digital still cameras. *Journal of Information Science and Engineering*, **22**, 497–509.
31. Huo, J., Chang, Y., and Wang, J., et al. (2006) Robust automatic white balance algorithm using gray color points in images. *IEEE Transactions on Consumer Electronics*, **52**(2), 541–546.
32. Lam, E.Y. and Fung, G.S.L. (2009) Automatic white balance in digital photography, Chapter 10, in *Single-Sensor Imaging* (ed. R. Lukac), CRC Press, Boca Raton, pp. 267–294.
33. Zapryanov, G., Ivanova, D., and Nikolova, I. (2012) Automatic white balance algorithms for digital still cameras – a comparative study. *Information Technologies and Control*, **1**, 16–22.
34. Lee, H.C. (1986) Method for computing the scene-illuminant from specular highlights. *Journal of the Optical Society of America A*, **3**, 1694–1699.
35. Imai, Y., Kato, Y., and Kaoi, H., Horiuchi, T. and Tominaga, S. (2011) Estimation of multiple illuminants based on specular highlight detection. Computational Color Imaging, Lecture Notes in Computer Science V 6626, Springer, Berlin, pp. 85–98.
36. Schwartz, M. S. (1996) Method for determining color of an illuminant in an image based on histogram data. US Patent 5495428 A.
37. Poynton, C.A. (2003) *Digital Video and HDTV: Algorithms and Interfaces*, Elsevier Science, San Francisco, pp. 233–240.
38. <http://www.cambridgeincolour.com/tutorials/sRGB-AdobeRGB1998.htm>
39. <http://www.quadibloc.com/other/cfaint.htm>
40. Holst, G.C. (1997) *Sampling, Aliasing, and Data Fidelity*, JCD Publishing, Winter Park, pp. 81–97.
41. http://www.digicamhistory.com/1980_1983.html
42. Kriss, M. (1998) *Tradeoff Between Aliasing Artifacts and Sharpness in Assessing Image Quality*. Proceedings of PICS IS&T Conference, May 17–20, pp. 247–256.
43. Topfer, K., Adams, J.E., and Keelan, B.W. (1998) *Modulation Transfer Function and Aliasing Patterns of CFA Interpolation Algorithms*. Proceedings of PICS IS&T Conference, May 17–20, pp. 367–370.
44. Kriss, M. (1996) *Color Filter Arrays for Digital Electronic Still Cameras*. Proceedings of IS&T 49th Annual Conference, p. 272.
45. Lukac, R., Plataniotis, K.N., Hatzinakos, D., and Smolka, D. (2004) *Improved CFA Interpolation Approach*. Proceeding of Second European Conference on Colour in Graphics, imaging, and Vision, April 5–8, pp. 303–310.
46. Alleyson, D. and Süsstrunk, S. (2004) *Spatio-Chromatic PCA of a Mosaiced Color Image*. Proceeding of Second European Conference on Colour in Graphics, Imaging, and Vision, April 5–8, pp. 311–314.
47. Alleyesson, D., Susstrunk, S., and Herault, J. (2002) *Color Demosaicing by Estimating Luminance and Opponent Chromatic Signals in the Fourier Domain*. Proceedings of the 10th IS&T Color Imaging Conference, November 12–15, pp. 331–336.

48. Lee, J.J., Tsai, Y.S., Liew, Y.C., et al. (2005) *Innovative Color Interpolation Using Fuzzy Logic and Linear Regression*. Proceedings of the 13th IS&T Color Imaging Conference, November 7–11, pp. 211–215.
49. Cok, D.R. (1994) *Reconstruction of CCD Images Using Template Matching*. Proceedings of the IS&T's Annual Conference/ICPS, pp. 380–385.
50. Kimmel, R. (1999) Demosaicing: image reconstruction from color CCD samples. *IEEE Transactions on Image Processing*, **8**(9), 1221–1228.
51. Ramanath, R., Snyder, W.E., and Bilbro, G.L. (2002) Demosaicing method for Bayer color arrays. *Journal of Electronic Imaging*, **11**(3), 306–315.
52. Lukac, G. and Plataniotis, K.N. (2005) Color filter arrays: design and performance. *IEEE Transactions on Consumer Electronics*, **51**(4), 1260–1267.
53. <http://www.kodak.com/global/en/service/professional/products/ekn017465.jhtml?pq-path=14015>
54. Greivenkamp, J.E. (1990) Color dependent optical prefilter for the suppression of aliasing artifacts. *Applied Optics*, **29**(5), 676–684.
55. Goodman, J.W. (1968) *Introduction to Fourier Optics*, McGraw Hill, New York, pp. 134–136.
56. Li, X., Gunturk, B.K., and Zhang, L. (2008) Image demosaicing: a systematic survey. *Proceedings of SPIE*, **6822**, 68221J-1–68221J-15.
57. Lee, W., Lee, S., and Kim, J. (2006) Cost-effective color filter array demosaicing using spatial correlation. *IEEE Transaction on Consumer Electronics*, **52**(2), 547–551.
58. Parulski, K.A. (1985) Color filters and processing alternatives for one-chip cameras. *IEEE Transactions on Electron Devices*, **32**(8), 1381–1389.