

Color conversion matrices in digital cameras: a tutorial

D. Andrew Rowlands*

Independent Consultant, Cambridge, United Kingdom

Abstract. Color conversion matrices and chromatic adaptation transforms (CATs) are of central importance when converting a scene captured by a digital camera in the camera raw space into a color image suitable for display using an output-referred color space. In this article, the nature of a typical camera raw space is investigated, including its gamut and reference white. Various color conversion strategies that are used in practice are subsequently derived and examined. The strategy used by internal image-processing engines of traditional digital cameras is shown to be based upon color rotation matrices accompanied by raw channel multipliers, in contrast to the approach used by smartphones and commercial raw converters, which is typically based upon characterization matrices accompanied by conventional CATs. Several advantages of the approach used by traditional digital cameras are discussed. The connections with the color conversion methods of the DCRaw open-source raw converter and the Adobe digital negative converter are also examined, along with the nature of the Adobe color and forward matrices. © 2020 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.OE.59.11.110801](https://doi.org/10.1117/1.OE.59.11.110801)]

Keywords: color matrix; white balance; chromatic adaptation; gamut; raw conversion; Adobe digital negative.

Paper 20201001T received Aug. 24, 2020; accepted for publication Oct. 27, 2020; published online Nov. 17, 2020.

1 Introduction

Consider converting a scene captured by a digital camera in the camera raw space into a digital image suitable for display using an output-referred color space. At the very least, there are two issues of fundamental importance that must be addressed when attempting to correctly reproduce the appearance of color. The first is that the response functions of digital cameras differ from those of the human visual system (HVS). A widely used approach to this issue is to consider color spaces as vector spaces and to account for the differences in response by introducing a color conversion matrix. A type of color conversion matrix that is commonly encountered is the 3×3 characterization matrix \underline{T} that defines the linear relationship between the camera raw space and the CIE XYZ reference color space:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \approx \underline{T} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}. \quad (1)$$

In general, camera raw spaces are not colorimetric, so the above conversion is approximate. The relationship can be optimized for a given illuminant by minimizing the color error. Significantly, this means that the optimum \underline{T} depends upon the nature of the scene illuminant,^{1,2} including its white point (WP). The characterization methodology for determining the optimum \underline{T} is described in Sec. 2.4, along with an illustration of how \underline{T} should be normalized in practice.

The second issue that must be addressed is the perception of the scene illumination WP. Although the various adaptation mechanisms employed by the HVS are complex and not fully understood, it is thought that the HVS naturally uses a chromatic adaptation mechanism to adjust its perception of the scene illumination WP to achieve color constancy under varying lighting conditions.^{3,4} Since digital camera sensors do not naturally adapt in this manner, incorrect white

*Address all correspondence to D. Andrew Rowlands, d.a.rowlands@physics.org

balance (WB) will arise when the WP of the scene illumination differs from the reference white of the output-referred color space used to encode the output image produced by the camera. As demonstrated in Sec. 3, digital cameras must attempt to emulate this chromatic adaptation mechanism by utilizing an appropriate chromatic adaptation transform (CAT).

As discussed in Sec. 4, modern smartphones and commercial raw converters typically calculate the optimum characterization matrix \underline{T} by interpolating between two preset characterization matrices according to an estimate of the scene illumination WP, and the CAT is implemented after applying \underline{T} . In traditional digital cameras, the color conversion is typically reformulated in terms of raw channel multipliers and color rotation matrices \underline{R} . This approach offers several advantages, as discussed in Sec. 5. A similar but computationally simpler approach is used by the DCRaw open-source raw converter, as discussed in Sec. 6. The open-source Adobe® digital negative (DNG) converter offers two color conversion methods, and the nature of the Adobe color matrices and forward matrices are discussed in Sec. 7. Finally, conclusions are drawn in Sec. 8.

2 Camera Raw Space

2.1 Gamut

The camera raw space for a given camera model arises from its set of spectral responsivity functions or camera response functions:

$$R_i(\lambda) = \text{QE}_i(\lambda) \frac{e\lambda}{hc}, \quad (2)$$

where e is the elemental charge, λ is the wavelength, h is Planck's constant, and c is the speed of light. The external quantum efficiency for mosaic i is defined by

$$\text{QE}_i(\lambda) = T_{\text{CFA},i}(\lambda)\eta(\lambda)T(\lambda)\text{FF}, \quad (3)$$

where $T_{\text{CFA},i}$ is the color filter array (CFA) transmittance function for mosaic i , $\eta(\lambda)$ is the charge collection efficiency or internal quantum efficiency of a photoelement, and $T(\lambda)$ is the SiO₂/Si interface transmittance function.⁵ The fill factor is defined by $\text{FF} = A_{\text{det}}/A_p$, where A_{det} is the photosensitive detection area at a photosite and A_p is the photosite area. The spectral passband of the camera should ideally correspond to the visible spectrum, so an infra-red blocking filter is required.

Analogous to the eye-cone response functions of the HVS, which can be interpreted as specifying the amounts of the eye cone primaries that the eye uses to sense color at a given λ , the camera response functions can be interpreted as specifying amounts of the camera raw space primaries at each λ . For example, the measured Nikon D700 camera response functions are shown in Fig. 1. However, a camera raw space is colorimetric only if the Luther-Ives condition is satisfied,⁷⁻⁹ meaning that the camera response functions must be an exact linear transformation of the eye-cone response functions, which are indirectly represented as a linear transformation from the CIE color-matching functions for the standard observer.

Although the eye-cone response functions are suited to capturing detail using the simple lens of the human eye, digital cameras use compound lenses that have been corrected for chromatic aberration. Consequently, camera response functions are designed with other considerations in mind.^{10,11} For example, better signal-to-noise performance is achieved by reducing the overlap of the response functions, which corresponds to a characterization matrix with smaller off-diagonal elements.¹⁰⁻¹² Indeed, minor color errors can be traded for better signal-to-noise performance.¹⁰⁻¹³ On the other hand, increased correlation in the wavelength dimension can improve the performance of the color demosaicing procedure.¹⁴ Due to such trade-offs along with filter manufacturing constraints, camera response functions are not exact linear transformations of the eye-cone response functions in practice. Consequently, camera raw spaces are not colorimetric, so cameras exhibit metameric error. Metamers are different spectral power distributions (SPDs) that are perceived by the HVS to be the same color when viewed under exactly the same conditions. Cameras that exhibit metameric error produce different color responses to these metamers.

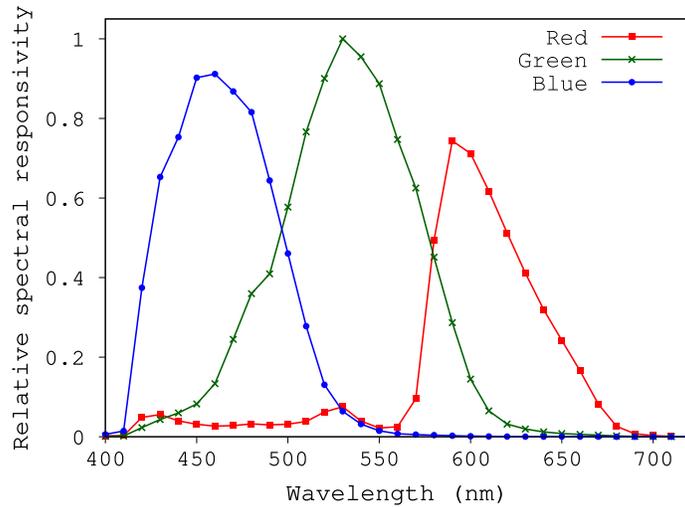


Fig. 1 Camera response functions for the Nikon D700 camera. The peak spectral responsivity has been normalized to unity. Data sourced from Ref. 6.

Camera metameric error can be determined experimentally and quantified by the digital still camera sensitivity metamerism index (DSC/SMI).^{8,15}

Figure 2 shows the spectral locus of the HVS on the xy chromaticity diagram, which is a 2D projection of the CIE XYZ color space that describes the relative proportions of the tristimulus values. Note that the spectral locus itself is horseshoe-shaped rather than triangular due to the fact

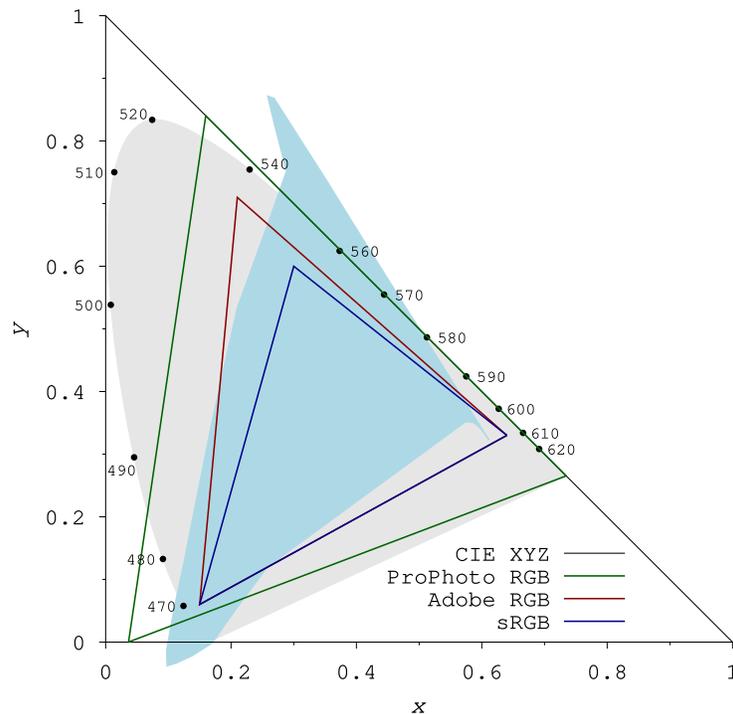


Fig. 2 Gamut of the camera raw space for the Nikon D700 (light-blue shaded area) plotted on the xy chromaticity diagram. The gamut is not a perfect triangle since the Luther-Ives condition is violated, which also explains why certain regions are pushed outside the triangle accessible to the CIE XYZ color space defined by the primaries located at (0,0), (0,1), and (1,0). The boundary of the horseshoe-shaped gray shaded area defines the spectral locus of the HVS. Saturation decreases with inward distance from the spectral locus. For comparison purposes, the (triangular) gamuts of several standard output-referred color spaces are indicated.

that overlap of the eye-cone response functions prevents the eye cones from being independently stimulated, so the chromaticities corresponding to (x, y) chromaticity coordinates positioned outside the spectral locus are invisible or imaginary as they are more saturated than pure spectrum colors. The gamut of the camera raw space for the Nikon D700 camera is also plotted on Fig. 2 and compared with several standard output-referred color spaces, namely sRGB,¹⁶ Adobe® RGB,¹⁷ and ProPhoto RGB.¹⁸ Due to the positions of the camera raw space primaries on the xy chromaticity diagram, certain regions of the camera raw space gamut do not reach the spectral locus of the HVS as these regions lie outside the triangular shape accessible to additive linear combinations of the three primaries. Furthermore, a notable consequence of camera metamerism error is that the camera raw space gamut is warped away from the triangular shape accessible to the additive linear combinations of the three primaries. Certain regions are even pushed outside of the triangle accessible to the CIE XYZ color space.¹⁹ See Ref. 19 for additional examples.

To determine the gamut of a camera raw space, the first step is to measure the camera response functions using a monochromator at a discrete set of wavelengths according to method A of the ISO 17321-1 standard.¹⁵ For each wavelength, the camera response functions yield raw \mathcal{RGB} relative tristimulus values in the camera raw space. The second step is to convert \mathcal{RGB} into relative CIE XYZ values by applying a characterization matrix that satisfies Eq. (1). Subsequently, the (x, y) chromaticity coordinates corresponding to the spectral locus of the camera raw space can be calculated using the usual formulas, $x = X/(X + Y + Z)$ and $y = Y/(X + Y + Z)$.

Since a given characterization matrix is optimized for use with the characterization illuminant, i.e., the scene illumination used to perform the characterization, another consequence of camera metamerism error is that the camera raw space gamut may vary according to the characterization matrix applied. The gamut of the Nikon D700 camera raw space shown in Fig. 2 was obtained using a characterization matrix optimized for CIE illuminant D65. Figure 3 shows how the gamut changes when a characterization matrix optimized for CIE illuminant A is applied instead.

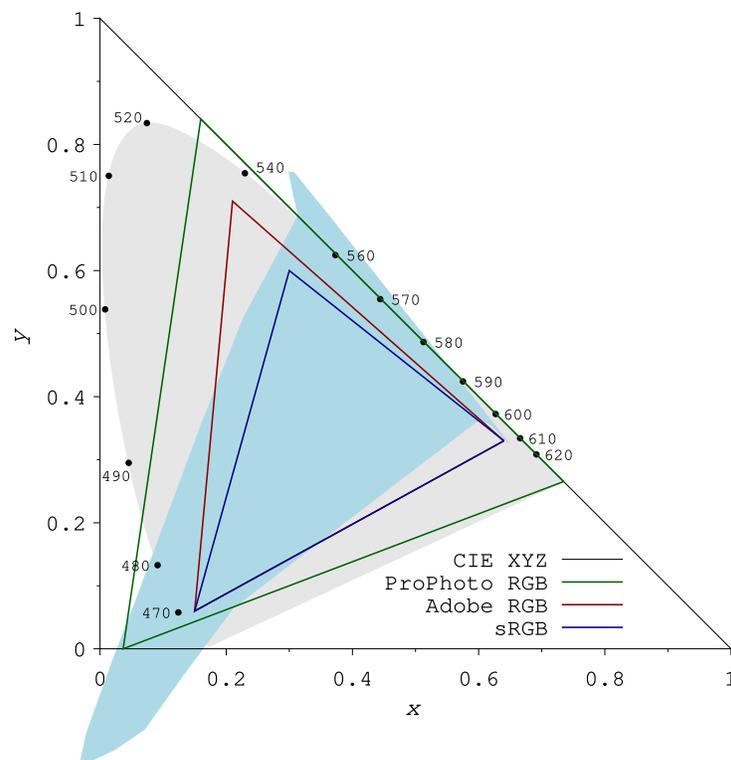


Fig. 3 Same as Fig. 2 except that a characterization matrix optimized for CIE illuminant A was used to obtain the camera raw space gamut rather than a characterization matrix optimized for CIE illuminant D65.

2.2 Raw Values

Color values in a camera raw space are expressed in terms of digital raw values for each raw color channel, which are analogous to tristimulus values in CIE color spaces. For a CFA that uses three types of color filters such as a Bayer CFA,²⁰ the raw values expressed using output-referred units, i.e., data/digital numbers (DN) or analog-to-digital units, belong to the following set of raw channels denoted here using calligraphic symbols:

$$\begin{bmatrix} n_{\text{DN},1} \\ n_{\text{DN},2} \\ n_{\text{DN},3} \\ n_{\text{DN},4} \end{bmatrix} = \begin{bmatrix} \mathcal{R} \\ \mathcal{G}_1 \\ \mathcal{G}_2 \\ \mathcal{B} \end{bmatrix}. \quad (4)$$

Although vector notation has been used here to represent a Bayer block, a true raw pixel vector is obtained only after the color demosaic has been performed, in which case there will be four raw values associated with each photosite. The Bayer CFA uses twice as many green filters as red and blue, which means that two values \mathcal{G}_1 and \mathcal{G}_2 associated with different positions in each Bayer block will be obtained in general. This is beneficial in terms of overall signal-to-noise ratio since photosites belonging to the green mosaics are more efficient in terms of photoconversion. Furthermore, the Bayer pattern is optimal in terms of reducing aliasing artifacts when three types of filters are arranged on a square grid.¹⁴ Although it is thought that a greater number of green filters provides enhanced resolution for the luminance signal since the standard 1924 CIE luminosity function for photopic vision peaks at 555 nm,²⁰ it has been argued that a Bayer CFA with two times more blue pixels than red and green would in fact be optimal for this purpose.¹⁴ When demosaicing raw data corresponding to a standard Bayer CFA, the final output will show false mazes or meshes if the ratio between \mathcal{G}_1 and \mathcal{G}_2 varies over the image.²¹ Software raw converters may average \mathcal{G}_1 and \mathcal{G}_2 together to eliminate such artifacts.²¹

Since there are fundamentally only three camera response functions, $R_1(\lambda)$, $R_2(\lambda)$, and $R_3(\lambda)$, color characterization for a Bayer CFA regards \mathcal{G}_1 and \mathcal{G}_2 as a single channel, \mathcal{G} . The raw values can be expressed as follows:

$$\mathcal{R} = k \int_{\lambda_1}^{\lambda_2} R_1(\lambda) \tilde{E}_{e,\lambda} d\lambda, \quad \mathcal{G} = k \int_{\lambda_1}^{\lambda_2} R_2(\lambda) \tilde{E}_{e,\lambda} d\lambda, \quad \mathcal{B} = k \int_{\lambda_1}^{\lambda_2} R_3(\lambda) \tilde{E}_{e,\lambda} d\lambda. \quad (5)$$

The camera response functions are defined by Eq. (2), the integration is over the spectral pass-band of the camera, $\tilde{E}_{e,\lambda}$ is the average spectral irradiance at the photosite, and k is a constant. Expressions for $\tilde{E}_{e,\lambda}$ and k are given in the [Appendix](#).

The actual raw values obtained in practice are quantized values modeled by taking the integer part of Eq. (5). When transforming from the camera raw space, it is useful to normalize the raw values to the range [0,1] by dividing Eq. (5) by the raw clipping point, which is the highest available DN.

2.3 Reference White

Using the above normalization, the reference white of a camera raw space is defined by the unit vector

$$\begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \quad (6)$$

Expressed in terms of CIE XYZ tristimulus values or (x, y) chromaticity coordinates with $Y = 1$, the reference white of a camera raw space is the WP of the scene illumination that yields maximum equal raw values for a neutral subject. (The WP of a SPD is defined by the CIE XYZ tristimulus values that correspond to a 100% neutral diffuse reflector illuminated by that SPD.)

It follows that the reference white of a camera raw space can in principle be determined experimentally by finding the illuminant that yields equal raw values for a neutral subject. Note that if the DCRaw open-source raw converter is used to decode the raw file, it is essential to disable WB. In terms of CIE colorimetry, the camera raw space reference white is formally defined by

$$\begin{bmatrix} X(\text{WP}) \\ Y(\text{WP}) \\ Z(\text{WP}) \end{bmatrix}_{\text{scene}} = \underline{T} \begin{bmatrix} \mathcal{R}(\text{WP}) = 1 \\ \mathcal{G}(\text{WP}) = 1 \\ \mathcal{B}(\text{WP}) = 1 \end{bmatrix}_{\text{scene}}, \quad (7)$$

where $Y(\text{WP}) = 1$ and the subscripts denote that the WP is that of the scene illumination. The 3×3 characterization matrix \underline{T} converts from the camera raw space to CIE XYZ and should be optimized for the required scene illumination. The optimum \underline{T} is unknown at this stage but can in principle be determined using the optimization procedure to be outlined in Sec. 2.4.

Although CIE color spaces use normalized units such that their reference whites correspond to WPs of CIE standard illuminants, camera raw spaces are not naturally normalized in such a manner. Consequently, the reference white of a camera raw space is not necessarily a neutral color as it is typically located far away from the Planckian locus and so does not necessarily have an associated correlated color temperature (CCT).

Note that a WP can be associated with a CCT provided its (x, y) chromaticity coordinates are sufficiently close to the Planckian locus, but there are many such coordinates that correspond to the same CCT. To distinguish between them, a D_{uv} value, informally referred to as a color tint, can be assigned.²² This is determined by converting (x, y) into (u, v) chromaticity coordinates on the CIE 1960 UCS chromaticity diagram,^{23,24} where isotherms are normal to the Planckian locus. In this representation, CCT is a valid concept only for (u, v) coordinates positioned a distance from the Planckian locus that is within $D_{uv} = \pm 0.05$ along an isotherm.²⁵

To see that the reference white of a camera raw space is far from the Planckian locus, consider the Nikon D700 raw values for a neutral diffuse reflector illuminated by CIE illuminants A and D65, respectively,

$$\begin{bmatrix} \mathcal{R}(\text{WP}) = 0.8878 \\ \mathcal{G}(\text{WP}) = 1.0000 \\ \mathcal{B}(\text{WP}) = 0.4017 \end{bmatrix}_A = \underline{T}_A^{-1} \begin{bmatrix} X(\text{WP}) = 1.0985 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 0.3558 \end{bmatrix}_A$$

$$\begin{bmatrix} \mathcal{R}(\text{WP}) = 0.4514 \\ \mathcal{G}(\text{WP}) = 1.0000 \\ \mathcal{B}(\text{WP}) = 0.8381 \end{bmatrix}_{D65} = \underline{T}_{D65}^{-1} \begin{bmatrix} X(\text{WP}) = 0.9504 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 1.0888 \end{bmatrix}_{D65}, \quad (8)$$

where \underline{T}_A and \underline{T}_{D65} are example characterization matrices optimized for CIE illuminants A and D65, respectively. As shown in Fig. 4, the WPs of these standard illuminants are very close to the Planckian locus. Illuminant A has CCT = 2856 K and $D_{uv} = 0.0$, and illuminant D65 has CCT = 6504 K and $D_{uv} = 0.0032$. Evidently, the above Nikon D700 raw values are very different from the camera raw space unit vector, and it would be necessary to apply large multipliers to the red and blue raw pixel values in both cases. These multipliers are known as raw channel multipliers as they are typically applied to the red and blue raw channels before the color demosaic as part of the color conversion strategy used by the internal image-processing engines of traditional digital cameras.

An estimate of the Nikon D700 reference white can be obtained by approximating Eq. (7) using a readily available characterization matrix in place of \underline{T} . Applying \underline{T}_A yields $(x, y) = (0.3849, 0.3058)$, which corresponds to $D_{uv} = -0.0378$. This has an associated CCT = 3155 K as the D_{uv} value is just within the allowed limit, but Fig. 4 shows that the color tint is a strong magenta. This is true of typical camera raw spaces in general.²¹ A similar estimate for the Olympus E-M1 camera yields $(x, y) = (0.3599, 0.2551)$, which corresponds to $D_{uv} = -0.0637$. This does not have an associated CCT, and the color tint is a very strong magenta.

Although the fact that camera raw space reference whites are not neutral in terms of CIE colorimetry has no bearing on the final reproduced image, it will be shown in Sec. 5 that the

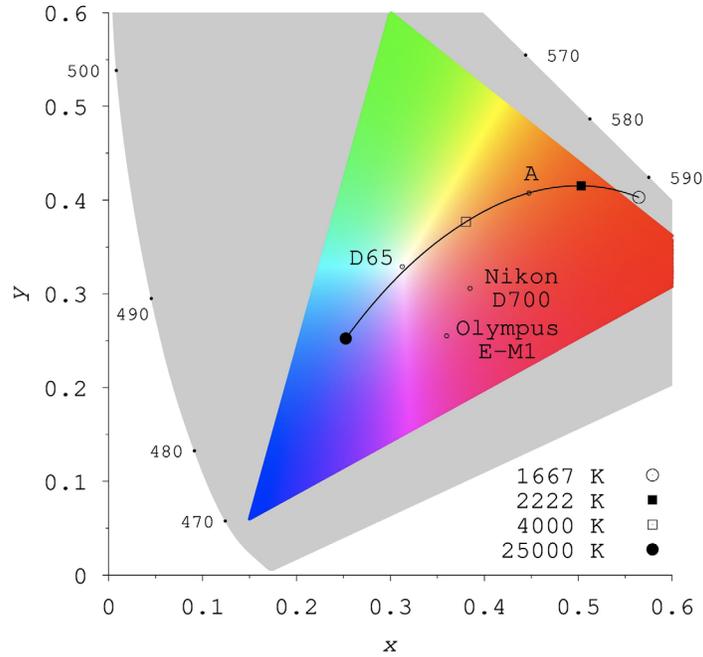


Fig. 4 Estimated reference whites of the Nikon D700 and Olympus E-M1 camera raw spaces in relation to the WPs of CIE illuminants A and D65. The Planckian locus is represented by the black curve. Only visible chromaticities contained within the sRGB color space are shown in color.

camera raw space reference white is utilized as a useful intermediary step in the color conversion strategy used by traditional digital cameras.

2.4 Camera Color Characterization

Recall the linear transformation from the camera raw space to CIE XYZ defined by Eq. (1):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \approx \underline{T} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix},$$

where \underline{T} is a 3×3 characterization matrix:

$$\underline{T} = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix}. \tag{9}$$

The color conversion is approximate since the Luther-Ives condition is not satisfied exactly. As mentioned in the introduction, \underline{T} can be optimized for the characterization illuminant, i.e., the scene illumination used to perform the characterization.^{1,2} The optimum matrix \underline{T} is dependent upon the SPD itself, but it largely depends upon the characterization illumination WP provided the illuminant is representative of a real world SPD.

Characterization matrices optimized for known illuminants can be determined by color-error minimization procedures based upon photographs taken of a standard color chart.² Although various minimization techniques have been developed, including WP-preserving techniques,²⁶ the procedure discussed below is based on the standardized method B of ISO 17321-1.¹⁵

Note that ISO 17321-1 uses processed images output by the camera rather than raw data and consequently requires inversion of the camera opto-electronic conversion function (OECF).²⁷ The OECF defines the nonlinear relationship between irradiance at the sensor plane and the digital output levels of a viewable output image such as a JPEG file produced by the camera.

To bypass the need to experimentally determine the OECF, a variation of method B from ISO 17321-1 is described below. This method uses the DCRaw open-source raw converter to decode the raw file so that the raw data can be used directly.^{28,29}

1. Take a photograph of a color chart illuminated by a specified illuminant. Since the raw values scale linearly, only their relative values are important. However, the f -number N and exposure duration t should be chosen so as to avoid clipping.
2. Calculate relative XYZ tristimulus values for each patch of the color chart:

$$X = k \int_{\lambda_1}^{\lambda_2} \bar{x}(\lambda) E_{e,\lambda} R(\lambda) d\lambda \quad Y = k \int_{\lambda_1}^{\lambda_2} \bar{y}(\lambda) E_{e,\lambda} R(\lambda) d\lambda \quad Z = k \int_{\lambda_1}^{\lambda_2} \bar{z}(\lambda) E_{e,\lambda} R(\lambda) d\lambda, \quad (10)$$

where $E_{e,\lambda}$ is the spectral irradiance incident at the color chart measured using a spectrometer; $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ are the color-matching functions of the CIE XYZ color space; and the integration is discretized into a sum with a 10-nm increment and limits $\lambda_1 = 380$ nm and $\lambda_2 = 780$ nm. Unless a tristimulus colorimeter is used, the calculation requires knowledge of the spectral reflectance of each patch. Spectral reflectance has been denoted by $R(\lambda)$ in the above equations, and this should not be confused with the camera response functions. The normalization constant k can be chosen so that Y is in the range $[0,1]$ using the white patch as a white reference.

3. Obtain a linear demosaiced output image directly in the camera raw space without converting to any other color space. Gamma encoding, tone curves, and WB must all be disabled. Since the present method bypasses the need to determine and invert the OECF, it is crucial to disable WB; otherwise, raw channel multipliers may be applied to the raw channels. If using the DCRaw open-source raw converter, an appropriate command is

`dcraw -v -r 1 1 1 1 -o 0 -4 -T filename.`

This yields a 16-bit linear demosaiced output TIFF file in the camera raw space. If working with raw channels rather than demosaiced raw pixel vectors, an appropriate command is

`dcraw -v -D -4 -T filename.`

The above DCRaw commands are explained in Table 3.

4. Measure average \mathcal{R} , \mathcal{G} , and \mathcal{B} values over a 64×64 block of pixels at the center of each patch. Each patch can then be associated with an appropriate average raw pixel vector.
5. Build a $3 \times n$ matrix \underline{A} containing the vectors of the XYZ color space for each patch $1, \dots, n$ as columns:

$$\underline{A} = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \\ Y_1 & Y_2 & \cdots & Y_n \\ Z_1 & Z_2 & \cdots & Z_n \end{bmatrix}. \quad (11)$$

Similarly, build a $3 \times n$ matrix \underline{B} containing the corresponding raw pixel vectors as columns:

$$\underline{B} = \begin{bmatrix} \mathcal{R}_1 & \mathcal{R}_2 & \cdots & \mathcal{R}_n \\ \mathcal{G}_1 & \mathcal{G}_2 & \cdots & \mathcal{G}_n \\ \mathcal{B}_1 & \mathcal{B}_2 & \cdots & \mathcal{B}_n \end{bmatrix}. \quad (12)$$

6. Estimate the 3×3 characterization matrix \underline{T} that transforms \underline{B} to \underline{A} :

$$\underline{A} \approx \underline{T} \underline{B}. \quad (13)$$

A preliminary solution is obtained using linear least-squares minimization:^{2,15}

$$\underline{T} = \underline{A} \underline{B}^T (\underline{B} \underline{B}^T)^{-1}, \quad (14)$$

where the T superscript denotes the transpose operator.

7. Use the preliminary \underline{T} estimate to calculate a new set of estimated CIE XYZ tristimulus values \underline{A}' according to Eq. (13). Transform \underline{A} and \underline{A}' to the perceptually uniform CIE LAB reference color space, and calculate the color difference ΔE_i between the estimated tristimulus values and real tristimulus values for each patch i . The set $\{\Delta E_i\}$ can be used to calculate the DSC/SMI.^{8,15} Note that, to satisfy the Luther-Ives condition exactly, it is necessary that $\underline{A}' = \underline{A}$, in which case a DSC/SMI score of 100 would be obtained.
8. Optimize \underline{T} by minimizing $\{\Delta E_i\}$ using the nonlinear optimization technique recommended by ISO 17321-1. The final DSC/SMI defines the final potential color error. Ideally, include a constraint that preserves the characterization illuminant WP.
9. Scale the final \underline{T} according to the normalization required for its practical implementation. This is discussed below.

Provided WB was disabled in step 3, the characterization matrix \underline{T} can be used with arbitrary scene illumination. However, optimum results will be obtained for scene illumination with a WP that closely matches that of the characterization illuminant.

Figure 5 shows how the matrix elements of an optimized characterization matrix vary as a function of characterization illuminant CCT for the Olympus E-M1 camera.

For the same camera, Fig. 6(a) shows a photo of a color chart in the camera raw space taken under D65 illumination. When the camera raw space \mathcal{RGB} values are interpreted as RGB values in the sRGB color space for display purposes without any color characterization matrix applied, a strong green color tint is revealed, which arises from the greater transmission of the green Bayer filter. Figure 6(b) shows the same photo converted into the sRGB color space by applying an optimized characterization matrix \underline{T} , followed by a matrix that converts the colors from the CIE XYZ color space to sRGB. Evidently, the colors are now displayed correctly.

2.5 Characterization Matrix Normalization

Normalization of a characterization matrix refers to scaling of the entire matrix so that all matrix elements are scaled identically. A typical normalization applied in practice is to ensure that the matrix maps between the characterization illuminant WP expressed using the CIE XYZ color space and the camera raw space such that the raw data just saturates when a 100% neutral diffuse reflector is photographed under the characterization illuminant. The green raw channel is typically the first to saturate.

For example, if the characterization illuminant is D65, then \underline{T} can be normalized such that its inverse provides the following mapping:

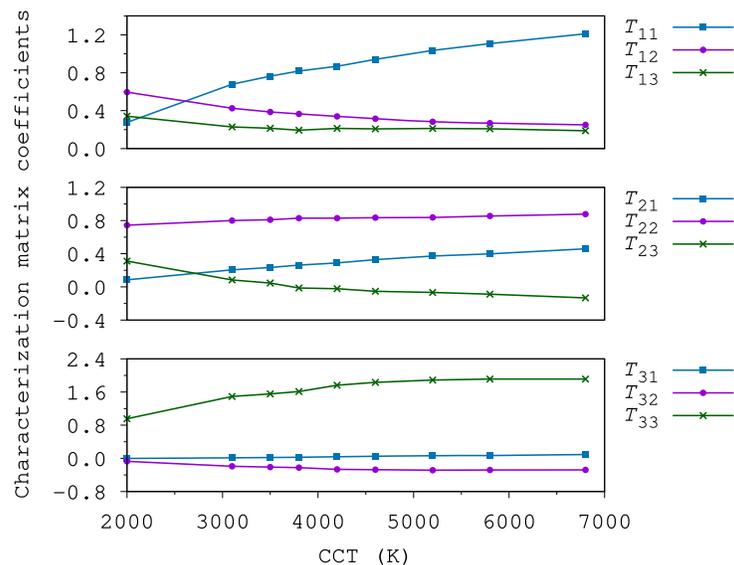


Fig. 5 Variation of the matrix elements of a characterization matrix for the Olympus E-M1 camera as a function of characterization illuminant CCT.

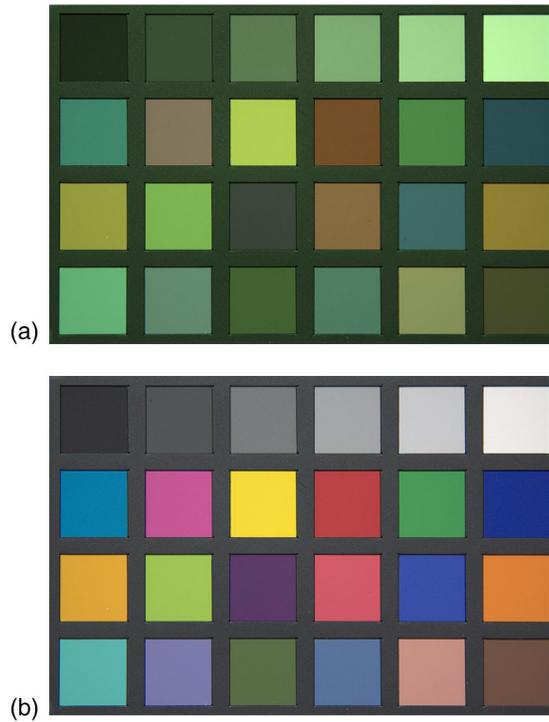


Fig. 6 (a) Photo of a color chart in the camera raw space taken under D65 illumination. (b) The same photo converted to the sRGB color space.

$$\begin{bmatrix} \mathcal{R}(\text{WP}) \\ \mathcal{G}(\text{WP}) \\ \mathcal{B}(\text{WP}) \end{bmatrix}_{\text{D65}} = \underline{T}^{-1} \begin{bmatrix} X(\text{WP}) = 0.9504 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 1.0888 \end{bmatrix}_{\text{D65}}, \quad (15)$$

where $\max\{\mathcal{R}(\text{WP}), \mathcal{G}(\text{WP}), \mathcal{B}(\text{WP})\} = 1$. Since the green raw channel is typically the first to saturate under most types of illumination, it will typically be the case that $\mathcal{G}(\text{WP}) = 1$, whereas $\mathcal{R}(\text{WP}) < 1$ and $\mathcal{B}(\text{WP}) < 1$.

For example, the Olympus E-M1 characterization matrices used by Fig. 5 for the 4200 and 6800 K calibration illuminants are defined by

$$\underline{T}_{4200 \text{ K}} = \begin{bmatrix} 0.8680 & 0.3395 & 0.2133 \\ 0.2883 & 0.8286 & -0.0216 \\ 0.0425 & -0.2647 & 1.7637 \end{bmatrix}, \quad \underline{T}_{6800 \text{ K}} = \begin{bmatrix} 1.2105 & 0.2502 & 0.1882 \\ 0.4586 & 0.8772 & -0.1328 \\ 0.0936 & -0.2788 & 1.9121 \end{bmatrix}. \quad (16)$$

These matrices are normalized such that the WP of the characterization illuminant maps to raw values where the green raw channel just reaches saturation:

$$\begin{bmatrix} \mathcal{R}(\text{WP}) = 0.6337 \\ \mathcal{G}(\text{WP}) = 1.0000 \\ \mathcal{B}(\text{WP}) = 0.5267 \end{bmatrix}_{4200 \text{ K}} = \underline{T}_{4200 \text{ K}}^{-1} \begin{bmatrix} X(\text{WP}) = 1.0019 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 0.6911 \end{bmatrix}_{4200 \text{ K}},$$

$$\begin{bmatrix} \mathcal{R}(\text{WP}) = 0.4793 \\ \mathcal{G}(\text{WP}) = 1.0000 \\ \mathcal{B}(\text{WP}) = 0.7312 \end{bmatrix}_{6800 \text{ K}} = \underline{T}_{6800 \text{ K}}^{-1} \begin{bmatrix} X(\text{WP}) = 0.9682 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 1.1642 \end{bmatrix}_{6800 \text{ K}}. \quad (17)$$

3 White Balance

A remarkable property of the HVS is its ability to naturally adjust to the ambient lighting conditions. For example, if a 100% neutral diffuse reflector is placed in a photographic scene illuminated by daylight, the reflector appears to be neutral white. Later in the day when there is a change in the chromaticity or CCT of the scene illumination, the color of the reflector would be expected to change accordingly. However, the reflector will continue to appear neutral white. In other words, the perceived color of objects remains relatively constant under varying types of scene illumination, which is known as color constancy.^{3,4}

The chromatic adaptation mechanism by which the HVS achieves color constancy is complex and not fully understood, but a simplified explanation is that the HVS aims to discount the chromaticity of the illuminant.³⁰ Back in 1902, von-Kries postulated that this is achieved by an independent scaling of each eye cone response function.^{3,4} The color stimulus that an observer adapted to the ambient conditions considers to be neutral white (perfectly achromatic with 100% relative luminance) is defined as the adapted white.³¹

Since camera response functions do not naturally emulate the HVS by discounting the chromaticity of the scene illumination, an output image will appear too warm or too cold if it is displayed using illumination with a WP that does not match the adapted white for the photographic scene at the time the photograph was taken. This is known as incorrect WB. The issue can be solved by implementing the following computational strategy.

1. Inform the camera of the adapted white before taking a photograph. Due to the complex dependence of the true adapted white on the ambient conditions, this task is replaced by a simpler one in practice, namely to identify the scene illumination WP. For example, a WB preset corresponding to the scene illumination can be manually selected, a scene illumination CCT estimate can be manually entered, or the camera can compute its own estimate by analyzing the raw data using the automatic WB function. In all cases, the camera estimate for the scene illumination WP is known as the camera neutral³² or adopted white (AW).³¹ (This illuminant estimation step should not be confused with WB. Illuminant estimation refers to the computational approaches used by the automatic WB function to estimate the scene illumination WP. A very simple illuminant estimation approach is the “gray world” method,³³ which assumes that the average of all of the scene colors will turn out to be achromatic. Another simple approach is to assume that the brightest white is likely to correspond to the scene illumination WP.³⁴ However, practical illuminant estimation algorithms are much more sophisticated.^{35,36})
2. Choose a standard reference white that will be used when displaying the output image. If the image will be displayed using a standard output-referred color space such as sRGB, the chosen reference white will be that of the output-referred color space, which is CIE illuminant D65 in the case of sRGB.
3. Chromatically adapt the image colors by adapting the scene illumination WP estimate (the AW) so that it becomes the reference white of the chosen output-referred color space. This white balancing step is achieved by applying a CAT.

The CAT needs to be applied as part of the overall color conversion from the camera raw space to the chosen output-referred color space. Different approaches for combining these components exist. The typical approach used in color science is to convert from the camera raw space to CIE XYZ, apply the CAT, and then convert to the chosen output-referred color space. In the case of sRGB,

$$\begin{bmatrix} R_L \\ G_L \\ B_L \end{bmatrix}_{D65} = \underline{M}_{sRGB}^{-1} \underline{CAT}_{AW \rightarrow D65} \underline{T} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{scene}, \quad (18)$$

where \underline{T} is a characterization matrix that converts from the camera raw space to CIE XYZ and is optimized for the scene AW, the matrix $\underline{CAT}_{AW \rightarrow D65}$ applied in the CIE XYZ color space is a CAT that adapts the AW to the D65 reference white of the sRGB color space, and finally

$\underline{M}_{\text{sRGB}}^{-1}$ is the matrix that converts from CIE XYZ to the linear form of the sRGB color space:

$$\underline{M}_{\text{sRGB}}^{-1} = \begin{bmatrix} 3.2410 & -1.5374 & -0.4986 \\ -0.9692 & 1.8760 & 0.0416 \\ 0.0556 & -0.2040 & 1.0570 \end{bmatrix}. \quad (19)$$

In particular, the AW in the camera raw space is mapped to the reference white of the output-referred color space defined by the unit vector in the output-referred color space:

$$\begin{bmatrix} R_L = 1 \\ G_L = 1 \\ B_L = 1 \end{bmatrix}_{\text{D65}} = \underline{M}_{\text{sRGB}}^{-1} \underline{\text{CAT}}_{\text{AW} \rightarrow \text{D65}} \underline{T} \begin{bmatrix} \mathcal{R}(\text{AW}) \\ \mathcal{G}(\text{AW}) \\ \mathcal{B}(\text{AW}) \end{bmatrix}_{\text{scene}}. \quad (20)$$

When the encoded output image is viewed on a calibrated display monitor, a scene object that the HVS regarded as being white at the time the photograph was taken will now be displayed using the D65 reference white. Ideally, the ambient viewing conditions should match those defined as appropriate for viewing the sRGB color space.

If the scene illumination WP estimate is far from the true scene illumination WP, then incorrect WB will be evident to the HVS. If the scene illumination CCT estimate is higher than the true CCT, then the photo will appear too warm. Conversely, if the scene illumination CCT estimate is lower than the true CCT, then the photo will appear too cold.

Figure 7(a) shows a photo of a color chart taken under 2700 K CCT tungsten illumination using the Olympus E-M1 camera. A characterization matrix \underline{T} was applied to convert the colors into CIE XYZ, followed by $\underline{M}_{\text{sRGB}}^{-1}$ to convert the colors to sRGB. Evidently, the true color of the scene illumination is revealed since no chromatic adaptation has been performed by the camera. In other words, the photo appears too warm in relation to the D65 reference white of the sRGB



Fig. 7 (a) Photo of a color chart taken under 2700 K CCT tungsten illumination and converted to the sRGB color space for display without any chromatic adaptation. (b) White balanced photo obtained by including a CAT to adapt the scene illumination WP to the D65 reference white of the sRGB color space.

color space. Figure 7(b) shows the same photo after white balancing by including a CAT that chromatically adapts the scene illumination WP to the sRGB color space D65 reference white, which has a 6504 K CCT and $D_{uv} = 0.0032$ color tint.

3.1 Chromatic Adaptation Transforms

A CAT is a computational technique for adjusting the WP of a given SPD. It achieves this goal by attempting to mimic the chromatic adaptation mechanism of the HVS. In the context of digital cameras, the most important CATs are the Bradford CAT and raw channel scaling.

In 1902, von-Kries postulated that the chromatic adaptation mechanism be modeled as an independent scaling of each eye cone response function,^{3,4} which is equivalent to scaling the L , M , and S tristimulus values in the LMS color space. To illustrate the von-Kries CAT, consider adapting the scene illumination WP estimate (the AW) to the WP of D65 illumination:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{\text{D65}} = \underline{\text{CAT}}_{\text{AW} \rightarrow \text{D65}} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{\text{scene}}. \quad (21)$$

In this case, the von-Kries CAT that must be applied to all raw pixel vectors can be written as

$$\underline{\text{CAT}}_{\text{AW} \rightarrow \text{D65}} = \underline{M}_{\text{vK}}^{-1} \begin{bmatrix} \frac{L(\text{D65})}{L(\text{AW})} & 0 & 0 \\ 0 & \frac{M(\text{D65})}{M(\text{AW})} & 0 \\ 0 & 0 & \frac{S(\text{D65})}{S(\text{AW})} \end{bmatrix} \underline{M}_{\text{vK}}. \quad (22)$$

The matrix $\underline{M}_{\text{vK}}$ transforms each raw pixel vector into a diagonal matrix in the LMS color space. Modern forms of $\underline{M}_{\text{vK}}$ include matrices based on the cone fundamentals defined by the CIE in 2006³⁷ and the Hunt–Pointer–Estevez transformation matrix³⁸ defined by

$$\underline{M}_{\text{vK}} = \begin{bmatrix} 0.38971 & 0.68898 & -0.07868 \\ -0.22981 & 1.18340 & 0.04641 \\ 0.00000 & 0.00000 & 1.00000 \end{bmatrix}. \quad (23)$$

After applying $\underline{M}_{\text{vK}}$, the L , M , and S values are independently scaled according to the von-Kries hypothesis. In the present example, the scaling factors arise from the ratio between the AW and D65 WPs. These can be obtained from the following WP vectors:

$$\begin{bmatrix} L(\text{AW}) \\ M(\text{AW}) \\ S(\text{AW}) \end{bmatrix} = \underline{M}_{\text{vK}} \begin{bmatrix} X(\text{WP}) \\ Y(\text{WP}) \\ Z(\text{WP}) \end{bmatrix}_{\text{scene}} \\ \begin{bmatrix} L(\text{D65}) \\ M(\text{D65}) \\ S(\text{D65}) \end{bmatrix} = \underline{M}_{\text{vK}} \begin{bmatrix} X(\text{WP}) = 0.9504 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 1.0888 \end{bmatrix}_{\text{D65}}. \quad (24)$$

Finally, the inverse of the transformation matrix $\underline{M}_{\text{vK}}$ is applied to convert each raw pixel vector back into the CIE XYZ color space.

The Bradford CAT³⁹ can be regarded as an improved version of the von-Kries CAT. A simplified linearized version is recommended by the ICC for use in digital imaging.⁴⁰ The linear Bradford CAT can be implemented in an analogous fashion as the von-Kries CAT, the difference being that the L , M , and S tristimulus values are replaced by ρ , γ , and β , which correspond to a “sharpened” artificial eye cone space. The transformation matrix is defined by

$$\underline{M}_{\text{BFD}} = \begin{bmatrix} 0.8951 & 0.2664 & -0.1614 \\ -0.7502 & 1.7135 & 0.0367 \\ 0.0389 & -0.0685 & 1.0296 \end{bmatrix}. \quad (25)$$

Analogous to the independent scaling of the eye cone response functions hypothesized by von-Kries, a type of CAT can be applied in the camera raw space by directly scaling the raw channels. Consider a Bayer block for the AW obtained by photographing a 100% neutral diffuse reflector under the scene illumination. The following operation will adapt the AW to the reference white of the camera raw space:

$$\begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{\text{RW}} = \underline{\text{CAT}}_{\text{AW} \rightarrow \text{RW}} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{\text{scene}}, \quad (26)$$

where

$$\underline{\text{CAT}}_{\text{AW} \rightarrow \text{RW}} = \underline{D} = \begin{bmatrix} \frac{1}{\mathcal{R}(\text{AW})} & 0 & 0 \\ 0 & \frac{1}{\mathcal{G}(\text{AW})} & 0 \\ 0 & 0 & \frac{1}{\mathcal{B}(\text{AW})} \end{bmatrix}_{\text{scene}}. \quad (27)$$

The diagonal scaling factors, known as raw channel multipliers, can be obtained directly from the raw data using the AW calculated by the camera. For example, AW = D65 for D65 scene illumination, in which case

$$\underline{\text{CAT}}_{\text{D65} \rightarrow \text{RW}} = \underline{D}_{\text{D65}} = \begin{bmatrix} \frac{1}{\mathcal{R}(\text{D65})} & 0 & 0 \\ 0 & \frac{1}{\mathcal{G}(\text{D65})} & 0 \\ 0 & 0 & \frac{1}{\mathcal{B}(\text{D65})} \end{bmatrix}, \quad (28)$$

where $\mathcal{R}(\text{D65})$, $\mathcal{G}(\text{D65})$, and $\mathcal{B}(\text{D65})$ are extracted from the Bayer block for a 100% neutral diffuse reflector photographed under D65 scene illumination.

In the context of digital cameras, the type of CAT defined by raw channel multipliers has been found to work better in practice, particularly for extreme cases.^{21,32} A reason for this is that the raw channel multipliers are applied in the camera raw space prior to application of a color conversion matrix. The camera raw space corresponds to a physical capture device, but CATs such as the linear Bradford CAT are applied in the CIE XYZ color space after applying a color conversion matrix that contains error. In particular, color errors that have been minimized in a nonlinear color space such as CIE LAB will be unevenly amplified, so the color conversion will no longer be optimal.⁴¹

4 Smartphone Cameras

Smartphone manufacturers, along with commercial raw conversion software developers, typically implement the conventional type of computational color conversion strategy used in color science that was introduced in Sec. 3. Since the camera raw space is transformed into CIE XYZ as the first step, image processing techniques can be applied in the CIE XYZ color space (or following a transformation into some other intermediate color space) before the final transformation to an output-referred RGB color space.

Consider the white-balanced transformation from the camera raw space to an output-referred RGB color space. Unlike in traditional digital cameras, the color demosaic is typically carried out first, so the vector notation used for the camera raw space below refers to raw pixel vectors

rather than Bayer blocks. In the case of sRGB, the transformation that must be applied to every raw pixel vector is defined by

$$\begin{bmatrix} R_L \\ G_L \\ B_L \end{bmatrix}_{D65} = \underline{M}_{sRGB}^{-1} \underline{CAT}_{AW \rightarrow D65} \underline{T} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{scene}. \quad (29)$$

The conversion can be decomposed into three steps.

1. After the camera has estimated the scene illumination WP (the AW), a characterization matrix \underline{T} optimized for the AW that converts from the camera raw space to CIE XYZ is applied:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{scene} = \underline{T} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{scene}. \quad (30)$$

The optimized matrix \underline{T} is typically normalized such that the AW in the CIE XYZ space is obtained when the raw pixel vector corresponding to a neutral diffuse reflector illuminated by the AW just reaches saturation:

$$\begin{bmatrix} X(AW) \\ Y(AW) \\ Z(AW) \end{bmatrix}_{scene} = \underline{T} \begin{bmatrix} \mathcal{R}(AW) \\ \mathcal{G}(AW) \\ \mathcal{B}(AW) \end{bmatrix}_{scene}, \quad (31)$$

where $Y(WP) = 1$ and $\max\{\mathcal{R}(AW), \mathcal{G}(AW), \mathcal{B}(AW)\} = 1$. As discussed in Sec. 2.5, the green component is typically the first to saturate, so $\mathcal{R}(AW) < 1$ and $\mathcal{B}(AW) < 1$ in general.

2. Since \underline{T} does not alter the AW, a CAT is applied to achieve WB by adapting the AW to the reference white of the chosen output-referred color space. This is D65 in the case of sRGB:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{D65} = \underline{CAT}_{AW \rightarrow D65} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{scene}. \quad (32)$$

The ICC recommends implementing the CAT using the linear Bradford CAT matrix defined by Eq. (25).

3. A matrix that converts from CIE XYZ to the linear form of the chosen output-referred color space is applied. In the case of sRGB,

$$\begin{bmatrix} R_L \\ G_L \\ B_L \end{bmatrix}_{D65} = \underline{M}_{sRGB}^{-1} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{D65}. \quad (33)$$

Finally, the digital output levels of the output image are determined by applying the nonlinear gamma encoding curve of the output-referred color space and reducing the bit depth to 8. In modern digital imaging, encoding gamma curves are designed to minimize visible banding artifacts when the bit depth is reduced, and the non-linearity introduced is later reversed by the display gamma.²⁸

To see that WB is correctly achieved, the above steps can be followed for the specific case of the raw pixel vector that corresponds to the AW. As required by Eq. (20), it is found that this maps to the reference white of the output-referred color space defined by the unit vector in that color space:

$$\begin{bmatrix} R_L = 1 \\ G_L = 1 \\ B_L = 1 \end{bmatrix}_{D65} = \underline{M}_{sRGB}^{-1} \underline{CAT}_{AW \rightarrow D65} \underline{T} \begin{bmatrix} \mathcal{R}(AW) \\ \mathcal{G}(AW) \\ \mathcal{B}(AW) \end{bmatrix}_{scene}.$$

Although the matrix transformation defined by Eq. (29) appears to be straightforward, the characterization matrix \underline{T} should in principle be optimized for the AW. However, it is impractical to determine a characterization matrix optimized for each possible scene illumination WP that could occur. For example, if CCTs are specified to the nearest Kelvin and color tint is neglected, then 12,000 matrices would be required to cover scene illumination WPs from 2000 to 14,000 K.

The computationally simplest solution used on some mobile phone cameras is to approximate the optimized characterization matrix \underline{T} using a single fixed matrix optimized for a representative illuminant. For example, this could be D65 illumination, in which case \underline{T} optimized for the AW is approximated as \underline{T}_{D65} . The drawback of this very simple approach is that the color conversion loses some accuracy when the scene illumination WP differs significantly from the WP of the representative illuminant.

As described below, an advanced solution to the problem is to adopt the type of approach used by the Adobe DNG converter.³² The idea is to interpolate between two preset characterization matrices that are optimized for use with either a low-CCT or high-CCT illuminant. For a given scene illumination, an interpolated matrix optimized for the CCT of the AW can be determined.

4.1 Interpolation Algorithm

If using the advanced approach mentioned above, the optimized characterization matrix \underline{T} required by Eq. (29) can be calculated by interpolating between two characterization matrices \underline{T}_1 and \underline{T}_2 based on the scene illumination CCT estimate denoted by $\text{CCT}(\text{AW})$, together with the CCTs of the two characterization illuminants denoted by CCT_1 and CCT_2 , respectively, with $\text{CCT}_1 < \text{CCT}_2$. For example, illuminant 1 could be a low-CCT illuminant such as CIE illuminant A, whereas illuminant 2 could be a high-CCT illuminant such as D65.

The first step is to appropriately normalize \underline{T}_1 and \underline{T}_2 . Although characterization matrices are typically normalized according to their corresponding characterization illuminant WPs as demonstrated in Sec. 2.5, it is more convenient to normalize \underline{T}_1 and \underline{T}_2 according to a common WP when implementing an interpolation algorithm. Unfortunately, the AW cannot be expressed using the CIE XYZ color space at this stage since \underline{T} is yet to be determined. Instead, the common WP could be chosen to be the reference white of the output-referred color space, which is D65 for sRGB. In this case, \underline{T}_1 and \underline{T}_2 should both be scaled according to Eq. (15):

$$\begin{aligned} \begin{bmatrix} \mathcal{R}(\text{WP}) \\ \mathcal{G}(\text{WP}) \\ \mathcal{B}(\text{WP}) \end{bmatrix}_{D65} &= \underline{T}_1^{-1} \begin{bmatrix} X(\text{WP}) = 0.9504 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 1.0888 \end{bmatrix}_{D65}, \\ \begin{bmatrix} \mathcal{R}(\text{WP}) \\ \mathcal{G}(\text{WP}) \\ \mathcal{B}(\text{WP}) \end{bmatrix}_{D65} &= \underline{T}_2^{-1} \begin{bmatrix} X(\text{WP}) = 0.9504 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 1.0888 \end{bmatrix}_{D65}, \end{aligned} \quad (34)$$

where $Y(\text{WP}) = 1$ and $\max\{\mathcal{R}(\text{WP}), \mathcal{G}(\text{WP}), \mathcal{B}(\text{WP})\} = 1$.

Unless the smartphone utilizes a color sensor that can directly estimate the scene illumination WP in terms of (x, y) chromaticity coordinates, the AW is calculated by the camera in terms of raw values $\mathcal{R}(\text{AW})$, $\mathcal{G}(\text{AW})$, and $\mathcal{B}(\text{AW})$, so the AW cannot be expressed using the CIE XYZ color space prior to the interpolation. However, the corresponding $\text{CCT}(\text{AW})$ requires knowledge of the (x, y) chromaticity coordinates, which means converting to CIE XYZ via a matrix transformation \underline{T} that itself depends upon the unknown $\text{CCT}(\text{AW})$. This problem can be solved using a self-consistent iteration procedure.³²

1. Make a guess for the AW chromaticity coordinates, $(x(\text{AW}), y(\text{AW}))$. For example, the chromaticity coordinates corresponding to one of the characterization illuminants could be used.
2. Find the CCT value $\text{CCT}(\text{AW})$ that corresponds to the chromaticity coordinates $(x(\text{AW}), y(\text{AW}))$. A widely used approach is to convert $(x(\text{AW}), y(\text{AW}))$ into the corresponding $(u(\text{AW}), v(\text{AW}))$ chromaticity coordinates on the 1960 UCS chromaticity diagram,^{23,24} where isotherms are normal to the Planckian locus. This enables $\text{CCT}(\text{AW})$ to

be determined using Robertson's method.⁴² Alternatively, approximate formulas⁴³⁻⁴⁵ or more recent algorithms⁴⁶ can be implemented.

3. Perform the interpolation so that

$$\underline{T}(AW) = f[\underline{T}_1(\text{CCT}_1), \underline{T}_2(\text{CCT}_2)], \quad (35)$$

where f is the interpolation function. The interpolation is valid for

$$\text{CCT}_1 \leq \text{CCT}(AW) \leq \text{CCT}_2.$$

If $\text{CCT}(AW) < \text{CCT}_1$, then \underline{T} should be set equal to \underline{T}_1 , and if $\text{CCT}(AW) > \text{CCT}_2$, then \underline{T} should be set equal to \underline{T}_2 .

4. Use \underline{T} to transform the AW from the camera raw space to the CIE XYZ color space:

$$\begin{bmatrix} X(AW) \\ Y(AW) \\ Z(AW) \end{bmatrix}_{\text{scene}} = \underline{T} \begin{bmatrix} \mathcal{R}(AW) \\ \mathcal{G}(AW) \\ \mathcal{B}(AW) \end{bmatrix}_{\text{scene}}. \quad (36)$$

This yields a new guess for $(x(AW), y(AW))$.

5. Repeat the procedure starting from step 2 until $(x(AW), y(AW))$, $\text{CCT}(AW)$, and \underline{T} all converge to a stable solution.

After the interpolation has been carried out, \underline{T} inherits the normalization of Eq. (34). However, the AW can now be expressed using the CIE XYZ color space, so \underline{T} can be renormalized to satisfy Eq. (31).

If the smartphone utilizes a color sensor that can directly estimate the scene illumination WP in terms of (x, y) chromaticity coordinates, then only steps 2 and 3 above are required.

5 Traditional Digital Cameras

Consider again the white-balanced transformation from the camera raw space to an output-referred RGB color space. In the case of sRGB, the transformation is defined by Eq. (29):

$$\begin{bmatrix} R_L \\ G_L \\ B_L \end{bmatrix}_{\text{D65}} = \underline{M}_{\text{sRGB}}^{-1} \underline{\text{CAT}}_{\text{AW} \rightarrow \text{D65}} \underline{T} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{\text{scene}},$$

where $\underline{\text{CAT}}_{\text{AW} \rightarrow \text{D65}}$ adapts the scene illumination WP estimate (the AW) to the sRGB color space D65 reference white. Traditional camera manufacturers typically re-express the above equation in the following manner:

$$\begin{bmatrix} R_L \\ G_L \\ B_L \end{bmatrix}_{\text{D65}} = \underline{R} \underline{D} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{\text{scene}}. \quad (37)$$

This equation can be interpreted by decomposing the conversion into two steps.

1. The matrix \underline{D} is a diagonal WB matrix containing raw channel multipliers appropriate for the AW:

$$\underline{D} = \begin{bmatrix} \frac{1}{\mathcal{R}(AW)} & 0 & 0 \\ 0 & \frac{1}{\mathcal{G}(AW)} & 0 \\ 0 & 0 & \frac{1}{\mathcal{B}(AW)} \end{bmatrix}_{\text{scene}}. \quad (38)$$

These are applied to the raw channels before the color demosaic. As shown by Eq. (27), the raw channel multipliers, in particular, serve to chromatically adapt the AW to the reference white of the camera raw space:

$$\begin{bmatrix} \mathcal{R} = 1 \\ \mathcal{G} = 1 \\ \mathcal{B} = 1 \end{bmatrix}_{\text{reference}} = \underline{D} \begin{bmatrix} \mathcal{R}(\text{AW}) \\ \mathcal{G}(\text{AW}) \\ \mathcal{B}(\text{AW}) \end{bmatrix}_{\text{scene}}. \quad (39)$$

2. The matrix \underline{R} is a color rotation matrix optimized for the scene illumination. After the color demosaic has been performed, \underline{R} is applied to convert directly from the camera raw space to the linear form of the chosen output-referred color space. By comparison of Eqs. (29) and (37), \underline{R} is algebraically defined as

$$\underline{R} = \underline{M}_{\text{sRGB}}^{-1} \underline{\text{CAT}}_{\text{AW} \rightarrow \text{D65}} \underline{T} \underline{D}^{-1}. \quad (40)$$

Color rotation matrices have the important property that each of their rows sums to unity:

$$\begin{aligned} R(1,1) + R(1,2) + R(1,3) &= 1, \\ R(2,1) + R(2,2) + R(2,3) &= 1, \\ R(3,1) + R(3,2) + R(3,3) &= 1. \end{aligned} \quad (41)$$

Consequently, \underline{R} maps the reference white of the camera raw space directly to the reference white of the output-referred color space.²¹ In the case of sRGB,

$$\begin{bmatrix} R_L = 1 \\ G_L = 1 \\ B_L = 1 \end{bmatrix}_{\text{D65}} = \underline{R} \begin{bmatrix} \mathcal{R} = 1 \\ \mathcal{G} = 1 \\ \mathcal{B} = 1 \end{bmatrix}_{\text{reference}}. \quad (42)$$

Combining Eqs. (39) and (42) shows that overall WB is achieved since the raw pixel vector corresponding to the AW is mapped to the reference white of the output-referred color space:

$$\begin{bmatrix} R_L = 1 \\ G_L = 1 \\ B_L = 1 \end{bmatrix}_{\text{D65}} = \underline{R} \underline{D} \begin{bmatrix} \mathcal{R}(\text{AW}) \\ \mathcal{G}(\text{AW}) \\ \mathcal{B}(\text{AW}) \end{bmatrix}_{\text{scene}}. \quad (43)$$

Like the characterization matrix \underline{T} , the color rotation matrix \underline{R} should in principle be optimized for the scene illumination. Rather than use an interpolation-based approach, the reformulation in the form of Eq. (37) enables traditional camera manufacturers to adopt an alternative and computationally simple approach that can be straightforwardly implemented on fixed-point number architecture.

5.1 Multiplier and Matrix Decoupling

Although Eq. (37) appears to be a straightforward reformulation of Eq. (29), it has several advantages that arise from the raw channel multipliers contained within the WB matrix \underline{D} having been extracted. As shown in Fig. 8, the variation of the elements of a color rotation matrix with respect to CCT is very small. The stability is greater than that of the elements of a conventional characterization matrix \underline{T} , as evident from comparison of Figs. 5 and 8.

Consequently, it suffices to determine a small set of n preset color rotation matrices that cover a range of WPs or CCTs, with each matrix optimized for a particular preset WP or CCT:

$$\underline{R}_i = \underline{M}_{\text{sRGB}}^{-1} \underline{\text{CAT}}_{\text{AW} \rightarrow \text{D65}} \underline{T}_i \underline{D}_i^{-1}, \quad (44)$$

where $i = 1 \dots n$. When the AW is calculated by the camera, the color rotation matrix \underline{R}_i optimized for the closest-matching WP or CCT preset can be selected. However, the WB matrix \underline{D}

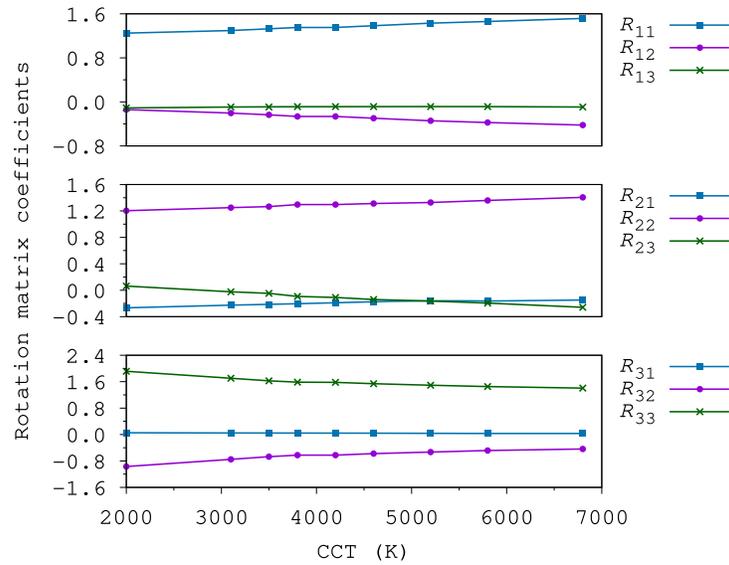


Fig. 8 Variation of the matrix elements of the raw-to-sRGB color rotation matrix \underline{R} used by the Olympus E-M1 camera as a function of CCT.

appropriate for the AW is always applied prior to \underline{R}_i , so the overall color conversion can be expressed as

$$\begin{bmatrix} R_L \\ G_L \\ B_L \end{bmatrix}_{D65} = (\underline{M}_{sRGB}^{-1} \underline{CAT}_{AW \rightarrow D65} \underline{T}_i \underline{D}_i^{-1}) \underline{D} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{scene}. \quad (45)$$

Since \underline{D} is decoupled from the rotation matrices, this approach will achieve correct WB without the need to interpolate the rotation matrices.

It should be noted that the camera raw space correctly represents the scene (albeit via a non-standard color model) and that the raw channel multipliers contained within \underline{D} are not applied to “correct” anything concerning the representation of the true scene white by the camera raw space, as often assumed. The multipliers are applied to chromatically adapt the AW to the reference white of the camera raw space as part of the overall CAT required to achieve WB by emulating the chromatic adaptation mechanism of the HVS. As shown in Fig. 4, the reference white of a camera raw space is typically a magenta color when expressed using CIE colorimetry, but it serves as a useful intermediary stage in the required color transformation as it facilitates the extraction of a channel scaling component that can be decoupled from the matrix operation. Other advantages of the reformulation include the following.

- The raw channel multipliers contained within \underline{D} can be applied to the raw channels before the color mosaic is performed. This results in a demosaic of better quality.²¹
- The method can be efficiently implemented on fixed-point architecture.⁴⁷
- If desired, part of the raw channel scaling can be carried out in the analog domain using analog amplification. This is beneficial for image quality if the analog-to-digital converter (ADC) does not have a sufficiently high bit depth. Note that this type of analog amplification will affect the input to output-referred unit conversion factors g_i defined by Eq. (80) in the Appendix.
- The raw channel multipliers contained within \underline{D} that appear in Eq. (37) are stored in the proprietary raw file metadata and are applied by the internal JPEG image-processing engine of the camera. Since the raw channel multipliers do not affect the raw data, they can be utilized by external raw conversion software provided by the camera manufacturer and can be easily adjusted by the user.

- Scene illumination presets that include a color tint can be straightforwardly implemented by storing the appropriate preset color rotation matrices and raw channel multipliers, as illustrated in Sec. 5.2.

5.2 Example: Olympus E-M1

Although the color matrices used by the camera manufacturers are generally unknown, certain manufacturers such as Sony and Olympus do reveal information about the color rotation matrices used by their cameras that can be extracted from the raw metadata.

Table 1 lists the data illustrated in Fig. 8 for the preset color rotation matrices used by the Olympus E-M1 digital camera, along with the scene illumination CCT ranges over which each matrix is applied. Figure 9 shows how the raw channel multipliers for the same camera vary as a function of CCT. The data was extracted from raw metadata using the freeware “ExifTool” application.⁴⁸ The color conversion strategy of the camera can be summarized as follows.

1. The camera determines the scene illumination WP estimate (the AW) using either an auto-WB algorithm, a selected scene illumination preset, or a custom CCT provided by the user.

Table 1 Raw-to-sRGB color rotation matrices corresponding to ranges of in-camera custom CCTs for the Olympus E-M1 camera with 12-100/4 lens and v4.1 firmware. The middle column lists the matrices extracted from the raw metadata, which are 8-bit fixed-point numbers. By dividing by 256, the right column lists to four decimal places the same matrices such that each row sums to unity rather than 256.

CCT range (K)	Rotation matrix (fixed point)	Rotation matrix
2000 → 3000	$\begin{bmatrix} 320 & -36 & -28 \\ -68 & 308 & 16 \\ 14 & -248 & 490 \end{bmatrix}$	$\begin{bmatrix} 1.2500 & -0.1406 & -0.1094 \\ -0.2656 & 1.2031 & 0.0625 \\ 0.0547 & -0.9688 & 1.9141 \end{bmatrix}$
3100 → 3400	$\begin{bmatrix} 332 & -52 & -24 \\ -58 & 320 & -6 \\ 12 & -192 & 436 \end{bmatrix}$	$\begin{bmatrix} 1.2969 & -0.2031 & -0.0938 \\ -0.2266 & 1.2500 & -0.0234 \\ 0.0469 & -0.7500 & 1.7031 \end{bmatrix}$
3500 → 3700	$\begin{bmatrix} 340 & -60 & -24 \\ -56 & 324 & -12 \\ 12 & -172 & 416 \end{bmatrix}$	$\begin{bmatrix} 1.3281 & -0.2344 & -0.0938 \\ -0.2188 & 1.2656 & -0.0469 \\ 0.0469 & -0.6719 & 1.6250 \end{bmatrix}$
3800 → 4000	$\begin{bmatrix} 346 & -68 & -22 \\ -52 & 332 & -24 \\ 10 & -160 & 406 \end{bmatrix}$	$\begin{bmatrix} 1.3516 & -0.2656 & -0.0859 \\ -0.2031 & 1.2969 & -0.0938 \\ 0.0391 & -0.6250 & 1.5859 \end{bmatrix}$
4200 → 4400	$\begin{bmatrix} 346 & -68 & -22 \\ -48 & 332 & -28 \\ 12 & -160 & 404 \end{bmatrix}$	$\begin{bmatrix} 1.3516 & -0.2656 & -0.0859 \\ -0.1875 & 1.2969 & -0.1094 \\ 0.0469 & -0.6250 & 1.5781 \end{bmatrix}$
4600 → 5000	$\begin{bmatrix} 354 & -76 & -22 \\ -44 & 336 & -36 \\ 10 & -148 & 394 \end{bmatrix}$	$\begin{bmatrix} 1.3828 & -0.2969 & -0.0859 \\ -0.1719 & 1.3125 & -0.1406 \\ 0.0391 & -0.5781 & 1.5391 \end{bmatrix}$
5200 → 5600	$\begin{bmatrix} 366 & -88 & -22 \\ -42 & 340 & -42 \\ 10 & -136 & 382 \end{bmatrix}$	$\begin{bmatrix} 1.4297 & -0.3438 & -0.0859 \\ -0.1641 & 1.3281 & -0.1641 \\ 0.0391 & -0.5313 & 1.4922 \end{bmatrix}$
5800 → 6600	$\begin{bmatrix} 374 & -96 & -22 \\ -42 & 348 & -50 \\ 8 & -124 & 372 \end{bmatrix}$	$\begin{bmatrix} 1.4609 & -0.3750 & -0.0859 \\ -0.1641 & 1.3594 & -0.1953 \\ 0.0313 & -0.4844 & 1.4531 \end{bmatrix}$
6800 → 14000	$\begin{bmatrix} 388 & -108 & -24 \\ -38 & 360 & -66 \\ 8 & -112 & 360 \end{bmatrix}$	$\begin{bmatrix} 1.5156 & -0.4219 & -0.0938 \\ -0.1484 & 1.4063 & -0.2578 \\ 0.0313 & -0.4375 & 1.4063 \end{bmatrix}$

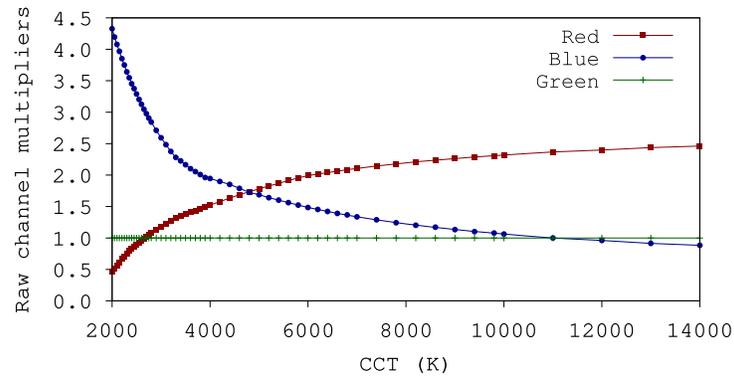


Fig. 9 Raw channel multipliers used by the Olympus E-M1 camera as a function of CCT. The camera uses the same multipliers for both of the green channels.

Table 2 Raw-to-sRGB color rotation matrices and associated raw channel multipliers corresponding to in-camera scene modes for the Olympus E-M1 camera with 12-100/4 lens and v4.1 firmware. All values are 8-bit fixed-point numbers that can be divided by 256. Since the scene mode presets include a color tint away from the Planckian locus, the multipliers and matrices do not necessarily have the same values as the custom CCT presets with the same CCT listed in Table 1.

Scene mode	CCT (K)	Multipliers	Rotation matrix (fixed point)
Fine weather	5300	474 256 414	$\begin{bmatrix} 366 & -88 & -22 \\ -42 & 340 & -42 \\ 10 & -136 & 382 \end{bmatrix}$
Fine weather with shade	7500	552 256 326	$\begin{bmatrix} 388 & -108 & -24 \\ -38 & 360 & -66 \\ 8 & -112 & 360 \end{bmatrix}$
Cloudy	6000	510 256 380	$\begin{bmatrix} 374 & -96 & -22 \\ -42 & 348 & -50 \\ 8 & -124 & 372 \end{bmatrix}$
Tungsten (incandescent)	3000	276 256 728	$\begin{bmatrix} 320 & -36 & -28 \\ -68 & 308 & 16 \\ 14 & -248 & 490 \end{bmatrix}$
Cool white fluorescent	4000	470 256 580	$\begin{bmatrix} 430 & -168 & -6 \\ -50 & 300 & 6 \\ 12 & -132 & 376 \end{bmatrix}$
Underwater		450 256 444	$\begin{bmatrix} 366 & -88 & -22 \\ -42 & 340 & -42 \\ 10 & -136 & 382 \end{bmatrix}$
Flash	5500	562 256 366	$\begin{bmatrix} 366 & -88 & -22 \\ -42 & 340 & -42 \\ 10 & -136 & 382 \end{bmatrix}$

The AW is used to calculate the appropriate raw channel multipliers via Eq. (38) so that the diagonal WB matrix \underline{D} can be applied to the raw channels. In particular, \underline{D} serves to adapt the AW to the reference white of the camera raw space.

- After the color demosaic is performed, the camera chooses the preset color rotation matrix \underline{R}_i optimized for illumination with a CCT that provides the closest match to the CCT associated with the AW, or the closest matching scene illumination preset.

3. The camera applies \underline{R}_i to convert to the output-referred color space selected in-camera by the user, such as sRGB. In particular, the camera raw space reference white is mapped to the reference white of the selected output-referred color space, which is D65 in the case of sRGB.

The Olympus E-M1 camera also includes several scene illumination presets. The color rotation matrices and associated raw channel multipliers for these scene presets are listed in Table 2. For a given CCT, notice that the scene preset matrices and multipliers are not necessarily the same as those listed in Table 1. This is because the scene preset renderings include a color tint away from the Planckian locus, so the chromaticity coordinates are not necessarily the same as those listed in Table 1 for a given CCT. For the same reason, notice that the “fine weather,” “underwater,” and “flash” scene mode presets actually use the same color rotation matrix but use very different raw channel multipliers.

For any given camera model, all preset color rotation matrices are dependent on factors such as the output-referred color space selected by the user in the camera settings (such as sRGB or Adobe® RGB), the lens model used to take the photograph, and the firmware version. Due to sensor calibration differences between different examples of the same camera model, there can also be a dependence on the individual camera used to take the photograph.

For example, Fig. 10(a) shows a photo of a color chart in the camera raw space taken under D65 illumination. Like Fig. 6(a), the green color tint arises from the fact that the camera raw space \mathcal{RGB} values are being interpreted as RGB values in the sRGB color space for display purposes without any color characterization matrix applied to convert the colors. Figure 10(b) shows the same photo after applying the diagonal WB matrix \underline{D} to chromatically adapt the AW to the camera raw space reference white. The raw channel multipliers remove the green tint, but the photo remains in the camera raw space. Remarkably, the colors appear realistic, although desaturated. To illustrate that the camera raw space reference white is actually a magenta color when expressed using CIE colorimetry, Fig. 10(c) converts (b) to the sRGB color space without any further chromatic adaptation by applying a conventional characterization matrix \underline{T} followed by $\underline{M}_{\text{sRGB}}^{-1}$. In contrast, Fig. 10(d) was obtained by applying the appropriate raw channel multipliers followed by the sRGB color rotation matrix \underline{R} in place of \underline{T} and $\underline{M}_{\text{sRGB}}^{-1}$. The color rotation matrix includes a CAT that adapts the camera raw space reference white to the sRGB color space D65

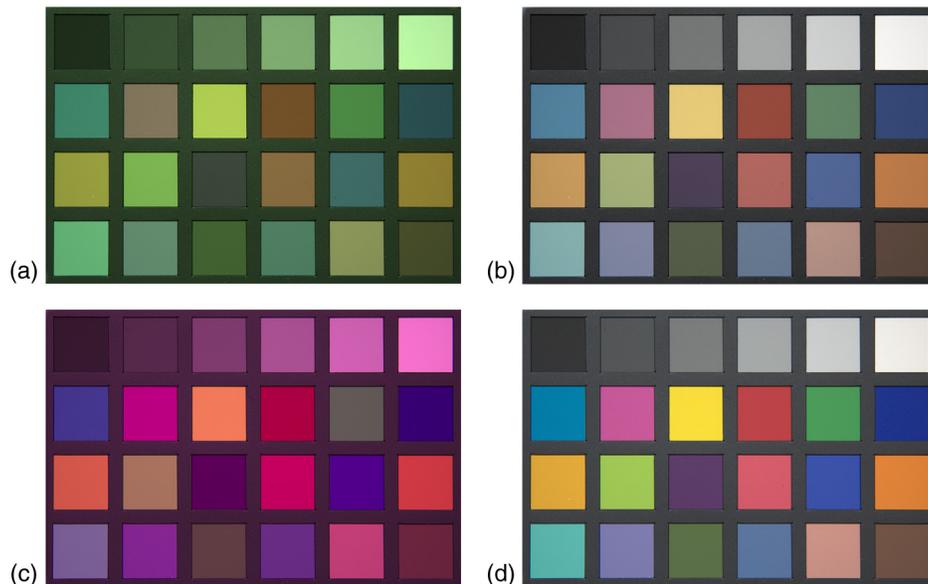


Fig. 10 (a) Photo of a color chart in the camera raw space taken under D65 illumination. (b) After application of the appropriate raw channel multipliers. These remove the green tint, but the photo remains in the camera raw space. (c) After application of the appropriate raw channel multipliers and converting to sRGB without any further chromatic adaptation. The white patch reveals the true color of the camera raw space reference white. (d) After application of the appropriate raw channel multipliers and an sRGB color rotation matrix \underline{R} .

reference white. In this particular case, $\underline{D} = \underline{D}_{D65}$, so the color rotation matrix \underline{R} defined by Eq. (40) becomes

$$\underline{R} \equiv \underline{R}_{D65} = \underline{M}_{sRGB}^{-1} \underline{T}_{D65} \underline{D}_{D65}^{-1}. \quad (46)$$

Substituting into Eq. (37) yields

$$\begin{bmatrix} R_L \\ G_L \\ B_L \end{bmatrix}_{D65} = \underline{M}_{sRGB}^{-1} \underline{T}_{D65} \underline{D}_{D65}^{-1} \underline{D}_{D65} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{scene}. \quad (47)$$

Consequently, the rotation matrix reverses the effect of the WB matrix since the scene and display illumination is the same.

6 DCRaw Open-Source Raw Converter

The widely used DCRaw open-source raw converter (pronounced “dee-see-raw”) written by D. Coffin can process a wide variety of raw image file formats. It is particularly useful for scientific analysis as it can decode raw files without demosaicing, it can apply linear tone curves, and it can directly output to the camera raw space and the CIE XYZ color space. Some relevant commands are listed in Table 3. However, DCRaw by default outputs directly to the sRGB color space with a D65 illumination WP by utilizing a variation of the traditional digital camera strategy described in the previous section.²⁸

Table 3 A selection of relevant DCRaw commands available in version 9.28. Note that the RGB output colorspace options use color rotation matrices and so should only be used with the correct raw channel multipliers due to the inbuilt CAT.

-v	Print verbose messages
-w	Use camera WB, if possible
-A <x y w h>	Average a gray box for WB
-r <r g b g>	Set custom WB
+M/-M	Use/do not use an embedded color matrix
-H [0-9]	Highlight mode (0 = clip, 1 = unclip, 2 = blend, 3+ = rebuild)
-o [0-6]	Output colorspace (raw, sRGB, Adobe, Wide, ProPhoto, XYZ, ACES)
-d	Document mode (no color, no interpolation)
-D	Document mode without scaling (totally raw)
-W	Do not automatically brighten the image
-b <num>	Adjust brightness (default = 1.0)
-g <p ts>	Set custom gamma curve (default = 2.222 4.5)
-q [0-3]	Set the interpolation quality
-h	Half-size color image (twice as fast as “-q 0”)
-f	Interpolate RGGB as four colors
-6	Write 16-bit instead of 8-bit
-4	Linear 16-bit, same as “-6 -W -g 1 1”
-T	Write TIFF instead of PPM

Recall that the color rotation matrix optimized for use with the scene illumination is defined by Eq. (40):

$$\underline{R} = \underline{M}_{\text{sRGB}}^{-1} \underline{\text{CAT}}_{\text{AW} \rightarrow \text{D65}} \underline{T} \underline{D}^{-1}.$$

Although digital cameras typically use a small set of preset rotation matrices optimized for a selection of preset illuminants, DCRaw instead takes a very computationally simple approach that uses only a single rotation matrix optimized for D65 scene illumination, $\underline{R} \approx \underline{R}_{\text{D65}}$. This is achieved using a characterization matrix $\underline{T}_{\text{D65}}$ optimized for D65 illumination, which means that the \underline{D}^{-1} matrix contained within \underline{R} is replaced by $\underline{D}_{\text{D65}}^{-1}$ and the $\underline{\text{CAT}}_{\text{AW} \rightarrow \text{D65}}$ matrix is not required:

$$\underline{R}_{\text{D65}} = \underline{M}_{\text{sRGB}}^{-1} \underline{T}_{\text{D65}} \underline{D}_{\text{D65}}^{-1}. \quad (48)$$

The diagonal WB matrix $\underline{D}_{\text{D65}}$ contains raw channel multipliers appropriate for D65 illumination:

$$\underline{D}_{\text{D65}} = \begin{bmatrix} \frac{1}{\mathcal{R}(\text{WP})} & 0 & 0 \\ 0 & \frac{1}{\mathcal{G}(\text{WP})} & 0 \\ 0 & 0 & \frac{1}{\mathcal{B}(\text{WP})} \end{bmatrix}_{\text{D65}} = \begin{bmatrix} \frac{1}{\mathcal{R}(\text{D65})} & 0 & 0 \\ 0 & \frac{1}{\mathcal{G}(\text{D65})} & 0 \\ 0 & 0 & \frac{1}{\mathcal{B}(\text{D65})} \end{bmatrix}. \quad (49)$$

The overall transformation from the camera raw space to the linear form of sRGB is defined by

$$\begin{bmatrix} R_L \\ G_L \\ B_L \end{bmatrix}_{\text{D65}} \approx \underline{R}_{\text{D65}} \underline{D} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{\text{scene}}, \quad (50)$$

which can be more explicitly written as

$$\begin{bmatrix} R_L \\ G_L \\ B_L \end{bmatrix}_{\text{D65}} \approx \underline{M}_{\text{sRGB}}^{-1} \underline{T}_{\text{D65}} \begin{bmatrix} \frac{\mathcal{R}(\text{D65})}{\mathcal{R}(\text{AW})} & 0 & 0 \\ 0 & \frac{\mathcal{G}(\text{D65})}{\mathcal{G}(\text{AW})} & 0 \\ 0 & 0 & \frac{\mathcal{B}(\text{D65})}{\mathcal{B}(\text{AW})} \end{bmatrix} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{\text{scene}}. \quad (51)$$

Consequently, all chromatic adaptation is performed using raw channel multipliers. Notice that the WB matrix \underline{D} appropriate for the scene illumination estimate is always applied to the raw data in Eq. (50), so WB is always correctly achieved in principle.

Although the color transformation matrix $\underline{T}_{\text{D65}}$ is optimized for D65 scene illumination, applying the color rotation matrix $\underline{R}_{\text{D65}}$ to transform from the camera raw space to sRGB is valid for any scene illumination CCT since color rotation matrices vary very slowly as a function of CCT, as evident from Fig. 8. However, $\underline{R}_{\text{D65}}$ is the optimum choice for D65 scene illumination, so a drawback of this simplified approach is that the overall color transformation loses some accuracy when the scene illumination differs significantly from D65.

6.1 Example: Olympus E-M1

DCRaw uses color rotation matrices obtained via Eq. (48), so a $\underline{T}_{\text{D65}}$ characterization matrix is required for a given camera model. For this purpose, DCRaw uses the Adobe ‘‘ColorMatrix2’’ matrices from the Adobe® DNG converter.³²

Due to highlight recovery logic requirements, the Adobe matrices map in the opposite direction to the conventional characterization matrices defined in Sec. 2.4, and therefore

$$\underline{T}_{\text{D65}} = \left(\frac{1}{c} \underline{\text{ColorMatrix2}} \right)^{-1}, \quad (52)$$

where c is a normalization constant. For the Olympus E-M1 digital camera, the DCRaw source code stores the ColorMatrix2 entries in the following manner:

$$7687, -1984, -606, -4327, 11928, 2721, -1381, 2339, 6452.$$

Dividing by 10,000 and rearranging in matrix form yields

$$\underline{\text{ColorMatrix2}} = \begin{bmatrix} 0.7687 & -0.1984 & -0.0606 \\ -0.4327 & 1.1928 & 0.2721 \\ -0.1381 & 0.2339 & 0.6452 \end{bmatrix}. \quad (53)$$

Recall from Sec. 2.5 that characterization matrices are typically normalized so that the WP of the characterization illuminant maps to raw values such that the maximum value (typically the green channel) just reaches saturation when a 100% neutral diffuse reflector is photographed under the characterization illuminant. Although the ColorMatrix2 matrices are optimized for CIE illuminant D65, they are by default normalized according to the WP of CIE illuminant D50 rather than D65:

$$\begin{bmatrix} \mathcal{R}(\text{WP}) \\ \mathcal{G}(\text{WP}) \\ \mathcal{B}(\text{WP}) \end{bmatrix}_{\text{D50}} = \underline{\text{ColorMatrix2}} \begin{bmatrix} X(\text{WP}) = 0.9642 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 0.8249 \end{bmatrix}_{\text{D50}}, \quad (54)$$

where $\max\{\mathcal{R}(\text{WP}), \mathcal{G}(\text{WP}), \mathcal{B}(\text{WP})\} = 1$. Accordingly, they need to be rescaled for use with DCRaw:

$$\begin{bmatrix} \mathcal{R}(\text{WP}) \\ \mathcal{G}(\text{WP}) \\ \mathcal{B}(\text{WP}) \end{bmatrix}_{\text{D65}} = \frac{1}{c} \underline{\text{ColorMatrix2}} \begin{bmatrix} X(\text{WP}) = 0.9504 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 1.0888 \end{bmatrix}_{\text{D65}}, \quad (55)$$

where $\max\{\mathcal{R}(\text{WP}), \mathcal{G}(\text{WP}), \mathcal{B}(\text{WP})\} = 1$. In the present example, it is found that $c = 1.0778$, so

$$\underline{T}_{\text{D65}}^{-1} = \begin{bmatrix} 0.7133 & -0.1841 & -0.0562 \\ -0.4015 & 1.1068 & 0.2525 \\ -0.1281 & 0.2170 & 0.5987 \end{bmatrix}. \quad (56)$$

By considering the unit vector in the sRGB color space, the above matrix can be used to obtain the raw tristimulus values for the D65 illumination WP:

$$\begin{bmatrix} \mathcal{R}(\text{WP}) = 0.4325 \\ \mathcal{G}(\text{WP}) = 1.0000 \\ \mathcal{B}(\text{WP}) = 0.7471 \end{bmatrix}_{\text{D65}} = \underline{T}_{\text{D65}}^{-1} \underline{M}_{\text{sRGB}} \begin{bmatrix} R_L = 1 \\ G_L = 1 \\ B_L = 1 \end{bmatrix}_{\text{D65}}, \quad (57)$$

where $\underline{M}_{\text{sRGB}}$ converts from the linear form of sRGB to CIE XYZ. Now Eq. (49) can be used to extract the raw channel multipliers for scene illumination with a D65 WP:

$$\underline{D}_{\text{D65}} = \begin{bmatrix} 2.3117 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1.3385 \end{bmatrix}. \quad (58)$$

Finally, the color rotation matrix can be calculated from Eq. (48):

$$\underline{R}_{\text{D65}} = \begin{bmatrix} 1.7901 & -0.6689 & -0.1212 \\ -0.2167 & 1.7521 & -0.5354 \\ 0.0543 & -0.5582 & 1.5039 \end{bmatrix}. \quad (59)$$

Each row sums to unity as required. The form of the matrix is similar to the in-camera Olympus matrices listed in Table 1. For comparison purposes, the appropriate listed matrix is the one valid

for scene illuminant CCTs ranging from 5800 to 6600 K. Some numerical differences are expected since D65 illumination has a $D_{uv} = 0.0032$ color tint. Other numerical differences are likely due to differing characterization practices between Olympus and Adobe. Additionally, Adobe uses HSV (hue, saturation, and value) tables to emulate the final color rendering of the in-camera JPEG processing engine.

6.2 DCRaw and MATLAB

As shown in Table 3, DCRaw includes many commands that are useful for scientific research. However, it is important to note that the RGB output color space options use color rotation matrices rather than the concatenation of the raw to CIE XYZ and CIE XYZ to RGB matrices. Since color rotation matrices include an inbuilt CAT, these options will only achieve the expected result in combination with the correct raw channel multipliers. For example, setting each raw channel multiplier to unity will not prevent some partial chromatic adaptation from being performed if the sRGB output is selected since the DCRaw color rotation matrix incorporates the \underline{D}_{D65}^{-1} matrix, which is a type of $\underline{CAT}_{RW \rightarrow D65}$.

A robust way to use DCRaw for scientific research is via the “dcrw -v -D -4 -T filename” command, which provides linear 16-bit TIFF output in the raw color space without white balancing, demosaicing, or color conversion. Subsequent processing can be performed after importing the TIFF file into MATLAB[®] using the conventional “imread” command. Reference 49 provides a processing tutorial. The color chart photos in the present article were produced using this methodology.

For example, after importing the file into MATLAB via the above commands, a viewable output image in the sRGB color space without any white balancing can be obtained by applying the appropriate characterization matrix \underline{T} after the color demosaic, followed by direct application of the standard CIE XYZ to sRGB matrix, $\underline{M}_{sRGB}^{-1}$.

7 Adobe DNG

The Adobe[®] DNG is an open-source raw file format developed by Adobe.^{32,50} The freeware DNG Converter can be used to convert any raw file into the DNG format.

Although the DNG converter does not aim to produce a viewable output image, it does perform a color conversion from the camera raw space into the profile connection space (PCS) based on the CIE XYZ color space with a D50 illumination WP.⁴⁰ (This is not the actual reference white of CIE XYZ, which is CIE illuminant E.) Consequently, the color processing model used by the DNG converter must provide appropriate characterization matrices along with a strategy for achieving correct WB in relation to the PCS. When processing DNG files, raw converters can straightforwardly map from the PCS to any chosen output-referred color space and associated reference white.

The DNG specification provides two different color processing models referred to here as method 1 and method 2. Method 1 adopts a similar strategy as smartphones and commercial raw converters, the difference being that the data remain in the PCS. By using raw channel multipliers, method 2 adopts a similar strategy as traditional digital cameras. However, the multipliers are applied in conjunction with a so-called forward matrix instead of a rotation matrix since the mapping is to the PCS rather than to an output-referred RGB color space.

7.1 Method 1: Color Matrices

The transformation from the camera raw space to the PCS is defined as follows:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{D50} = \underline{CAT}_{AW \rightarrow D50} \underline{C}^{-1} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{\text{scene}} . \quad (60)$$

Here \underline{C} is an Adobe color matrix optimized for the scene AW. Due to highlight recovery logic requirements, Adobe color matrices map in the direction from the CIE XYZ color space to the camera raw space:

$$\begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{\text{scene}} = \underline{\mathcal{C}} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{\text{scene}}. \quad (61)$$

This is the opposite direction to a conventional characterization matrix \underline{T} , so

$$\underline{\mathcal{C}} \propto \underline{T}^{-1}. \quad (62)$$

After the $\underline{\mathcal{C}}$ inverse maps from the camera raw space to CIE XYZ, the linear Bradford CAT is applied to adapt the AW to the WP of the PCS.

Analogous to the issue described in Sec. 4 for smartphones, the implementation of Eq. (60) is complicated by the fact that $\underline{\mathcal{C}}$ should be optimized for the scene AW. The optimized $\underline{\mathcal{C}}$ matrix is determined by interpolating between two color matrices labeled ColorMatrix1 and ColorMatrix2, where ColorMatrix1 should be obtained from a characterization performed using a low-CCT illuminant such as CIE illuminant A and ColorMatrix2 should be obtained from a characterization performed using a high-CCT illuminant such as CIE illuminant D65.³²

The optimized matrix $\underline{\mathcal{C}}$ is calculated by interpolating between ColorMatrix1 and ColorMatrix2 based on the scene illumination CCT estimate denoted by CCT(AW), together with the CCTs associated with each of the two characterization illuminants denoted by CCT₁ and CCT₂, respectively, with CCT₁ < CCT₂.

7.2 Color Matrix Normalization

Recall from Sec. 2.5 that characterization matrices are typically normalized so that the characterization illuminant WP in the CIE XYZ color space just saturates the raw data in the camera raw space and that the green raw channel is typically the first to saturate. However, in the present context, the Adobe ColorMatrix1 and ColorMatrix2 matrices require a common normalization that is convenient for performing the interpolation. Analogous to Sec. 4.1, the AW is not known in terms of the CIE XYZ color space prior to the interpolation. Instead, ColorMatrix1 and ColorMatrix2 are by default normalized so that the WP of the PCS just saturates the raw data:

$$\begin{aligned} \begin{bmatrix} \mathcal{R}(\text{WP}) \\ \mathcal{G}(\text{WP}) \\ \mathcal{B}(\text{WP}) \end{bmatrix}_{\text{D50}} &= \underline{\text{ColorMatrix1}} \begin{bmatrix} X(\text{WP}) = 0.9642 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 0.8249 \end{bmatrix}_{\text{D50}}, \\ \begin{bmatrix} \mathcal{R}(\text{WP}) \\ \mathcal{G}(\text{WP}) \\ \mathcal{B}(\text{WP}) \end{bmatrix}_{\text{D50}} &= \underline{\text{ColorMatrix2}} \begin{bmatrix} X(\text{WP}) = 0.9642 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 0.8249 \end{bmatrix}_{\text{D50}}, \end{aligned} \quad (63)$$

where $\max\{\mathcal{R}(\text{WP}), \mathcal{G}(\text{WP}), \mathcal{B}(\text{WP})\} = 1$. For example, the default ColorMatrix1 and ColorMatrix2 for the Olympus E-M1 camera are, respectively, normalized as follows:

$$\begin{aligned} \begin{bmatrix} \mathcal{R}(\text{WP}) = 0.5471 \\ \mathcal{G}(\text{WP}) = 1.0000 \\ \mathcal{B}(\text{WP}) = 0.6560 \end{bmatrix}_{\text{D50}} &= \begin{bmatrix} 1.1528 & -0.5742 & 0.0118 \\ -0.2453 & 1.0205 & 0.2619 \\ -0.0751 & 0.189 & 0.6539 \end{bmatrix} \begin{bmatrix} X(\text{WP}) = 0.9642 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 0.8249 \end{bmatrix}_{\text{D50}}, \\ \begin{bmatrix} \mathcal{R}(\text{WP}) = 0.4928 \\ \mathcal{G}(\text{WP}) = 1.0000 \\ \mathcal{B}(\text{WP}) = 0.6330 \end{bmatrix}_{\text{D50}} &= \begin{bmatrix} 0.7687 & -0.1984 & -0.0606 \\ -0.4327 & 1.1928 & 0.2721 \\ -0.1381 & 0.2339 & 0.6452 \end{bmatrix} \begin{bmatrix} X(\text{WP}) = 0.9642 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 0.8249 \end{bmatrix}_{\text{D50}}. \end{aligned} \quad (64)$$

The interpolated $\underline{\mathcal{C}}$ initially inherits this normalization. However, after $\underline{\mathcal{C}}$ has been determined, the CIE XYZ values for the AW will be known. Consequently, the Adobe DNG SDK source code later re-normalizes Eq. (60) so that the AW in the camera raw space maps to the WP of the PCS when the raw data just saturates:

$$\begin{bmatrix} X(\text{WP}) = 0.9641 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 0.8249 \end{bmatrix}_{\text{D50}} = \underline{\text{CAT}}_{\text{AW} \rightarrow \text{D50}} \underline{\text{C}}^{-1} \begin{bmatrix} \mathcal{R}(\text{AW}) \\ \mathcal{G}(\text{AW}) \\ \mathcal{B}(\text{AW}) \end{bmatrix}_{\text{scene}}, \quad (65)$$

where $\max\{\mathcal{R}(\text{WP}), \mathcal{G}(\text{WP}), \mathcal{B}(\text{WP})\} = 1$. This is equivalent to re-normalizing $\underline{\text{C}}$ as follows:

$$\begin{bmatrix} \mathcal{R}(\text{AW}) \\ \mathcal{G}(\text{AW}) \\ \mathcal{B}(\text{AW}) \end{bmatrix}_{\text{scene}} = \underline{\text{C}} \begin{bmatrix} X(\text{AW}) \\ Y(\text{AW}) \\ Z(\text{AW}) \end{bmatrix}_{\text{scene}}, \quad (66)$$

where $Y(\text{AW}) = 1$ and $\max\{\mathcal{R}(\text{WP}), \mathcal{G}(\text{WP}), \mathcal{B}(\text{WP})\} = 1$.

7.3 Linear Interpolation Based on Inverse CCT

The method 1 interpolation algorithm is the same as that described in Sec. 4.1, except that ColorMatrix1, ColorMatrix2, and $\underline{\text{C}}$ replace $\underline{\text{T}}_1$, $\underline{\text{T}}_2$, and $\underline{\text{T}}$, respectively. Furthermore, the Adobe DNG specification requires the interpolation method to be linear interpolation based upon inverse CCT.³²

Again, the interpolation itself is complicated by the fact that the AW is typically calculated by the camera in terms of raw values $\mathcal{R}(\text{AW})$, $\mathcal{G}(\text{AW})$, and $\mathcal{B}(\text{AW})$, but the corresponding CCT(AW) requires knowledge of the (x, y) chromaticity coordinates. This means converting to CIE XYZ via a matrix transformation $\underline{\text{C}}$ that itself depends upon the unknown CCT(AW), which can be solved using a self-consistent iteration procedure.

1. Make a guess for the AW chromaticity coordinates, $(x(\text{AW}), y(\text{AW}))$. For example, the chromaticity coordinates corresponding to one of the characterization illuminants could be used.
2. Find the CCT value CCT(AW) that corresponds to the chromaticity coordinates $(x(\text{AW}), y(\text{AW}))$ using one of the methods listed in step 2 of Sec. 4.1.
3. Perform a linear interpolation:

$$\underline{\text{C}} = \alpha \underline{\text{ColorMatrix1}} + (1 - \alpha) \underline{\text{ColorMatrix2}}, \quad (67)$$

where α is the CCT-dependent weighting that depends upon inverse-CCT:

$$\alpha = \frac{(\text{CCT}(\text{AW}))^{-1} - (\text{CCT}_2)^{-1}}{(\text{CCT}_1)^{-1} - (\text{CCT}_2)^{-1}}. \quad (68)$$

These weights (denoted by g and $1 - g$ in the Adobe DNG SDK source code) are illustrated in Fig. 11 for a pair of example CCT₁ and CCT₂ values. The interpolation is valid for

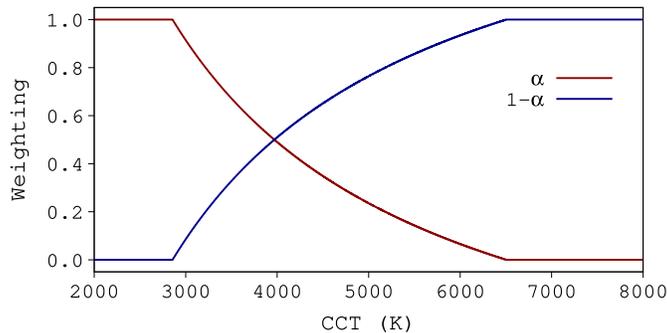


Fig. 11 Linear interpolation weighting factors α and $1 - \alpha$ -based on inverse CCT with CCT₁ = 2855 K and CCT₂ = 6504 K.

$$\text{CCT}(1) \leq \text{CCT}(\text{AW}) \leq \text{CCT}(2).$$

If $\text{CCT}(\text{AW}) < \text{CCT}_1$, then \underline{C} should be set equal to ColorMatrix1, and if $\text{CCT}(\text{AW}) > \text{CCT}_2$, then \underline{C} should be set equal to ColorMatrix2.

- Use \underline{C} to transform the AW from the camera raw space to CIE XYZ:

$$\begin{bmatrix} X(\text{AW}) \\ Y(\text{AW}) \\ Z(\text{AW}) \end{bmatrix}_{\text{scene}} = \underline{C}^{-1} \begin{bmatrix} \mathcal{R}(\text{AW}) \\ \mathcal{G}(\text{AW}) \\ \mathcal{B}(\text{AW}) \end{bmatrix}_{\text{scene}}. \quad (69)$$

This yields a new guess for $(x(\text{AW}), y(\text{AW}))$.

- Repeat the procedure starting from step 2 until $(x(\text{AW}), y(\text{AW}))$, $\text{CCT}(\text{AW})$, and \underline{C} all converge to a stable solution.
- Normalize the color conversion according to Eq. (65).

Figure 12 illustrates the results of inverse-CCT based linear interpolation using the Adobe color matrices defined by Eq. (64) for the Olympus E-M1 camera. Note that the ColorMatrix2 is the same as that defined by Eq. (53), which was extracted from the DCRaw source code.

Since \underline{C} maps in the direction from the CIE XYZ color space to the camera raw space, the inverse of the interpolated \underline{C} can be compared to a conventional characterization matrix \underline{T} at a given illuminant CCT. Figure 13 shows the inverse of the interpolated \underline{C} plotted as a function of CCT, and this figure can be compared with Fig. 5, which shows conventional characterization matrices for the same camera optimized for a selection of CCTs. Although the two plots use different normalizations since the characterization matrices are normalized according to their characterization illuminant WP rather than the WP of the PCS, the variation with respect to CCT is similar. However, it is evident that the interpolated \underline{C} loses accuracy for CCTs below CCT_1 .

7.4 Method 2: Forward Matrices

Consider the transformation from the camera raw space to the PCS defined by Eq. (60):

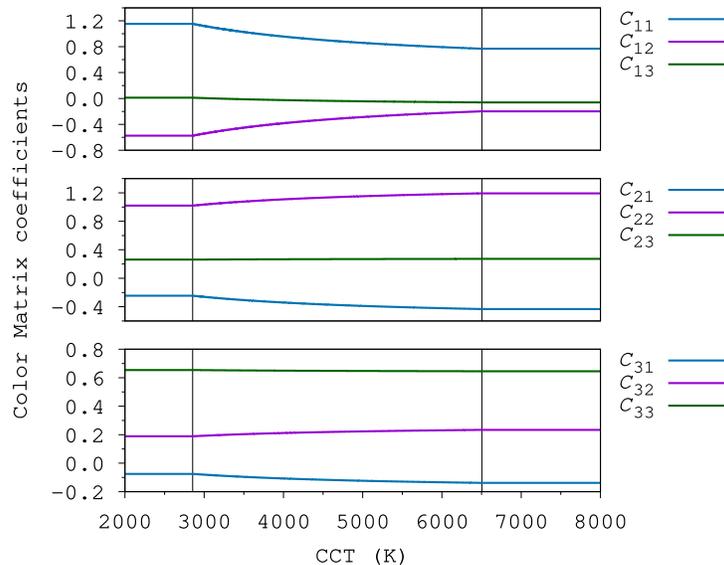


Fig. 12 Optimized color matrix \underline{C} plotted as a function of CCT and obtained via inverse-CCT-based linear interpolation of the Adobe ColorMatrix1 (illuminant A, $\text{CCT}_2 = 2855$ K) and ColorMatrix2 (illuminant D65, $\text{CCT}_2 = 6504$ K) color conversion matrices for the Olympus E-M1 camera.

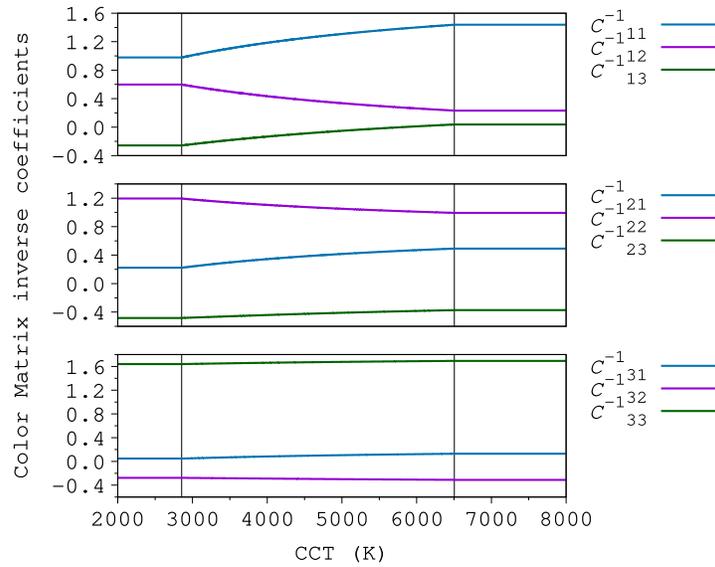


Fig. 13 Inverse of the interpolated color matrix \underline{C} plotted in Fig. 12.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{\text{D50}} = \underline{\text{CAT}}_{\text{AW} \rightarrow \text{D50}} \underline{C}^{-1} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{\text{scene}},$$

where \underline{C} is the Adobe color matrix optimized for the scene AW. Method 2 reformulates the above transformation in the following manner:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{\text{D50}} = \underline{F} \underline{D} \begin{bmatrix} \mathcal{R} \\ \mathcal{G} \\ \mathcal{B} \end{bmatrix}_{\text{scene}}. \quad (70)$$

The color conversion can be decomposed into two steps.

1. Analogous to the color conversion strategy of traditional digital cameras described in Sec. 5, the diagonal matrix \underline{D} defined by Eq. (38) contains raw channel multipliers appropriate for the AW, i.e., the estimated scene illumination WP estimate:

$$\underline{D} = \begin{bmatrix} \frac{1}{\mathcal{R}(\text{AW})} & 0 & 0 \\ 0 & \frac{1}{\mathcal{G}(\text{AW})} & 0 \\ 0 & 0 & \frac{1}{\mathcal{B}(\text{AW})} \end{bmatrix}_{\text{scene}}.$$

In particular, the raw channel multipliers serve to chromatically adapt the AW to the reference white of the camera raw space:

$$\begin{bmatrix} \mathcal{R} = 1 \\ \mathcal{G} = 1 \\ \mathcal{B} = 1 \end{bmatrix} = \underline{D} \begin{bmatrix} \mathcal{R}(\text{AW}) \\ \mathcal{G}(\text{AW}) \\ \mathcal{B}(\text{AW}) \end{bmatrix}_{\text{scene}}. \quad (71)$$

Note that the Adobe DNG specification also accommodates raw channel multipliers applied in the analog domain.³² However, the latest digital cameras utilize ADCs with a relatively high bit depth of order 12 or 14 and consequently elect to apply raw channel multipliers in the digital domain.

2. The forward matrix \underline{F} is a type of characterization matrix that maps from the camera raw space to the PCS and is optimized for the scene illumination. Since the PCS is based on the

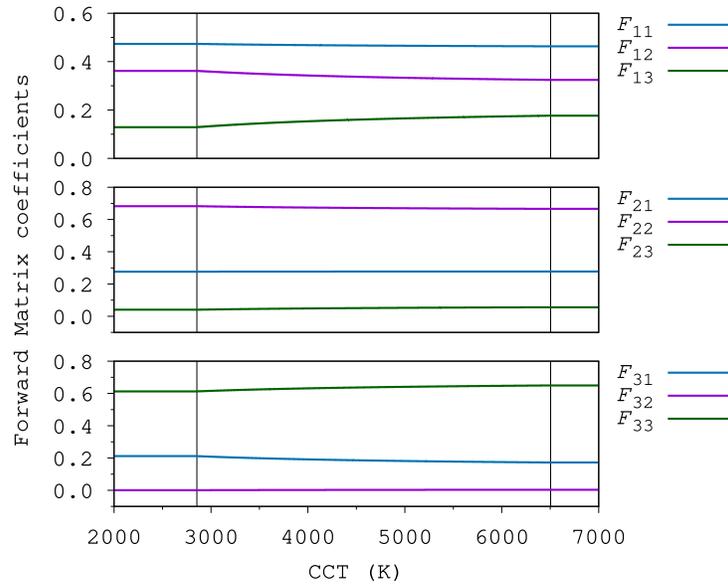


Fig. 14 Optimized forward matrix \underline{F} plotted as a function of CCT and obtained via inverse-CCT-based linear interpolation of the Adobe ForwardMatrix1 (illuminant A, CCT₂ = 2855 K) and Forward Matrix2 (illuminant D65, CCT₂ = 6504 K) matrices for the Olympus E-M1 camera. Evidently, the elements of the optimized forward matrix \underline{F} vary very slowly and in a stable manner as a function of CCT, analogous to the color rotation matrix elements illustrated in Fig. 8.

CIE XYZ color space with a D50 illumination WP, the forward matrix \underline{F} includes a built-in CAT as it must also adapt the reference white of the camera raw space to the WP of D50 illumination:

$$\begin{bmatrix} X(\text{WP}) = 0.9642 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 0.8249 \end{bmatrix}_{\text{D50}} = \underline{F} \begin{bmatrix} \mathcal{R} = 1 \\ \mathcal{G} = 1 \\ \mathcal{B} = 1 \end{bmatrix}. \quad (72)$$

Since the forward matrix \underline{F} should be optimized for the scene AW, it is in practice determined by interpolating between two forward matrices analogous to the interpolation approach used by method 1. The Adobe DNG specification provides tags for two forward matrices labeled ForwardMatrix1 and ForwardMatrix2, which should again be obtained from characterizations performed using a low-CCT illuminant and a high-CCT illuminant, respectively. The same interpolation method described in the previous section should be used, with ForwardMatrix1, ForwardMatrix2, and \underline{F} replacing ColorMatrix1, ColorMatrix2, and \underline{C} , respectively,

$$\underline{F} = \alpha \underline{\text{ForwardMatrix1}} + (1 - \alpha) \underline{\text{ForwardMatrix2}}. \quad (73)$$

Figure 14 shows the optimized forward matrix interpolated from ForwardMatrix1 and ForwardMatrix2 and expressed as a function of CCT for the Olympus E-M1 camera.

7.5 Forward Matrix Specification

By comparing Eqs. (60) and (70), \underline{F} is algebraically related to the color matrix \underline{C} as follows:

$$\underline{F} = \underline{\text{CAT}}_{\text{AW} \rightarrow \text{D50}} \underline{C}^{-1} \underline{D}^{-1}. \quad (74)$$

Since \underline{F} is interpolated from ForwardMatrix1 and ForwardMatrix2 in practice, these are defined as

$$\begin{aligned}\underline{\text{ForwardMatrix1}} &= \underline{\text{CAT}}_{\text{AW} \rightarrow \text{D50}} \underline{\text{ColorMatrix1}}^{-1} \underline{D}^{-1} \\ \underline{\text{ForwardMatrix2}} &= \underline{\text{CAT}}_{\text{AW} \rightarrow \text{D50}} \underline{\text{ColorMatrix2}}^{-1} \underline{D}^{-1}.\end{aligned}\quad (75)$$

According to Eq. (72), the optimized forward matrix \underline{F} is by definition normalized such that the unit vector in the camera raw space maps to the D50 WP of the PCS.³² This means that ForwardMatrix1 and ForwardMatrix2 must also be normalized in this manner. For example, the default ForwardMatrix1 and ForwardMatrix2 for the Olympus E-M1 camera are, respectively, normalized as follows:

$$\begin{aligned}\begin{bmatrix} X(\text{WP}) = 0.9643 \\ Y(\text{WP}) = 0.9999 \\ Z(\text{WP}) = 0.8251 \end{bmatrix}_{\text{D50}} &= \begin{bmatrix} 0.4734 & 0.3618 & 0.1291 \\ 0.2765 & 0.6827 & 0.0407 \\ 0.2116 & 0.0006 & 0.6129 \end{bmatrix} \begin{bmatrix} \mathcal{R} = 1 \\ \mathcal{G} = 1 \\ \mathcal{B} = 1 \end{bmatrix}, \\ \begin{bmatrix} X(\text{WP}) = 0.9643 \\ Y(\text{WP}) = 1.0000 \\ Z(\text{WP}) = 0.8252 \end{bmatrix}_{\text{D50}} &= \begin{bmatrix} 0.4633 & 0.3244 & 0.1766 \\ 0.2779 & 0.6661 & 0.0560 \\ 0.1722 & 0.0033 & 0.6497 \end{bmatrix} \begin{bmatrix} \mathcal{R} = 1 \\ \mathcal{G} = 1 \\ \mathcal{B} = 1 \end{bmatrix}.\end{aligned}\quad (76)$$

The official D50 WP of the PCS is actually $X = 0.9642$, $Y = 1.0000$, and $Z = 0.8249$,⁴⁰ which is a 16-bit fractional approximation of the true D50 WP defined by $X = 0.9642$, $Y = 1.0000$, and $Z = 0.8251$.

8 Conclusions

The opening section of this paper showed how the DCRaw open-source raw converter can be used to directly characterize a camera without needing to determine and invert the OECF and illustrated how characterization matrices are normalized in practice. As a consequence of camera metameric error, the camera raw space for a typical camera was shown to be warped away from the triangular shape accessible to additive linear combinations of three fixed primaries on the xy chromaticity diagram, and the available gamut was shown to be dependent on the characterization illuminant. It was also shown that the reference white of a typical camera raw space has a strong magenta color tint.

Subsequently, this paper investigated and compared the type of color conversion strategies used by smartphone cameras and commercial raw converters, the image-processing engines of traditional digital cameras, DCRaw, and the Adobe DNG converter.

Smartphones and raw conversion software applications typically adopt the type of color conversion strategy familiar in color science. This involves the application of a characterization matrix \underline{T} to transform from the camera raw space to the CIE XYZ color space, a CAT to chromatically adapt the estimated WP of the scene illumination to the reference white of an output-referred color space (such as D65 for sRGB), and finally a transformation from CIE XYZ to the linear form of the chosen output-referred color space. Since the optimized characterization matrix is CCT-dependent unless the Luther-Ives condition is satisfied, an optimized matrix can be determined by interpolating between two preset characterization matrices, one optimized for a low-CCT illuminant and the other optimized for a high-CCT illuminant. Simpler solutions include using a fixed characterization matrix optimized for representative scene illumination.

For traditional digital cameras, this paper showed how the overall color conversion is typically reformulated in terms of raw channel multipliers \underline{D} along with a set of color rotation matrices \underline{R} . The raw channel multipliers act as a type of CAT by chromatically adapting the scene illumination WP estimate to the reference white of the camera raw space. Since the rows of a color rotation matrix each sum to unity, the rotation matrix subsequently transforms from the camera raw space directly to the chosen output-referred RGB color space and at the same time chromatically adapts the camera raw space reference white to that of the output-referred color space. It was shown that the variation of the elements of a color rotation matrix with respect to CCT is very small, so only a small selection of preset rotation matrices are needed, each optimized for a specified preset illuminant. This enables raw channel multipliers appropriate for the scene illumination WP estimate to be applied in combination with the preset rotation matrix

associated with the closest-matching WP. The primary advantages of the reformulation are that interpolation is not required and the method can be efficiently implemented on fixed-point architecture. Furthermore, image quality can be improved by applying the raw channel multipliers prior to the color demosaic.

It was shown that DCRaw uses a similar model as traditional digital cameras, except that only a single color rotation matrix is used for each camera, specifically a matrix optimized for D65 illumination, \underline{R}_{D65} . Although the overall color conversion loses some accuracy when the scene illumination differs significantly from D65, an advantage of decoupling the raw channel multipliers from the characterization information represented by the color rotation matrix is that WB can be correctly achieved for any type of scene illumination provided raw channel multipliers appropriate for the scene illumination are applied. It was shown that the rotation matrices used by DCRaw can be derived from the inverses of the ‘‘ColorMatrix2’’ color characterization matrices used by the Adobe DNG converter.

The Adobe DNG converter maps the camera raw space and scene illumination WP estimate to an intermediate stage in the overall color conversion, namely the PCS based on the CIE XYZ color space with a D50 WP. Method 1 defines the approach that is also used in commercial raw converters and advanced smartphones. A color matrix \underline{C} optimized for the scene illumination is obtained by interpolating between the ‘‘ColorMatrix1’’ low-CCT and ‘‘ColorMatrix2’’ high-CCT preset matrices. Due to highlight recovery logic requirements, these color matrices map in the opposite direction to conventional characterization matrices. Furthermore, the ColorMatrix1 and ColorMatrix2 matrices are initially normalized according to the WP of the PCS rather than their corresponding characterization illuminants. Since the Adobe color matrices are freely available, their appropriately normalized inverses can serve as useful high-quality characterization matrices when camera characterization equipment is unavailable.

Method 2 offered by the Adobe DNG converter uses raw channel multipliers in a similar manner as traditional digital cameras. However, these are applied in combination with a so-called forward matrix rather than a rotation matrix since the Adobe DNG converter does not directly map to an output-referred RGB color space, so the forward matrix rows do not each sum to unity. Although the optimized forward matrix is determined by interpolating the ‘‘ForwardMatrix1’’ and ‘‘ForwardMatrix2’’ preset matrices, the variation of the optimized forward matrix with respect to CCT is very small, analogous to a rotation matrix.

9 Appendix: Raw Data Model

Consider the raw values expressed as an integration over the spectral passband of the camera according to Eq. (5):

$$\mathcal{R} = k \int_{\lambda_1}^{\lambda_2} R_1(\lambda) \tilde{E}_{e,\lambda} d\lambda, \quad \mathcal{G} = k \int_{\lambda_1}^{\lambda_2} R_2(\lambda) \tilde{E}_{e,\lambda} d\lambda, \quad \mathcal{B} = k \int_{\lambda_1}^{\lambda_2} R_3(\lambda) \tilde{E}_{e,\lambda} d\lambda.$$

Although $\tilde{E}_{e,\lambda}$ can be regarded as the average spectral irradiance at a photosite, it is more precisely described as the spectral irradiance convolved with the camera system point-spread function (PSF) $h(x, y, \lambda)$ and sampled at positional coordinates (x, y) on the sensor plane:

$$\tilde{E}_{e,\lambda}(x, y) = [E_{e,\lambda,\text{ideal}}(x, y) * h(x, y, \lambda)] \text{comb} \left[\frac{x}{p_x}, \frac{y}{p_y} \right], \quad (77)$$

where p_x and p_y are the pixel pitches in the horizontal and vertical directions. A noise model can also be included.^{28,51} The quantity denoted by $E_{e,\lambda,\text{ideal}}(x, y)$ is the ideal spectral irradiance at the sensor plane that would theoretically be obtained in the absence of the system PSF:

$$E_{\lambda,\text{ideal}}(x, y) = \frac{\pi}{4} L_{e,\lambda} \left(\frac{x}{m}, \frac{y}{m} \right) \frac{1}{N_w^2} T \cos^4 \left\{ \varphi \left(\frac{x}{m}, \frac{y}{m} \right) \right\}, \quad (78)$$

where $L_{e,\lambda}$ is the corresponding scene spectral radiance, m is the system magnification, N_w is the working f -number of the lens, T is the lens transmittance factor, and φ is the object-space angle

between the optical axis and the indicated scene coordinates. If the vignetting profile of the lens is known, the cosine fourth term can be replaced by the relative illumination factor, which is an image-space function describing the real vignetting profile.⁵²

The constant k that appears in Eq. (5) places an upper bound on the magnitude of the raw values. It can be shown²⁸ that k is given by

$$k = \frac{A_p t}{g_i e}, \quad (79)$$

where t is the exposure duration and g_i is the conversion factor between electron counts and raw values for mosaic i , expressed using e_-/DN units.^{53,54} The conversion factor is inversely proportional to the ISO gain G_{ISO} , which is the analog gain setting of the programmable gain amplifier situated upstream from the ADC:

$$g_i = \frac{U}{G_{\text{ISO},i}}, \quad U = \frac{n_{e,i,\text{FWC}}}{n_{\text{DN},i,\text{clip}}}. \quad (80)$$

Here U is the unity gain, which is the gain setting at which $g_i = 1$. Full-well capacity is denoted by $n_{e,i,\text{FWC}}$, and $n_{\text{DN},i,\text{clip}}$ is the raw clipping point, which is the maximum available raw level. This value is not necessarily as high as the maximum raw level provided by the ADC given its bit-depth M , which is $2^M - 1$ DN, particularly if the camera includes a bias offset that is subtracted before the raw data is written.^{28,53}

The least analog amplification is defined by $G_{\text{ISO}} = 1$, which corresponds to the base ISO gain.^{28,51} The numerical values of the corresponding camera ISO settings S are defined using the JPEG output rather than the raw data.^{55,56} These user values also take into account digital gain applied via the JPEG tone curve. When comparing raw output from cameras based on different sensor formats, equivalent rather than the same exposure settings should be used when possible.⁵⁷

As noted in Sec. 2.2, the actual raw values obtained in practice are quantized values modeled by taking the integer part of Eq. (5), and it is useful to subsequently normalize them to the range $[0,1]$ by dividing Eq. (5) by the raw clipping point.

References

1. E. Machishima, "Color CCD imager having a system for preventing reproducibility degradation of chromatic colors," U.S. Patent 5,253,047 (1993).
2. K. E. Spaulding, R. M. Vogel, and J. R. Szczepanski, "Method and apparatus for color-correcting multi-channel signals of a digital camera," U.S. Patent 5,805,213 (1998).
3. J. Von Kries, "Chromatic adaptation," in *Sources of Color Science*, D. L. MacAdam, Ed., pp. 145–148, MIT Press, Cambridge (1970). [Originally published in *Festschrift der Albrecht-Ludwigs-Universitat* (1902)].
4. H. E. Ives, "The relation between the color of the illuminant and the color of the illuminated object," *Trans. Illum. Eng. Soc.* **7**, 62–72 (1912). [Reprinted in *Color Res. Appl.* **20**, pp. 70–75 (1995)].
5. J. Nakamura, "Basics of image sensors," Chapter 3 in *Image Sensors and Signal Processing for Digital Still Cameras*, J. Nakamura, Ed., pp. 53–93, CRC Press, Taylor & Francis Group, Boca Raton, Florida (2006).
6. J. Jiang et al., "What is the space of spectral sensitivity functions for digital color cameras?" in *IEEE Workshop Appl. Comput. Vision*, pp. 168–179 (2013).
7. R. Luther, "Aus dem Gebiet der Farbreizmetrik (On color stimulus metrics)," *Z. Tech. Phys.* **12**, 540–558 (1927).
8. P.-C. Hung, "Sensitivity metamerism index for digital still camera," *Proc. SPIE* **4922**, 1–14 (2002).
9. P.-C. Hung, "Color theory and its application to digital still cameras," Chapter 7 in *Image Sensors and Signal Processing for Digital Still Cameras*, J. Nakamura, Ed., pp. 205–222, CRC Press, Taylor & Francis Group, Boca Raton, Florida (2006).

10. P. M. Hubel et al., "Matrix calculations for digital photography," in *Proc., IS&T Fifth Color Imaging Conf.*, pp. 105–111 (1997).
11. J. Holm, I. Tastl, and S. Hordley, "Evaluation of DSC (digital still camera) scene analysis error metrics—Part 1," in *Proc., IS&T/SID Eighth Color Imaging Conf.*, pp. 279–287 (2000).
12. R. M. Vogel, "Digital imaging device optimized for color performance," U.S. Patent 5,668,596 (1997).
13. G. D. Finlayson and Y. Zhu, "Finding a colour filter to make a camera colorimetric by optimisation," *Lect. Notes Comput. Sci.* **11418**, 53–62 (2019).
14. D. Alleysson, S. Susstrunk, and J. Hérault, "Linear demosaicing inspired by the human visual system," *IEEE Trans. Image Process.* **14**(4), 439–449 (2005).
15. International Organization for Standardization, "Graphic technology and photography—colour target and procedures for the colour characterisation of digital still cameras (DCSs)," ISO 17321-1:2012 (2012).
16. International Electrotechnical Commission, "Multimedia systems and equipment—colour measurement and management—Part 2-1: colour management—default RGB colour space—sRGB," IEC 61966-2-1:1999 (1999).
17. Adobe Systems Incorporated, "Adobe® RGB (1998) color image encoding," Version 2005-05 (2005).
18. International Organization for Standardization, "Photography and graphic technology—extended colour encodings for digital image storage, manipulation and interchange—Part 2: reference output medium metric RGB colour image encoding (ROMM RGB)," ISO 22028-2:2013 (2013).
19. J. Holm, "Capture color analysis gamuts," in *Proc. 14th Color and Imaging Conf., IS&T*, pp. 108–113 (2006).
20. B. E. Bayer, "Color imaging array," U.S. Patent 3,971,065 (1976).
21. D. Coffin, Private Communication (2015).
22. Y. Ohno, "Practical use and calculation of CCT and DUV," *LEUKOS* **10**(1), 47–55 (2014).
23. D. L. MacAdam, "Projective transformations of I. C. I. color specifications," *J. Opt. Soc. Am.* **27**, 294 (1937).
24. CIE (Commission Internationale de l'Eclairage), *Proc. 14th Session*, Brussels, Vol. A, p. 36 (1959).
25. Commission Internationale de l'Eclairage, "Colorimetry," CIE 15:2004, 3rd ed., Vienna, (2004).
26. G. D. Finlayson and M. S. Drew, "White-point preserving color correction," in *Proc., IS&T Fifth Color Imaging Conf.*, pp. 258–261 (1997).
27. International Organization for Standardization, "Photography—electronic still picture cameras—methods for measuring opto-electronic conversion functions (OECFs)," ISO 14524:2009 (2009).
28. D. A. Rowlands, *Physics of Digital Photography*, IOP Publishing Ltd., Bristol (2017).
29. D. Varghese, R. Wanat, and R. K. Mantiuk, "Colorimetric calibration of high dynamic range images with a ColorChecker chart," in *2nd Int. Conf. and SME Workshop HDR Imaging*, A. Chalmers and K. Bouatouch, Eds., The Eurographics Association (2014).
30. J. McCann, "Do humans discount the illuminant?" *Proc. SPIE* **5666**, 9–16 (2005).
31. International Organization for Standardization, "Photography—electronic still picture imaging—vocabulary," ISO 12231:2012 (2012).
32. Adobe Systems Incorporated, "Adobe digital negative (DNG) specification," Version 1.4 (2012).
33. G. Buchsbaum, "A spatial processor model for object color perception," *J. Franklin Inst.* **310**, 1–26 (1980).
34. E. H. Land and J. McCann, "Lightness and Retinex theory," *J. Opt. Soc. Am.* **61**(1), 1–11 (1971).
35. S. Hordley, "Scene illuminant estimation: past, present and future," *Color Res. Appl.* **31**, 303 (2006).
36. E. Y. Lam and G. S. K. Fung, "Automatic white balancing in digital photography," Chapter 10 in *Single-Sensor Imaging: Methods and Applications for Digital Cameras*, R. Lukac, Ed., pp. 267–294, CRC Press, Boca Raton, Florida (2009).

37. Commission Internationale de l'Eclairage, "Fundamental chromaticity diagram with physiological axes—Part 1," CIE 170-1:2006, Vienna (2006).
38. M. D. Fairchild, *Color Appearance Models*, 3rd ed., Wiley, New York (2013).
39. K. M. Lam, "Metamerism and colour constancy," PhD Thesis, University of Bradford (1985).
40. International Color Consortium, "Image technology colour management—architecture, profile format, and data structure," Specification ICC.1:2010 (profile version 4.3.0.0) (2010).
41. J. Holm et al., "Color processing for digital photography," in *Color Engineering: Achieving Device Independent Color*, P. Green and L. MacDonald, Eds., pp. 179–220, John Wiley & Sons, Ltd., Chichester (2002).
42. A. R. Robertson, "Computation of correlated color temperature and distribution temperature," *J. Opt. Soc. Am.* **58**, 1528 (1968).
43. Q. Xingzhong, "Formulas for computing correlated color temperature," *Color Res. Appl.* **12**(5), 285 (1987).
44. C. S. McCamy, "Correlated color temperature as an explicit function of chromaticity coordinates," *Color Res. Appl.* **17**, 142 (1992).
45. J. Hernández-Andrés, R. L. Lee, and J. Romero, "Calculating correlated color temperatures across the entire gamut of daylight and skylight chromaticities," *Appl. Opt.* **38**, 5703 (1999).
46. C. Li et al., "Accurate method for computing correlated color temperature," *Opt. Express* **24**(13), 14066 (2016).
47. K. A. Parulski, R. H. Hibbard, and J. D'Luna, "Programmable Digital Circuit for Performing a Matrix Multiplication," U.S. Patent No. 5,001,663 (1991).
48. "ExifTool", application by P. Harvey, <https://exiftool.org>.
49. R. Sumner, "Processing RAW Images in MATLAB," unpublished, Department of Electrical Engineering, UC Santa Cruz (2014).
50. Adobe Systems Incorporated, "Introducing the digital negative specification: information for manufacturers" (2004).
51. D. A. Rowlands, *Field Guide to Photographic Science*, SPIE Field Guides, Vol. FG46, SPIE Press, Bellingham (2020).
52. P. Maeda, P. Catrysse, and B. Wandell, "Integrating lens design with digital camera simulation," *Proc. SPIE* **5678**, 48 (2005).
53. E. Martinec, "Noise, dynamic range and bit depth in digital SLRs," unpublished (2008).
54. T. Mizoguchi, "Evaluation of image sensors," Chapter 6 in *Image Sensors and Signal Processing for Digital Still Cameras*, J. Nakamura, Ed., pp. 179–203, CRC Press, Taylor & Francis Group, Boca Raton, Florida (2006).
55. Camera & Imaging Products Association, "Sensitivity of digital cameras," CIPA DC-004 (2004).
56. International Organization for Standardization, "Photography—digital still cameras—determination of exposure index, ISO speed ratings, standard output sensitivity, and recommended exposure index," ISO 12232:2006 (2006).
57. D. A. Rowlands, "Equivalence theory for cross-format photographic image quality comparisons," *Opt. Eng.* **57** (11), 110801 (2018).

D. Andrew Rowlands received his BSc degree in mathematics and physics and his PhD in physics from the University of Warwick, UK, in 2000 and 2004, respectively. He has held research positions at the University of Bristol, UK, Lawrence Livermore National Laboratory, USA, Tongji University, China, and the University of Cambridge, UK, and he has authored three books on the science of digital photography.