Colorimetric Calibration of a Video Digitizing System: Algorithm and Applications

In principle, digitized color video images should be rich and convenient sources of colorimetric information. In practice, these advantages are offset by the difficulty of reliably translating the video camera’s output into colorimetric variables. A solution to this problem is outlined here, one which exploits the fact that spectral reflectances of many natural materials vary slowly in the visible. A characteristic vector analysis of reflectances for a set of such materials leads to an algorithm that gives colorimetrically accurate spectral reflectances from the red-green-blue output of a video digitizing system. Prior knowledge about the illumination leads to chromaticity and luminance information, which can be comparable in quality to that obtained from a spectroradiometer. Some sample retrievals are shown for the algorithm. Since it is designed to correct color biases that are unknown initially, the algorithm has the advantage that images from many sources can be analyzed.

Introduction

Obtaining accurate colorimetric information about color samples may seem to be a hopeless task unless we have access to elaborate spectroradiometers and sophisticated illumination standards. However, it is possible to obtain good estimates of chromaticity values even if we use broad-band devices such as color video cameras, provided that we have a reliable way of estimating their color and luminance biases. Our technique was developed using the CIE 1931 standard observer, but clearly it could be adapted to other colorimetric systems.

Maloney and Wandell’s study of color constancy in human and machine vision provides a theoretical basis for our work. However, our algorithm differs from theirs in some of its assumptions. In particular, we initially assume only an approximate knowledge of the digitizing system’s (camera plus digitizer) color bias. The calibration procedure then improves this estimate of the system’s spectral transfer function. Maloney and Wandell considered at length the general case in which both the lighting and spectral reflectances are unknown. We simplify this problem by assuming that the spectral illumination is known, a constraint that is unlikely to be troublesome in many cases. Finally, they are interested in the general issue of how visual systems achieve color constancy, while we want to examine this problem in a particular type of system. Despite these differences in approach, the goal of our calibration is essentially the same as that outlined in Maloney and Wandell’s work: to allow accurate recovery of spectral reflectances or transmittances (and thus colorimetric data) from samples viewed under a variety of illuminants.

Theory and Notation

We may be tempted to assume that red-green-blue (RGB) pixel values are proxies for colorimetric tristimulus values. This assumption amounts to requiring that the two trichromatic systems be related linearly. Describing the consequences and mathematical nature of this linear relationship is an old issue in color reproduction, and it is referred to as the Maxwell–Ives criterion. Maloney has recently reexamined this issue and demonstrated that if a digitizing system’s spectral sensitivities are not a linear transform of the human color-matching functions, then stimuli that are metamer for one of these visual systems need not be metamer for the other. Mathematically, the absence of a linear transform means that if we use linear equations to relate the two systems’ spectral transfer functions, we can only estimate how the systems will respond to identical stimuli. Physically, the lack of a linear transform means that the digitizing system will be blind to some color differences that are perceptible to humans, and vice versa. We will call
this failure to agree on color matches eye-versus-camera metamereism.

Despite this difficulty, if we know that the digitizing system can make color distinctions over a wide range of chromaticities and luminances, then we can be confident that the system is a useful analog of human vision.\(^5\) In our work, we have addressed the transform problem by empirically modifying a linear relationship between the spectral response of the digitizer and the human visual system. This approach sidesteps the thorny problem that each link in the optical and electronic chain of the digitizer makes its own, usually unknown, contribution to the overall system bias.

We start by noting that each RGB pixel in a digitized image has a value proportional to a weighted integral over the visible spectrum, an integral that depends on three spectral variables. These are the spectral irradiance of the illuminant [denoted \(E(\lambda)\), where \(\lambda\) indicates wavelength], the spectral reflectance \(S(\lambda)\) at some point in the image, and the spectral response of the \(k\)th channel in the digitizing system, \(R_k(\lambda)\). In our case, \(k\) varies from 1 to 3 (red through blue). If we denote a pixel value for the \(k\)th color channel as \(p_k\), then:

\[
p_k \propto \int S(\lambda) E(\lambda) R_k(\lambda) \, d\lambda
\]

is the fundamental relationship between pixel values and spectral quantities which Maloney and Wandell describe, and it is analogous to the CIE definitions of tristimulus values.\(^6\) However, in the CIE definitions, the digitizer’s response is replaced by the corresponding spectral transfer functions for the human eye, the color-matching functions.

Our next step is to approximate the three integrals above as summations over wavelength, using the weighted-ordinate method. If we subsume the proportionality factor in the \(R_k(\lambda)\), we can construct the matrix equation:

\[
p = S \cdot E \cdot R.
\]

In our case, \(p\) is a 1-by-3 row vector (the RGB pixel values at a given location in the image). \(S\) is a 1-by-\(m\) row vector whose elements are the surface reflectances at \(m\) equally-spaced wavelengths across the spectrum. \(E\) is an \(m\)-by-\(m\) diagonal matrix whose nonzero entries are the \(E(\lambda)\), and \(R\) is a \(m\)-by-3 matrix representing the digitizing system’s spectral transfer function.

We may think of the right side of Eq. 2 in two ways. It may consist of the known 1-by-\(m\) vector \(SE\) (the combined effects of the illuminant and reflectances) and the unknown \(R\). Alternatively, if we know the spectral illumination and the transfer function, then the 1-by-3 matrix \(ER\) is a given and the unknown quantity is the surface reflectance \(S\).

In principle, if we know elements of two of the arrays on the right side of Eq. 2 and the corresponding RGB pixel values on the left side, we can solve for the unknown array. Since we assume only an approximate knowledge of the matrix \(R\) at the outset, our tack will be to: (1) specify the system transfer function \(R\) more accurately by analyzing color samples with known reflectances; (2) use this new information to find the unknown spectral reflectances of other samples that are illuminated by the same light source.

From this knowledge of spectral reflectances, it is a straightforward matter to calculate colorimetric variables such as chromaticity and luminous reflectance.

### Description of the Calibration Procedure and Algorithm

One obvious difficulty with this approach is that for all \(m > 3\), Eq. 2 is underdetermined. As matters stand, we cannot uniquely retrieve reflectances at more than three locations across the visible spectrum. However, \(S(\lambda)\) varies smoothly in the visible for many pigments and naturally occurring materials (see Ref. 6, Section 1.4 and Ref. 4, Chapter 2). This suggests that we can accurately represent the spectral reflectances of a set of color standards with the first few components of a characteristic vector analysis.\(^7\) In effect, this analysis allows us to reduce the dimensionality of \(S\) and thus solve Eq. 2. Consistent with Cohen’s findings,\(^8\) our own analysis of the chips on a Macbeth ColorChecker chart\(^4\) shows that upwards of 98% of the variance in their spectral reflectances can be accounted for by the mean vector and weighted combinations of the first three characteristic vectors. Our choice of the ColorChecker (or any other statistical sample drawn from the population of color standards) clearly imposes a bias on the characteristic vector analysis. However, if the resulting basis vectors ultimately yield spectral reflectances that are colorimetrically accurate over a wide range of chromaticities, then the ColorChecker will be satisfactory for our purposes.

First we need to determine the system’s spectral transfer function. We start with an approximate transfer function \(R_1\) (an \(m\)-by-3 matrix), defined by the spectral sensitivity of each channel in the color camera. Next we digitize an image of the color chart, which is illuminated by the same light source that will be used when we evaluate unknown color samples. This image gives us a \(q\)-by-3 matrix of RGB values, where \(q\) is the number of ColorChecker chips \((q = 24)\). We denote this matrix as \(P\). Since we know the illumination \(E\), the set of \(q\) reflectances \(S\), and \(R_1\), we can form the \(q\)-by-3 matrix \(SER1\) which combines all of them. This leads to:

\[
P = S_{\text{mean}} (R_1 \cdot R_2),
\]

where \(R_2\) is an unknown 3-by-3 matrix. A least-squares solution of Eq. 3 for \(R_2\) leads to an improved estimate of the system’s spectral transfer function. That is, when we multiply \(R_1\) by \(R_2\) to form a new \(R\), we have included an approximation of the digitizer’s color bias. All subsequent references to \(R\) will be to this improved estimate. Despite our calculation of a new linear basis for the system’s spectral sensitivities, we have not in any way changed those sensitivities. Put another way, eye-versus-camera metamers remain metamers until we change the system hardware, either optically or electronically.

The next step in the algorithm is to rewrite Eq. 2 in terms of characteristic vectors. We define the mean vector of the ColorChecker reflectances as \(S_{\text{mean}}\) (a 1-by-\(m\) vector), and
from the first three characteristic vectors of our analysis, we form the \( b \)–by–\( m \) matrix \( S_{\text{basis}} \) (\( b = 3 \), the number of characteristic vectors used). When we multiply \( S_{\text{basis}} \) by the appropriate \( 1 \)–by–\( b \) vector of basis weights (denoted \( \mathbf{B} \)) and add the result to \( S_{\text{mean}} \), we can reconstruct any spectral reflectance \( S \) in our original set of colors. (For sake of clarity, we will refer to the algorithm’s reflectances as \( \text{reconstructed} \), and those determined independently as \( \text{measured} \).) In other words, the reconstructed reflectances are given by:

\[
S = S_{\text{mean}} + \mathbf{B} S_{\text{basis}}.
\]

Now we may rewrite Eq. 2 as:

\[
\mathbf{B} = (\mathbf{p} - S_{\text{mean}} \mathbf{E} \mathbf{R}) (S_{\text{basis}} \mathbf{E} \mathbf{R})^{-1},
\]

where \( \mathbf{E} \mathbf{R} \) is as noted above. Equation 5 is at the heart of the algorithm. We solve it by inserting into \( \mathbf{p} \) the digitized RGB values. The solution \( \mathbf{B} \) is then used in Eq. 4 to calculate the spectral reflectance vector \( S \).

As our solution to Eq. 3 implies, if we use the least-squares estimate of \( \mathbf{R} \) to reconstruct the reflectances of the color chips, some reflectances will yield significantly larger chromaticity errors than others. (We define chromaticity error as the Cartesian distance between a reconstructed chromaticity and one calculated from the corresponding measured \( S(\lambda) \).) In addition, the integrated values of the measured and reconstructed spectra will differ, and we want to correct this error as well. To improve the reconstructed spectra, we do not introduce correction factors directly into \( \mathbf{R} \), but rather apply them as multipliers to the observed RGB values for the color chips (matrix \( \mathbf{P} \)). We denote this \( q \)-by–\( 3 \) matrix of empirically determined correction factors as \( \mathbf{C} \). In other words, when we multiply corresponding elements from the \( i \)th row of \( \mathbf{C} \) and the \( i \)th row of \( \mathbf{P} \), we obtain a triplet of values \( \mathbf{p} \) that is used to solve Eq. 5. From this, we can reconstruct spectral reflectances that are colorimetrically accurate.

We calculate the elements of the correction matrix \( \mathbf{C} \) in the following way. First we perform a binary search for those multiplicative factors that minimize chromaticity errors under the given illumination. This initial \( \mathbf{C} \) still leaves us with errors in the integrated reflectances. We define these reflectance errors as the ratio of the sum of the reconstructed \( S(\lambda) \) to the sum of the measured \( S(\lambda) \), and we call this ratio \( F \). We could just as easily define \( F \) as the ratio of the integrated luminous reflectances, and this would give us a more rigorously defined estimate of relative luminance error. However, our work indicates that the two definitions of \( F \) differ very little in terms of correcting average spectral reflectance errors. Noting that \( F \) changes for each color chip, we correct the reflectance errors by dividing each row of \( \mathbf{C} \) by the appropriate value of \( F \). In changing the magnitudes of the elements of \( \mathbf{C} \), we often increase the chromaticity errors of the reconstructed \( S(\lambda) \). However, if our chromaticity error tolerance is very small in the first round of calculations, any additional error introduced at this stage is usually acceptable. If not, the offending row of \( \mathbf{C} \) is recalculated, using its current elements as a starting point.

With this approach, we can obtain good fits to the chromaticities and integrated reflectances of our original ColorChecker chips. That is, we have a way to correct the color and relative luminance bias of the digitizing system for this particular illumination. The remaining problem is to extend our reflectance retrieval technique to other RGB values from new images.

Our solution is to calculate interpolated correction factors that are based on matrix \( \mathbf{C} \). We start with the plausible assumption that the same interpolation scheme which generates a new RGB value in RGB space should also generate the appropriate correction factor from matrix \( \mathbf{C} \). In addition, we want the interpolation scheme to vary smoothly over the domain of \( \mathbf{P} \) (the RGB values of the color chips), based on the assumption that gradual changes in these RGB values are the result of gradual changes in spectral reflectance curves. One method that works well is to determine the Cartesian distance of the new RGB value from each element of \( \mathbf{P} \). We call the new color’s distance from the \( i \)th color chip \( D[\text{RGB}(i)] \), and we use the negative exponential of this distance to calculate the \( i \)th element of a \( 1 \)-by–\( q \) vector of inverse-distance weights (vector \( \mathbf{W} \)). Specifically, the \( i \)th element of \( \mathbf{W} \) is:

\[
W(i) = \exp[h \cdot D[\text{RGB}(i)]/D[\text{maximum}]],
\]

where \( h \) is a constant less than zero, and \( D[\text{maximum}] \) is the distance between the minimum and maximum RGB values that can be generated by the digitizer. The value of \( h \) determines how rapidly the weights decrease as we move away from the original ColorChecker RGB values (\( h = \ln(10^{-9}) \) in our work). If \( h \) is made more negative, the weights are more sharply peaked about that ColorChecker swatch which is closest in RGB space to the new color. In essence, Eq. 6 scales RGB distances so that if a ColorChecker chip is sufficiently different from the new color, the chip’s effect on the retrieval is nil.

If we normalize the \( 1 \)-by–\( q \) weighting vector \( \mathbf{W} \) by its sum and multiply the result by our matrix of empirical corrections (the \( q \)-by–\( 3 \) matrix \( \mathbf{C} \)), we obtain a smoothly interpolated correction factor (the \( 1 \)-by–\( 3 \) vector \( \mathbf{c} \)). However, matrix \( \mathbf{C} \) was calculated on the tacit assumption that, at the \( i \)th ColorChecker swatch, all elements of \( \mathbf{W} \) are zero except the \( i \)th element, which equals one. Our interpolation scheme approximates this by making the \( i \)th element of \( \mathbf{W} \) the largest, but it does not equal one, nor are the other elements equal to zero.

We can solve this new problem by forming the \( q \)-by–\( q \) matrix \( \mathbf{W} \), whose \( i \)th row is the vector \( \mathbf{W} \) that results from calculating the inverse-distance weights for the \( i \)th row of the ColorChecker RGB data (matrix \( \mathbf{P} \)). This allows us to define a new correction matrix \( \mathbf{C} \) (\( q \)-by–\( 3 \)) which will yield the original correction factors \( \mathbf{C} \) when multiplied by \( \mathbf{W} \):

\[
\mathbf{C} = \mathbf{W}^{-1} \mathbf{C}.
\]

Now we may redefine the vector of corrections \( \mathbf{c} \) as:

\[
\mathbf{c} = \mathbf{W} \mathbf{C}.
\]
generated the elements of $W$. In practice, the $k$th value of the RGB vector $p$ at any pixel is modified so that:

$$p_{k, \text{modified}} = p_k c_k. \quad (9)$$

It is these altered $p_k$ that we use in solving Eq. 5 and in recovering the spectral reflectance $S$ that generated the pixel.

Sample Retrieval

To date, tests of the algorithm show that it can do well in estimating reflectances and chromaticities. The color samples used in these tests are pieces of uniformly-colored paper which are illuminated by a known source of light (we used artist’s papers and an ordinary slide projector). Care needs to be taken that all spectral measurements and digitized images are made in the same part of the projector’s illumination field, which varies appreciably in luminance. First we measure the spectral reflectances of the paper samples with a spectroradiometer. These spectral reflectance data give us standards for assessing the accuracy of the retrieval algorithm. As before, we distinguish between these measured reflectance standards and the algorithm’s reconstructed reflectances.

The papers tested are fairly chromatic samples of red, green, blue, and orange. Their measured spectral reflectances are shown in Figs. 1–4. To test the algorithm, we digitize video images of these papers, averaging each sample’s RGB values over both area and time. Next, the ColorChecker card’s color chips are similarly averaged, and we form the data matrix $P$. Using this, we calculate the improved spectral transfer function $R$. Now we can determine the empirical corrections (matrix $C_2$) that further improve the algorithm’s fit to the known ColorChecker spectral reflectances.

In Figs. 5–6, we show reconstructed reflectances for both the worst and one of the best fits to the ColorChecker samples. One of our criteria for goodness of fit is the distance between the measured and reconstructed chromaticities. Another is the root mean square (RMS) percentage difference between the reconstructed and measured spectral reflectances. Note that our constraints in calculating $C_2$ did not rule out spectral reflectances less than 0% or greater than 100%, but merely required that the integrated reflectances differ as little as possible between the measured and reconstructed cases. The additional constraint of minimizing chromaticity errors usually precludes such nonphysical reflectances, but not always. As we might expect from our constraints, chromaticity and reflectance errors for the ColorChecker are small, averaging 0.00076 and 3.55%, respectively. Typical chromaticity distance errors are shown in Fig. 7. For sake of illustration, the chromaticity of the illuminant is also included. Note that this yellowish light source shifts the entire gamut of ColorChecker chromaticities.
FIG. 4. Measured and reconstructed spectral reflectances for an orange test sample. RMS reflectance difference 12.2%.

ities toward longer wavelengths, although it affects their positions with respect to each other only slightly.

When we turn to chromaticity distance errors in the test samples (Fig. 8), the agreement is still good. The average error for all four samples is 0.0047, and the maximum error is less than 0.0092 (the red sample). This largest error is roughly seven times the length of the minor semiaxis of the nearest measured MacAdam ellipse. These results are encouraging, but we need to ask whether we are really interpolating between chromaticities from the color chart, or whether we have fortuitously chosen test samples that nearly coincide with that data. Figure 9 shows that we are indeed interpolating (apparently accurately) between the chromaticities of the ColorChecker chips. Further, if we interpolate linearly between the RGB values of matrix P and plot the resulting chromaticities, we see that smooth (but not necessarily straight) lines connect the ColorChecker chromaticity coordinates (Fig. 10). This is another of our goals in designing the algorithm: gradual changes in digitized RGB values should result in gradual changes in chromaticity.

The RMS reflectance differences for our test samples average 7.77%, a figure which is slightly more than double that for the calibration data. However, individual errors in spectral reflectances can be noticeably larger, as is evident in Figures 1–4. In particular, the nonphysical reflectances for the green and orange test samples show that we cannot regard a video camera that uses our algorithm as the equivalent of a spectroradiometer. However, we neither claim nor aim to do so. We are principally interested in obtaining accurate chromaticity data and in estimating the integrated values of the spectral reflectances. Tests show that the range

FIG. 5. Measured and reconstructed spectral reflectances for the ColorChecker chip designated "Yellow". RMS reflectance difference 1.38%; chromaticity distance 0.0004.

FIG. 6. Measured and reconstructed spectral reflectances for the ColorChecker chip "Bluish Green". RMS reflectance difference 6.66%; chromaticity distance 0.0005.

FIG. 7. CIE 1931 chromaticity diagram (2° standard observer), showing measured and reconstructed chromaticities for most of the ColorChecker chips.
of physically realistic reflectances calculated by the algorithm is roughly defined by the gamut of ColorChecker chromaticities. Within limits, increasing this gamut would expand the range of reliably reconstructed reflectances. However, for many naturally occurring materials, this range may be sufficient. Even without altering the color card, we see that the reconstructed reflectance curves are nearly metameric for chromaticities and can approximate the measured curves' tristimulus values.

Given the constraints derived from the calibration data, it is unlikely (but not impossible) that the estimated curves will depart radically from the shapes of the true transmittances or reflectances. Clearly there are some radiant energy spectra that can be measured reliably only with a spectroradiometer. As we have demonstrated, though, there are many cases where such elaborate instrumentation is unnecessary. Equally important is the fact that it may not be available. Often a video or still camera will be our only data-gathering equipment when we are confronted with a colorimetrically interesting scene. Then by definition each part of the captured image contains only three pieces of information, whether RGB values or film dye densities. Using a spectroradiometer to analyze this kind of image would likely offer no additional information. We had this common situation in mind when we wrote the colorimetric calibration routine.

**Capabilities and Caveats**

Clearly, one has a potentially powerful image analysis tool with this scheme for recovering spectral reflectances. Our impetus for developing it was to perform colorimetric analysis of phenomena in meteorological optics such as rainbows and halos. However, the technique can extend far beyond this range, into a variety of scientific and commercial applications.

One asset of this scheme is that it can account for bias not only in the camera and digitizer, but also in any other links in the optical chain, such as camera optics. If we incorporate any new links in the optical chain and repeat
the calibration procedure with the color chart, we should attain nearly the same degree of accuracy. In fact, even if we use storage media such as color transparencies or video tape in making digitized images, it is possible to compensate for their color biases too.

While we have claimed that an accurate knowledge of the spectral illumination is a prerequisite for calibration, it seems reasonable that slight errors in these data are unlikely to seriously affect the accuracy of the recovery. This is because illumination errors are accounted for in $R$ and the empirical corrections $C_2$. Provided that the same illumination holds for test images, the interpolation scheme still requires that reconstructed reflectances vary between known spectral reflectances. Naturally, errors in our knowledge of the spectral quality of the illumination will affect calculation of chromaticities. Another potential complication is changes in the spectral character of the illumination across the image, although this is not a problem in many natural scenes. Other practical pitfalls in color analysis are discussed in Reference 5.

As is true in any calibration scheme, violation of one's initial assumptions can quickly lead to foolishness. Corrections derived at one calibration session are unlikely to be valid at the next. Color settings on the digitizer and camera must be maintained between the calibration and test images. If we fail to account for differences in ambient illumination levels in the test and calibration images, errors will follow. Each time that we change to a different type of color film, video tape, or video recorder, we must recalibrate.

More basic than these problems are some limitations inherent in the equipment itself and in our assumptions. While our scheme provides smooth interpolations based on an ensemble of corrections, some minimal precautions need to be observed. All types of imaging systems can be made hopelessly distorted, as would be the case with using faded or overexposed color slides. A video camera with severe chroma or alignment errors will render meaningless results. In addition, our assumption in applying the characteristic vector analysis (namely, that $S(\lambda)$ varies smoothly with wavelength) means that our technique can only approximate spectra that have a substantial amount of fine detail. Finally, we should reemphasize that we have not eliminated eyeversus-camera metamerism. To do so would require changes in the system's hardware. However, where such metamers do not exist, our algorithm substantially improves the mapping between the two visual systems.

Conclusions

We have described a simple method of calibrating a color video digitizing system which permits retrieval of colorimetrically accurate spectral reflectances or transmittances.

This and related methods should prove helpful to many workers concerned with color appearance and color rendition. As we indicate above, our system has some limitations, both functional and theoretical. Preliminary tests of the algorithm and equipment show that, with a modicum of care, these problems can be minimized. Accordingly, we expect to realize the system's promise of quick and accurate measurement of colorimetric variables from a wide variety of sources (film, video tape, and live video). These advantages should make our system a valuable tool in image analysis.

Acknowledgments

The author gratefully acknowledges the assistance of Alastair B. Fraser and Craig F. Bohren in suggesting and perceptively criticizing this work. I am also indebted to Robert T. Marcus of Macbeth, whose spectral reflectance measurements of the ColorChecker card are crucial to many of the calculations made here. Brian A. Wandell and Laurence T. Maloney, who have presented in print the theoretical basis for this work, also have generously provided me with valuable papers from their own files. Much of this project was sponsored by National Science Foundation grant number ATM-8607577.

1. Strictly speaking, luminance is defined only for the CIE standard photopic observer, and also requires knowledge of the absolute spectral radiance. However, our algorithm constructs a photometrically scaled measure of the digitizing system's response to spectral radiance and reflectance. This means that we can calculate luminous reflectances in a digitized image, although radiance measurements would be needed to determine absolute luminances.


Received August 13, 1987; accepted December 26, 1987.