Improving color measurements of displays

Jonathan E. Hardis

Radiometric Physics Division, National Institute of Standards & Technology
Technology Administration, U. S. Department of Commerce
Gaithersburg, MD 20899-0001

ABSTRACT

It is generally believed that the most accurate means of measuring the CIE tristimulus values X, Y, Z or chromaticity coordinates x, y from a display is by using a spectroradiometer. Nevertheless, tristimulus colorimeters employing three or four colored filters find wide use because of their simplicity and lower cost. These devices cannot be calibrated to give accurate results in all situations because the spectral responsivities of their filtered detectors are not exactly the CIE color-matching functions. However, for a display that produces a linear superposition of three primary colored lights of fixed spectra, a tristimulus colorimeter can be correctly calibrated to measure all colors on that display. Signals from all of the filtered detectors are used to compute each of the X, Y, Z values. The calibration matrix is computed by data fitting to a reference colorimeter. An improvement to the previously published method is reported, and a numerical example is shown. This technique is more tractable with today's digital instrumentation than it was when it was discovered, yet it remains underused. The American Society for Testing and Materials (ASTM), through its Committee on Color and Appearance, is revising its standard on display measurements using tristimulus colorimeters to encourage the adoption of the technique.

Keywords: ASTM, calibration, chromaticity, CIE, color, colorimeter, matrix transformation, measurement, standards, tristimulus

1. INTRODUCTION

It pays to remind ourselves that we cannot truly measure color. Color is a subjective quality, not a physical quantity. Nevertheless, color metrology is an important science. Measurements made of light spectra, interpreted as to color, serve the purpose in specifications, regulations, and many other aspects of everyday commerce. Our common concern is that these data should be accurate, faithfully reporting what actually occurred, however limited they might be by the applicability of the visual models in which they are used.

This paper speaks to a long-known but underutilized method of improving the accuracy of color measurements made on displays, such as CRTs and self-luminous flat-panel displays. It discusses improvements to the method, and it puts into context additional issues that may be explored further.

2. BACKGROUND

2.1 Colorimetry

Color measurement instruments consist, in general, of means to measure radiometric power as transmitted through a number of bandpass filters. Most commonly, electrical devices are used to measure the filtered light. They may be used with different filters in succession, or multiple devices may be used concurrently. In instruments called spectroradiometers, the radiometric power is measured through a large number (typically 30 to 500) of narrowband filters. In instruments called tristimulus colorimeters, the radiometric power is measured through three or four wideband filters. These filters may be constructed from dispersive elements (prisms and gratings) or from materials with selective spectral transmission or reflection. The latter may be either uniform or comprised of different patches, in a mosaic pattern, that provide the desired overall effect.

No matter how many filters are used, or in what manner, the goal of the measurement process is to determine tristimulus values X, Y, Z, as defined by the Commission Internationale de L'Éclairage (CIE)[1] in its publications.[2] For light with a spectral power distribution $\varphi(\lambda)$,

---

0-8194-2032-8/96/$6.00

$$
\begin{aligned}
X &= k\int_{360\,\mathrm{nm}}^{830\,\mathrm{nm}} \varphi(\lambda)\,\bar{x}(\lambda)\,d\lambda \\
Y &= k\int_{360\,\mathrm{nm}}^{830\,\mathrm{nm}} \varphi(\lambda)\,\bar{y}(\lambda)\,d\lambda \quad , \\
Z &= k\int_{360\,\mathrm{nm}}^{830\,\mathrm{nm}} \varphi(\lambda)\,\bar{z}(\lambda)\,d\lambda
\end{aligned}
\tag{1}
$$

where k is a constant that depends on the circumstances (and which is irrelevant if X, Y, Z will be used only to compute chromaticity coordinates x, y), and $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ are color-matching functions.[3] While the standard definition of X, Y, Z requires the use of the CIE 1931 2° color-matching functions, the mathematics described in this paper would also be applicable to any other set of color-matching functions, such as the CIE 1964 10° functions.

In practice, color measurement instruments compute X, Y, Z by the summation of the signals as measured through the various filters, each signal being multiplied by an appropriate calibration factor. In matrix notation:

$$
\begin{pmatrix} X_\mathrm{m} \\ Y_\mathrm{m} \\ Z_\mathrm{m} \end{pmatrix} = \begin{pmatrix} C_{\mathrm{X}1} & C_{\mathrm{X}2} & C_{\mathrm{X}3} & \cdots & C_{\mathrm{X}f} \\ C_{\mathrm{Y}1} & C_{\mathrm{Y}2} & C_{\mathrm{Y}3} & \cdots & C_{\mathrm{Y}f} \\ C_{\mathrm{Z}1} & C_{\mathrm{Z}2} & C_{\mathrm{Z}3} & \cdots & C_{\mathrm{Z}f} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \\ F_3 \\ \vdots \\ F_f \end{pmatrix} \quad , \tag{2}
$$

where F_1, F_2, F_3, through F_f are the electrical signals from the f filtered detectors and the C_{ij} are calibration coefficients. X_m, Y_m, Z_m have subscripts to indicate that they are measured values rather than ideal ones.

We presume that the color measuring instrument is linear: that each signal F_a is strictly proportional to the received optical power, that any zero-offset (background in darkness) is removed, that the proportionality for signal F_a is not affected by the value of signal F_b, and in the case of closely packed detectors (such as CCD detector elements) no signal F_a spills over and affects signal F_b as it approaches saturation. These presumptions are amenable to experimental verification using methods beyond the scope of this paper.[4]

The values of the matrix elements C_{ij} may be determined using criteria that depends on the design and intended application of the instrument. The full extent of this issue is beyond the scope of this paper. However, in general, for spectroradiometers ($f \approx 30$ to 500), $C_{\mathrm{X}j}$ reflects the tabulated value of $\bar{x}(\lambda)$ near the center wavelength of Filter j as well as the spectral responsivity of the corresponding detector channel. (Likewise, $C_{\mathrm{Y}j}$ and $C_{\mathrm{Z}j}$ reflect $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$, respectively.) For tristimulus colorimeters, the choice of C_{ij} is discussed further, below. As a general matter, the instrument designer should choose passbands and matrix elements that balance accuracy, sensitivity, and other design requirements.

Tristimulus colorimeters are generally designed with filters that are intended to match the spectral responsivities of their detector channels to the CIE $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ functions. For such an instrument,

$$
\begin{pmatrix} X_\mathrm{m} \\ Y_\mathrm{m} \\ Z_\mathrm{m} \end{pmatrix} = \begin{pmatrix} C_{\mathrm{X}1} & 0 & 0 \\ 0 & C_{\mathrm{Y}2} & 0 \\ 0 & 0 & C_{\mathrm{Z}3} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix} \quad , \tag{3}
$$

where the non-zero C_{ij} matrix elements represent adjustable gains of the detector channels. However, the $\bar{x}(\lambda)$ function has two distinct lobes. This may be dealt with by splitting $\bar{x}(\lambda)$ into $\bar{x}_\mathrm{short}(\lambda)$ and $\bar{x}_\mathrm{long}(\lambda)$, each with a separate filter (F_1 and F_2, respectively). For such an instrument,

$$
\begin{pmatrix} X_\mathrm{m} \\ Y_\mathrm{m} \\ Z_\mathrm{m} \end{pmatrix} = \begin{pmatrix} C_{\mathrm{X}1} & C_{\mathrm{X}2} & 0 & 0 \\ 0 & 0 & C_{\mathrm{Y}3} & 0 \\ 0 & 0 & 0 & C_{\mathrm{Z}4} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \end{pmatrix} \quad . \tag{4}
$$

Alternatively, the $\bar{z}(\lambda)$ function may serve the role of $\bar{x}_\mathrm{short}(\lambda)$ since they have a similar shape,

$$
\begin{pmatrix} X_\mathrm{m} \\ Y_\mathrm{m} \\ Z_\mathrm{m} \end{pmatrix} = \begin{pmatrix} C_{\mathrm{X}1} & 0 & C_{\mathrm{X}3} \\ 0 & C_{\mathrm{Y}2} & 0 \\ 0 & 0 & C_{\mathrm{Z}3} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix} \quad . \tag{5}
$$

In all of these cases, it is difficult to realize an exact match between the CIE color-matching functions and the actual spectral responsivities of the corresponding detector channels. This means that no choice of C_{ij} will provide perfect calibration for all applications of the instrument. The criteria for setting the C_{ij} might not be well documented for a particular instrument.

It is generally believed that spectroradiometers, with their many detector channels, may be calibrated to yield superior measurements of X, Y, Z in many applications. Nevertheless, the relative simplicity of tristimulus colorimeters and their commensurately lower cost have made them popular where the highest accuracy is not required.

2.2 Self-luminous displays

A self-luminous display, such as a CRT, an electroluminescent (EL) panel, a field emission display (FED), or a backlit liquid crystal display (LCD) generates colored light by the proportional superposition (addition) of primary colored lights $p_r(\lambda)$, $p_g(\lambda)$, $p_b(\lambda)$. The subscripts represent red, green, and blue, the primary colors of an additive set. An arbitrarily colored patch on the visual display has one and only one spectral power distribution $\varphi(\lambda)$,

$$\varphi(\lambda) = a p_r(\lambda) + b p_g(\lambda) + c p_b(\lambda) \quad , \tag{6}$$

where a, b, c are coefficients that are determined by the display electronics.

The display electronics vary a, b, c over the face of the display in order to generate a colored image. We presume that the display electronics may be set to make a, b, c uniform (perhaps after averaging non-obvious fine-structure) over a sufficient area of the display to permit measurements to be made on that area.

This is the key fact from which the present method derives: the spectrum from such a display is not entirely arbitrary; it must fall within a restricted domain. The spectral radiance from such a display is highly correlated between different wavelengths. While it is a requirement for the applicability of this method that the display device behaves as stated in Eq. 6, and while those within the mentioned classes of devices might do so, the procedure for experimental verification of this property for any specific display device is beyond the scope of this paper.[5]

2.3 Colorimetric measurement of displays

Each of the primary spectral power distributions $p_r(\lambda)$, $p_g(\lambda)$, $p_b(\lambda)$ stimulates responses in the f detector channels that may be represented by a vector $\boldsymbol{F}$ (*i.e.*, $\boldsymbol{F}_r$, $\boldsymbol{F}_g$, $\boldsymbol{F}_b$). Given their construction, these vectors are linearly independent. (Any of the three cannot be expressed as a linear combination of the other two.) While $\boldsymbol{F}$ is an element of an f-dimensional vector space, it is clear that only a three-dimensional subspace is spanned by the $\boldsymbol{F}$'s of all possible spectral power distributions following Eq. 6. Further, the mapping of $\boldsymbol{F}$ into (X_m, Y_m, Z_m) space by Eq. 2 remains three dimensional. In other words, there is a one-to-one mapping of the vector (a, b, c) onto (X, Y, Z) by application of Eq. 1; and, for a particular instrument with a fixed calibration matrix C, there is also a one-to-one mapping of the vector (a, b, c) onto (X_m, Y_m, Z_m). From this we deduce that a matrix R exists that can be used to translate (X_m, Y_m, Z_m) values into actual (X, Y, Z) values.

A colorimeter that takes advantage of this fact must provide means for implementing the matrix R. That is, all f filtered detector signals should contribute linearly toward the computation of each output, X_m, Y_m, Z_m, instead of using different detectors for each output. This idea was reported as long ago as 1973 by Wagner,[6] and it has been expanded upon and rediscovered by others since then.[7,8,9,10,11]

On the basis of this property, a tristimulus colorimeter can be optimized for use on a self-luminous display by the proper derivation of a matrix R for that display. We proceed on the assumptions that the components are sufficiently stable, and that similarly built displays have similar enough spectral primaries to make a derivation of R worthwhile. However, these assumptions should be quantified before accuracy claims are made in any specific situation.

If we are careful to limit the use of a colorimeter to a display for which R is well chosen, we need not be concerned that the spectral responsivities of the filtered detectors do not exactly replicate the CIE color-matching functions. Indeed, it may be advantageous if they do not. Signal/noise may be improved by matching the spectral responsivities of the filtered detectors to the spectra of the primary colors. (For example, the emission peaks of common CRT red phosphors lie in the long-wavelength wing of the $\bar{x}(\lambda)$ function.) When there are more than three detectors, the C_{ij} may be chosen to best discriminate between the primary spectra in (X_m, Y_m, Z_m) space.

3. OPTIMIZATION OF COLORIMETERS

Given the existence of a matrix R, how does one determine it? Experimentally, the problem is one of comparing the data X, Y, Z from a reference colorimeter with the data X_m, Y_m, Z_m from the colorimeter being optimized for a number of color samples at different display settings. From these data, R is calculated.

This method is not directly concerned with the absolute accuracy of the measurements. It concerns the transfer of calibration from a reference instrument to another instrument, regardless of the absolute accuracy of the reference. The later is an interesting problem in its own right, but beyond the scope of this paper.

3.1 Noiseless data

For clarity, we first consider the ideal case where the colorimeters are free of noise. We define vectors **n** and **m** as:

$$\boldsymbol{n} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}; \; \boldsymbol{m} = \begin{pmatrix} X_m \\ Y_m \\ Z_m \end{pmatrix} \quad , \tag{7}$$

where each is an element of its corresponding vector space, and both derive from the same (a, b, c) setting on a display. The matrix R maps between them,

$$\boldsymbol{n} = R \; \boldsymbol{m} \quad . \tag{8}$$

This relationship may be stated for more than one such pairs of vectors (for multiple display settings) at the same time:

$$N = R \; M \quad , \tag{9}$$

where

$$N = \begin{pmatrix} X_1 \, X_2 \, X_3 \cdots X_i \\ Y_1 \, Y_2 \, Y_3 \cdots Y_i \\ Z_1 \, Z_2 \, Z_3 \cdots Z_i \end{pmatrix} \tag{10}$$

and

$$M = \begin{pmatrix} X_{m1} \, X_{m2} \, X_{m3} \cdots X_{mi} \\ Y_{m1} \, Y_{m2} \, Y_{m3} \cdots Y_{mi} \\ Z_{m1} \, Z_{m2} \, Z_{m3} \cdots Z_{mi} \end{pmatrix} \quad . \tag{11}$$

When matrices N and M have exactly three columns, and when their columns are linearly independent, R may be easily determined:

$$R = N \; M^{-1} \quad . \tag{12}$$

The determination of matrix R in this case requires the reference values and the test colorimeter measurements of exactly three distinct colors on the display.

3.2 Real-world transformations of X, Y, Z

In practice, neither measurements of $\boldsymbol{n}$ nor $\boldsymbol{m}$ are made with perfect accuracy. Noise affects the measurements, the linearity presumptions mentioned in Section 2 may not be perfectly true, and there may be other unexpected systematic effects in the data. Therefore, it is prudent to determine R by using more than three color samples and by using statistical methods.

For a given R_0 and for a pair of related $\boldsymbol{n}$ and $\boldsymbol{m}$ vectors,

$$\boldsymbol{n} = R_0 \; \boldsymbol{m} + \boldsymbol{v} \quad , \tag{13}$$

where $\boldsymbol{n}$ and $\boldsymbol{m}$ are as in Eq. 7, and $\boldsymbol{v}$ is the difference vector between the reference tristimulus values $\boldsymbol{n}$ and those computed by the use of Eq. 8. For several different colors Eq. 13 can be expressed as:

$$N = R_0 \; M + V \quad , \tag{14}$$

where N is as in Eq. 10, M is as in Eq. 11, and

$$V = \begin{pmatrix} v_{11} \, v_{12} \, v_{13} \cdots v_{1i} \\ v_{21} \, v_{22} \, v_{23} \cdots v_{2i} \\ v_{31} \, v_{32} \, v_{33} \cdots v_{3i} \end{pmatrix} \quad . \tag{15}$$

When many measurements $\boldsymbol{n}$ and $\boldsymbol{m}$ are available, the optimal R_0, hereafter called R', may be determined by minimization of the sum of the squares of the elements of V with respect to all nine elements of R'. This is the statement of least-squares fitting when applied to the vector space of R. It follows that:

$$R' = N\ [(M\,M^T)^{-1}\,M]^T \quad , \tag{16}$$

where the superscript T implies matrix transposition, and where we presume that the statistical uncertainty (as opposed to the measurement standard uncertainty) of all X, Y, Z and X_m, Y_m, Z_m are the same, that is, each is known to ± the same value. As with all least-squares fitting procedures, the results will be affected by the portions of R space that are sampled the most and by the presumption of identical statistical weights (uncertainties) for each data sample. When R' is used for R_0 in Eq. 14, the residuals V may be examined for trends that might indicate systematic effects.

Some published versions of Eq. 16 (and Eq. 12) show M as potentially having four (or more) rows, one for each of the f filtered detectors. This is incorrect. If M had more than three rows, the physics here are such that $M\,M^T$ would be singular, but for measurement noise and small systematic effects. The stability of the solution for R' requires that the dimensionality of the problem be reduced from f to 3 by use of the C matrix in Eq. 2. The optimal choice of C for the purpose of dimensionality reduction goes beyond the scope of this paper. As implied in Section 2, this aspect of C must be based on additional factors besides display measurement accuracy. For example, when $f > 3$, the reduction in dimensionality can follow the Hilbert-space minimization approach of Lux and Schanda.[7]

3.3 Real-world transformations of Y, x, y

Equation 16 is the principal result of Wagner.[6] However, there is an important situation where all of X, Y, Z and X_m, Y_m, Z_m are not known with equal uncertainties. Many commercial instruments report Y, x, y instead of X, Y, Z, where

$$x = \frac{X}{X+Y+Z}\ ;\quad y = \frac{Y}{X+Y+Z}\quad . \tag{17}$$

(We have dropped the m subscripts in Eq. 17 for clarity.)

There are at least two good reasons for this convention. First, displays typically flicker as they are scanned, and measuring instruments may not synchronize to, or sufficiently average over the temporal variations in spectral radiance. The resultant measurement noise would be less pronounced in x, y than it would be in X, Y, Z. Second, in many cases the user really does want to know Y, x, y rather than X, Y, Z, and the instrument vendor is meeting that need. However, in order for this method to apply, the instrument must transform X, Y, Z to Y, x, y accurately. Instruments with analog circuits that only approximate Eq. 17 are not covered here.

Those who compute X, Y, Z from Y, x, y data (likewise for X_m, Y_m, Z_m) and try to apply Eq. 16 to it get a rude shock. Application of that R' to optimize a colorimeter may well make it worse in x, y rather than better. The root of the problem is that the transformations from X, Y, Z to Y, x, y (internally, in the colorimeter) and back again to X, Y, Z do not preserve the statistical uncertainties needed for the best fit.

Calculation of R' for this case cannot be done as compactly as it was in Eq. 16. Instead, we must calculate it a row at a time. We express

$$R' = \begin{pmatrix} R'_{XX} & R'_{XY} & R'_{XZ} \\ R'_{YX} & R'_{YY} & R'_{YZ} \\ R'_{ZX} & R'_{ZY} & R'_{ZZ} \end{pmatrix} \quad , \tag{18}$$

and we say that

$$\begin{pmatrix} R'_{YX} & R'_{YY} & R'_{YZ} \end{pmatrix} = N_Y\ [(M_Y M_Y^T)^{-1}\, M_Y]^T \quad , \tag{19}$$

where

$$N_Y = \begin{pmatrix} Y_1 & Y_2 & Y_3 & \cdots & Y_i \end{pmatrix} \quad , \tag{20}$$

and

$$M_Y = \begin{pmatrix} X_{m1} & X_{m2} & X_{m3} & \cdots & X_{mi} \\ Y_{m1} & Y_{m2} & Y_{m3} & \cdots & Y_{mi} \\ Z_{m1} & Z_{m2} & Z_{m3} & \cdots & Z_{mi} \end{pmatrix} \quad \text{(as in Eq. 11)}, \tag{21}$$

with

$$X_{mi} = Y_{mi}\frac{x_{mi}}{y_{mi}}\ , \quad Z_{mi} = Y_{mi}\frac{\left(1 - x_{mi} - y_{mi}\right)}{y_{mi}}\quad . \tag{22}$$

With this partial result, it is useful to compute

$$\left(Y'_1\, Y'_2\, Y'_3 \cdots Y'_i\right) = \left(R'_{\mathrm{YX}}\, R'_{\mathrm{YY}}\, R'_{\mathrm{YZ}}\right) M_{\mathrm{Y}} \quad , \tag{23}$$

where the Y'_i are the best-fit luminances of the sample display settings.

Turning to X, at first one might think that N_{X} would be

$$\left(Y_1\frac{x_1}{y_1}\; Y_2\frac{x_2}{y_2}\; Y_3\frac{x_3}{y_3} \cdots Y_i\frac{x_i}{y_i}\right) \quad ,$$

and that M_{X} would be the same as M_{Y}. However, this is not the case. We need to incorporate the statistical weighting in these matrix elements. Also, in some circumstances it would be better to use Y' rather than Y values in N_{X}. For example, if the flicker effect mentioned earlier is significant, it already has been accounted for to the extent that it will be in these calculations. Use of Y rather than Y' values at this point would transfer measurement noise from the luminance to the chromaticity coordinates. On the other hand, if some aspect of the display is varying slowly in time, or actually varies based on the display setting, then it is better to use Y rather than Y' values in order to make the best comparisons between individual readings of the two instruments. While we will display the equations using Y' rather than Y, both sets of formulae should be considered.

With statistical weighting,[12]

$$N_{\mathrm{X}} = \left(\frac{Y'_1\frac{x_1}{y_1}}{\sigma_{\mathrm{X}1}}\; \frac{Y'_2\frac{x_2}{y_2}}{\sigma_{\mathrm{X}2}}\; \frac{Y'_3\frac{x_3}{y_3}}{\sigma_{\mathrm{X}3}} \cdots \frac{Y'_i\frac{x_i}{y_i}}{\sigma_{\mathrm{X}i}}\right) \tag{24}$$

and

$$M_{\mathrm{X}} = \begin{pmatrix} \frac{X_{\mathrm{m}1}}{\sigma_{\mathrm{X}1}} & \frac{X_{\mathrm{m}2}}{\sigma_{\mathrm{X}2}} & \frac{X_{\mathrm{m}3}}{\sigma_{\mathrm{X}3}} & \cdots & \frac{X_{\mathrm{m}i}}{\sigma_{\mathrm{X}i}} \\ \frac{Y_{\mathrm{m}1}}{\sigma_{\mathrm{X}1}} & \frac{Y_{\mathrm{m}2}}{\sigma_{\mathrm{X}2}} & \frac{Y_{\mathrm{m}3}}{\sigma_{\mathrm{X}3}} & \cdots & \frac{Y_{\mathrm{m}i}}{\sigma_{\mathrm{X}i}} \\ \frac{Z_{\mathrm{m}1}}{\sigma_{\mathrm{X}1}} & \frac{Z_{\mathrm{m}2}}{\sigma_{\mathrm{X}2}} & \frac{Z_{\mathrm{m}3}}{\sigma_{\mathrm{X}3}} & \cdots & \frac{Z_{\mathrm{m}i}}{\sigma_{\mathrm{X}i}} \end{pmatrix} \quad . \tag{25}$$

Ideally we would know the actual statistical uncertainties $\sigma_{\mathrm{X}i}$, but we consider the situation where the two instruments (reference and test) both report x, y to ± 0.001. In this case, the measurement precision is limited by the truncation and rounding to three digits. Following the standard rules for propagation of uncertainties, the relative statistical uncertainties are

$$\sigma_{\mathrm{X}i} = Y'_i\,\frac{x_i}{y_i}\sqrt{\left(\frac{\Delta x}{x_i}\right)^2 + \left(\frac{\Delta y}{y_i}\right)^2 + \left(\frac{\Delta Y}{Y'_i}\right)^2} \quad , \tag{26}$$

where Δx, Δy (and later, Δz) are taken as 0.001 (for the sake of example, as these are roughly the standard uncertainties that might be expected in this situation), and ΔY can be taken as the experimental standard deviation of the set of values $\{(Y'_i - Y_i)\}$. From this, we can calculate

$$\left(R'_{\mathrm{XX}}\, R'_{\mathrm{XY}}\, R'_{\mathrm{XZ}}\right) = N_{\mathrm{X}}\; [(M_{\mathrm{X}} M_{\mathrm{X}}^T)^{-1} M_{\mathrm{X}}]^T \quad . \tag{27}$$

Similarly,

$$N_{\mathrm{Z}} = \left(\frac{Y'_1\frac{z_1}{y_1}}{\sigma_{\mathrm{Z}1}}\; \frac{Y'_2\frac{z_2}{y_2}}{\sigma_{\mathrm{Z}2}}\; \frac{Y'_3\frac{z_3}{y_3}}{\sigma_{\mathrm{Z}3}} \cdots \frac{Y'_i\frac{z_i}{y_i}}{\sigma_{\mathrm{Z}i}}\right) \tag{28}$$

and

$$M_{\mathrm{Z}} = \begin{pmatrix} \frac{X_{\mathrm{m}1}}{\sigma_{\mathrm{Z}1}} & \frac{X_{\mathrm{m}2}}{\sigma_{\mathrm{Z}2}} & \frac{X_{\mathrm{m}3}}{\sigma_{\mathrm{Z}3}} & \cdots & \frac{X_{\mathrm{m}i}}{\sigma_{\mathrm{Z}i}} \\ \frac{Y_{\mathrm{m}1}}{\sigma_{\mathrm{Z}1}} & \frac{Y_{\mathrm{m}2}}{\sigma_{\mathrm{Z}2}} & \frac{Y_{\mathrm{m}3}}{\sigma_{\mathrm{Z}3}} & \cdots & \frac{Y_{\mathrm{m}i}}{\sigma_{\mathrm{Z}i}} \\ \frac{Z_{\mathrm{m}1}}{\sigma_{\mathrm{Z}1}} & \frac{Z_{\mathrm{m}2}}{\sigma_{\mathrm{Z}2}} & \frac{Z_{\mathrm{m}3}}{\sigma_{\mathrm{Z}3}} & \cdots & \frac{Z_{\mathrm{m}i}}{\sigma_{\mathrm{Z}i}} \end{pmatrix} \quad , \tag{29}$$

where
$$z_i = 1 - x_i - y_i \ , \quad \sigma_{Zi} = Y'_i \frac{z_i}{y_i} \sqrt{\left(\frac{\Delta z}{z_i}\right)^2 + \left(\frac{\Delta y}{y_i}\right)^2 + \left(\frac{\Delta Y}{Y'_i}\right)^2} \ , \tag{30}$$

and
$$\left(R'_{\mathrm{ZX}}\, R'_{\mathrm{ZY}}\, R'_{\mathrm{ZZ}}\right) = N_{\mathrm{Z}}\ [(M_{\mathrm{Z}} M_{\mathrm{Z}}^T)^{-1} M_{\mathrm{Z}}]^T \ . \tag{31}$$

4. DISCUSSION

This paper primarily concerns the spectral characteristics of colorimeters, but not the errors that might arise due to the geometries of the light collection in either the reference or the target instrument. For example, the spectral radiance of an active matrix liquid crystal (AMLCD) display is highly directional. Such a display could not be used to transfer calibration between instruments with different numerical apertures. Also, displays exhibit spatial nonuniformities in their radiance. The inability of dissimilar instruments to measure the same area of a display can limit the accuracy of this technique.

Similarly, care must be taken to consider temporal effects. Many types of displays are scanned, and the resultant flicker might impart an error if the two instruments are dissimilar. The performance of this technique may also be affected by the environment in which it is performed. Ambient light may affect the measurements if they are not made in a dark room, and changes in ambient temperature may alter the characteristics of one or both of the instruments.

The present work takes advantage of the fact that the spectra of light from displays is not entirely arbitrary. In the future, similar methods may be developed to exploit other situations where the spectra that form colors fall into limited domains. A simple example is well-known: with only two filtered detectors, one "red" and one "blue," the ratio of their signals locates the position of an incandescent source on the black body locus (i.e., determines its radiance temperature). However, such measurements may be improved by using additional filtered detectors and a C matrix designed to make the instrument insensitive to the forms of the likely deviations from truly Planckian spectra. (These deviations might arise from variable obscuration between the source and the instrument, a spectral emissivity that varies due to possible variations in the source material, or other factors that would depend on the application of the instrument.) Using principal-component analysis, Parkkinen, Hallikainen, and Jaaskelainen[13] have shown that the reflectance spectra of 1257 samples in a *Munsell Book of Color* can be well represented by a basis set of only eight functions. Moreover, Jaaskelainen, Parkkinen, and Toyooka[14] went on to show that the same basis applies to a sample set of many natural objects. This suggests that a tristimulus colorimeter may be constructed with $f \approx 8$ and a C matrix that takes advantage of the discrimination that such a basis set offers. In both of these cases, the C matrix would cleverly combine the signals from *all* well-chosen f filtered detectors for *each* determination of X, Y, or Z.

5. EXPERIMENTAL VERIFICATION

The R' transformation was demonstrated by optimizing a commercial, four-channel tristimulus colorimeter, using as a reference a commercial spectroradiometer. Each instrument held its original, factory calibration. Both instruments were connected to electric outlets (as opposed to being operated solely on their internal storage batteries), and both were allowed one hour to reach operating equilibrium.

The display used for this test was a CRT connected to a popular type of microcomputer. The computer was programmed to display a circle 12 cm in diameter near the center of the screen, either in an elementary color (R, G, B full on or full off; or all at 50% for gray) or randomly. In a darkened room, alternately the tristimulus colorimeter was attached to the screen with its suction cup near the center of the test circle, and the spectroradiometer was used from its tripod mount ≈1 m away. The spectroradiometer was aimed at the center of the test circle and well defocused to integrate the light from many pixels. Data were recorded for the eight non-black elementary colors and for 20 randomly generated colors. Particularly dark colors were skipped and not recorded.

Table 1 shows the data of the elementary colors. These data from the reference spectroradiometer and the tristimulus colorimeter being optimized were used to calculate R' (Eq. 18):

$$R' = \begin{pmatrix} 1.0536 & 0.0007 & 0.0088 \\ 0.0144 & 1.0519 & 0.0138 \\ 0.0081 & -0.0080 & 1.0861 \end{pmatrix} \tag{32}$$

Using this R', the last three columns in the Table were calculated. They show that, in this example, the optimization made a substantial improvement in the luminance calibration of the target colorimeter. This is primarily due to the ≈5% scale shift that the optimization includes. The procedure made a modest improvement to the x chromaticity coordinate measurements, and no significant change to the y chromaticity coordinate measurements. This particular tristimulus colorimeter was well adjusted before the optimization.

The reader is reminded that the optimization procedure minimizes differences in X, Y, Z, rather than Y, x, y. In this example, the limiting factor to further improvements in x, y matching appears to be caused by non-ideal behavior in one or the other of the instruments in its Z channel. This can be seen in the residuals (Eq. 15), particularly for the blue elementary color. This instrumental anomaly was reproduced using another computer and monitor.

It is fair to ask whether another R_0 matrix might do a better job at minimizing the RMS differences in x, y, rather than X, Z. This requires non-linear (i.e., iterative) data fitting. The reader may verify that

$$R_0 = \begin{pmatrix} 1.0641 & -0.0030 & 0.0031 \\ 0.0144 & 1.0519 & 0.0138 \\ 0.0064 & -0.0080 & 1.0922 \end{pmatrix} \tag{33}$$

reduces the RMS Δx to 0.0014, and the RMS Δy to 0.0025. For the same Y transformation (Eq. 19), this matrix minimizes the sum-squared distances in x, y space between the reference and transformed data.

The value of the optimization procedure is shown in how the reference and optimized instruments continue to track each other after the R' matrix is determined and fixed. Table 2 shows 20 randomly selected colors, both before and after the R' transformation. The improvements in Y, x, y are similar to those seen with the elementary colors. The R_0 matrix shown in Eq. 33 leads an RMS Δx of 0.0019, and an RMS Δy of 0.0020.

6. CONCLUSION

It makes good sense for users of tristimulus colorimeters to be able to determine and apply R transformations to their data in order to optimize their instruments for use with particular display devices, or types of display devices. The data in Tables 1 and 2 show that the R' transformation, which is motivated by the nature of the measurement process, worked as well in transferring calibration as did an *ad hoc* method, which was based on "brute force" minimization in x, y space. The R' transformation has the advantage of being determined in closed form by the data, instead of requiring an iterative procedure that has a risk of converging to a local but not the global minimum.

Although R transformations have been recognized for over 20 years as being potentially valuable, industry has been slow to adopt them. Clearly the computations described in this paper require the sort of microprocessor control in instruments that was not available 20 years ago. More to the point, few manufacturers have designed their tristimulus colorimeters to allow each detector channel to contribute to the computation of all of X, Y, Z. That is, they do not have the nine degrees of calibration freedom required for a complete R matrix.

True, the user of such a colorimeter can apply an R matrix *ex post facto*. However this may be clumsy to do, and such calculations may suffer from non-linear effects (including the truncation of precision in the data) caused by the instrument. Previously published formulae for R' have not reflected all of the subtlety of the situation.

It is the intent of the author to encourage manufacturers of tristimulus colorimeters to design their products to include means to calculate and apply R' matrices automatically, or to include libraries of R' matrices in their instruments for different reference phosphor sets and other display colorants. This would lead to wider use of the technique, and potentially more uniformly accurate measurements of display color throughout the community of display producers and users. Progress in developing device-independent color standards is well known, but such efforts need to be backed by accurate measurements.

Recognizing this need, the American Society for Testing and Materials (ASTM), through its Committee E-12 on Appearance, adopted the Wagner formulation of R' as a voluntary industry standard in 1992.[15] The author is working with Committee E-12 to refine and improve the currently published standard. Participation of additional producers and users of tristimulus colorimeters in the activities of the committee, including standards development and balloting, would be most welcome.[16]

7. ACKNOWLEDGEMENTS

Special thanks are due to the members of ASTM Subcommittee E-12.06 on the Appearance of Displays who are endeavoring to put this work to practical use. Among many others, we value the contributions of Alan Kravetz (Minolta Corp.), and Cal Leone and Spike Sojourner (Graseby Optronics). The author also gratefully acknowledges the fruitful discussions and collaboration with his colleagues at NIST on this project. Thanks are due, among others, to Ed Kelley in the flat panel display laboratory, and to Yoshi Ohno, Chris Cromer, and George Eppeldauer in the Radiometric Physics Division.

8. REFERENCES

1. http://www.hike.te.chiba-u.ac.jp/ikeda/CIE/home.html
2. *Colorimetry*, 2nd Ed., Publ. 15.2; *Colorimetric Observers*, Publ. S002 (Commission Internationale de L'Éclairage, Vienna, 1986). (CIE publications are currently available through the U. S. National Committee of the CIE, c/o Mr. Thomas M. Lemons, TLA–Lighting Consultants, Inc., 7 Pond Street, Salem, MA 01970-4819.)
3. *Colorimetric Observers*, Publ. S002, is also available in electronic form as disk D001. This most recent revision to the standard contains data at 1 nm intervals. Older data at 5 nm intervals from Publ. 15.2 are available via:
   ftp://ftp.hike.te.chiba-u.ac.jp/cie/README.txt ftp://ftp.hike.te.chiba-u.ac.jp/cie/x2.txt
   ftp://ftp.hike.te.chiba-u.ac.jp/cie/y2.txt ftp://ftp.hike.te.chiba-u.ac.jp/cie/z2.txt
4. See also: R. S. Berns, M. E. Gorzynski, and R. J. Motta, "CRT Colorimetry. Part II: Theory and Practice," *Color Res. Appl.* **18**(5), 315-325 (1993).
5. See also: R. S. Berns, R. J. Motta, and M. E. Gorzynski, "CRT Colorimetry. Part I: Theory and Practice," *Color Res. Appl.* **18**(5), 299-314 (1993).
6. H. G. Wagner, "Farbmessungen an Fernseh-Farbbildröhren nach dem Dreibereichsverfahren," *Farbe 22*, 31–43 (1973).
7. G. Lux and J. Schanda, "On the Vectorial Representation of Basic Colour Perception and its Use in Colour-Measurement," *Reports of the Scientific Staff of the Research Institute for Technical Physics of the Hungarian Academy of Sciences*, 5–27 (1974).
8. G. Eppeldauer, J. Schanda, and G. Lux, "Berendezés szinösszetevök és szinkoordináták meghatározására," ("Instrument to determine CIE tristimulus values and chromaticity coordinates,") Hungarian Patent #168,258, Feb. 28, 1977 (Expired).
9. S. Kokoschka, "Matrixkorrektur fehlangepaßter Dreibereichsfarbmeßgeräten," ("Matrix Correction of Wrongly Adjusted Tristimulus Colorimeters,") *Farbe + Design* **11**, 22–25 (1979).
10. W. Erb, M. Krystek, and W. Budde, "A Method for Improving the Accuracy of Tristimulus Colorimeters," *Color Res. Appl.* **9**, 84–88 (1984).
11. Y. Ohno and R. Suzuki, "Colorimeter for Color Displays," United States Patent #5,363,197, Nov. 8, 1994.
12. P. R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences*, pp. 106–108, McGraw Hill, New York, 1969.
13. J. P. S. Parkkinen, J. Hallikainen, and T. Jaaskelainen, "Characteristic spectra of Munsell colors," *J. Opt. Soc. Am. A* **6**(2), 318–322 (1989).
14. T. Jaaskelainen, J. Parkkinen, and S. Toyooka, "Vector-subspace model for color representation," *J. Opt. Soc. Am. A* **7**(4), 725–730 (1990).
15. E 1455–92, "Standard Practice for Obtaining Colorimetric Data from a Visual Display Unit Using Tristimulus Colorimeters," *ASTM Standards on Color and Appearance Measurement*, 4th Ed. (Available from ASTM, 100 Barr Harbor Dr., West Conshohocken, PA 19428-2959.)
16. Apply through the ASTM membership office, at the above address.

Table 1. Luminance and chromaticity coordinates of the elementary colors as measured by the reference colorimeter and the colorimeter being optimized (the target colorimeter). The data from the target colorimeter is shown both before and after the optimization matrix is applied. *Y* stated in candela/square-meter.

Elementary Color	Reference			Target			Optimized Target		
	Y	*x*	*y*	*Y*	*x*	*y*	*Y*	*x*	*y*
Red	12.25	0.617	0.351	11.50	0.620	0.350	12.40	0.613	0.354
Green	38.45	0.294	0.604	36.25	0.293	0.605	38.47	0.292	0.605
Yellow	50.85	0.409	0.516	47.53	0.413	0.514	50.64	0.411	0.515
Blue	6.43	0.150	0.075	5.11	0.153	0.070	6.32	0.153	0.079
Magenta	18.65	0.286	0.154	16.43	0.293	0.153	18.55	0.288	0.159
Cyan	44.70	0.211	0.301	41.25	0.213	0.299	44.74	0.211	0.300
White	56.75	0.289	0.311	52.55	0.293	0.309	56.92	0.289	0.310
Gray	11.90	0.297	0.323	10.95	0.302	0.320	11.84	0.298	0.321
RMS difference from reference:				2.587	0.0040	0.0025	0.125	0.0021	0.0026

Table 2. Luminance and chromaticity coordinates of randomly generated colors as measured by the reference colorimeter and the target colorimeter. The data from the target colorimeter is shown both before and after the optimization matrix is applied. *Y* stated in candela/square-meter.

Reference			Target			Optimized Target		
Y	*x*	*y*	*Y*	*x*	*y*	*Y*	*x*	*y*
36.10	0.236	0.258	33.00	0.239	0.256	36.05	0.236	0.258
17.70	0.551	0.404	16.55	0.556	0.401	17.76	0.551	0.404
5.82	0.314	0.573	5.49	0.315	0.576	5.83	0.314	0.576
28.50	0.256	0.216	25.85	0.262	0.215	28.51	0.258	0.218
16.90	0.331	0.186	15.10	0.339	0.185	16.81	0.333	0.190
12.80	0.608	0.357	11.95	0.613	0.356	12.88	0.606	0.360
12.30	0.290	0.236	11.10	0.298	0.234	12.18	0.294	0.237
40.20	0.318	0.585	37.70	0.318	0.585	40.04	0.317	0.586
5.59	0.295	0.587	5.27	0.294	0.591	5.60	0.293	0.591
10.10	0.481	0.454	9.35	0.488	0.450	10.00	0.484	0.452
16.90	0.298	0.429	15.70	0.301	0.427	16.81	0.298	0.427
31.10	0.232	0.305	28.65	0.234	0.305	31.05	0.232	0.306
8.56	0.238	0.127	7.41	0.244	0.124	8.52	0.240	0.131
44.90	0.293	0.425	41.85	0.295	0.425	44.82	0.293	0.425
18.30	0.193	0.234	16.60	0.195	0.231	18.23	0.194	0.233
46.10	0.379	0.534	43.25	0.378	0.535	46.03	0.376	0.536
37.20	0.443	0.490	34.95	0.443	0.490	37.28	0.440	0.492
26.90	0.329	0.575	25.40	0.330	0.575	26.98	0.329	0.576
26.00	0.265	0.498	24.40	0.266	0.498	26.01	0.265	0.497
7.41	0.395	0.488	6.97	0.401	0.488	7.44	0.398	0.489
RMS difference from reference:			1.865	0.0043	0.0020	0.076	0.0019	0.0021