Linear models of surface and illuminant spectra

David H. Marimont and Brian A. Wandell

Xerox Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, California 94304

Received February 11, 1992; accepted April 15, 1992; revised manuscript received June 11, 1992

We describe procedures for creating efficient spectral representations for color. The representations generalize conventional tristimulus representations, which are based on the peripheral encoding by the human eye. We use low-dimensional linear models to approximate the spectral properties of surfaces and illuminants with respect to a collection of sensing devices. We choose the linear-model basis functions by minimizing the error in approximating sensor responses for collections of surfaces and illuminants. These linear models offer some conceptual simplifications for applications such as printer calibration; they also perform substantially better than principal-components approximations for computer-graphics applications.

1. INTRODUCTION

Current dogma in color science emphasizes the distinction between the physical variables of image formation and the perceptual variables of color appearance. Color is a psychological phenomenon, of course; but we must not forget that color also serves to estimate the physical factors of image formation. In this and related papers we propose color representations and techniques for computing with them that incorporate the physical factors of image formation accurately and naturally. Incorporating physical variables explicitly in color representations leads to more realistic imagery in computer graphics and provides a sound basis for inferring the physical variables as part of conventional colorimetry.

The high dimensionality of surface and illuminant spectral functions poses a challenge to their inclusion in color image representations. In recent years there has been interest in finding efficient, low-dimensional linear representations of surface reflectance and illuminant spectral-power distribution functions. Efficiency is essential if we are to succeed in creating useful spectral representations of color information. Efficient linear representations have potential applications for rendering in computer graphics and for material estimation in computer vision.

Linear models have two useful properties. First, linear models offer a compact description of the data. For example, Parkkinen *et al.*[1] recently measured more than 1200 Munsell chips at a 5-nm sampling interval over the wavelength range from 400 to 700 nm (61 numbers/samples). Parkkinen *et al.* found that the reflectance data can be represented with no loss of precision with the use of a small number of basis functions (8 numbers/samples). Their measurements confirmed earlier studies of surfaces by Cohen[2] and Maloney.[3]

But efficiency is only part of the motivation for using linear models; many other compression schemes would do just as well or better. A second important reason is that linear models preserve the simplicity of graphics and estimation algorithms. When sensor encoding is linear with incident light (as in the human photopigments or in CCD sensors), linear models fit well into the computational algorithms for material and illuminant estimation algorithms and for computer-graphics calculations.

2. THE MAIN IDEA

Classically, linear models are built by approximating the spectral functions in the wavelength domain (see, e.g., Refs. 2 and 4). For example, suppose that we build a d-dimensional linear model to approximate the surfaces in a collection. The linear model will consist of a set of d-basis functions; we approximate each surface-reflectance function, $S(\lambda)$, in the collection as the weighted sum of d-basis functions:

$$S(\lambda) \approx \sum_{i=1}^{i=d} \sigma_i S_i(\lambda). \tag{1}$$

The basis functions, S_i, are chosen to minimize the error

$$\sum_S \int \left[S(\lambda) - \sum_{i=1}^{i=d} \sigma_i S_i(\lambda) \right]^2 d\lambda. \tag{2}$$

As the dimension d of the linear model increases, the approximation improves. The basis functions that minimize the quantity in expression (2) can be found by using many standard techniques, all of which can be derived from the singular-value decomposition of the matrix whose columns contain the surface-reflectance data.

For many applications, however, a linear model designed to minimize the error in expression (2) is inappropriate. For example, suppose that we want to represent the spectral-reflectance functions of print samples in order to predict the response of a flatbed scanner. Scanner sensors do not respond equally well to all wavelengths; the conventional minimization based on expression (2) is ill suited for predicting the scanner responses. If we want a linear model that helps us to predict the scanner response, then we should derive our basis functions by minimizing the error in the predicted scanner responses.

As a second example, suppose that we design linear models to represent surface and illuminant spectral functions in computer-graphics simulations. The graphics-simulation objective is to predict the initial human encoding (e.g., tristimulus coordinates) expected from various surface–illuminant combinations. Linear models for surface and illuminant functions derived by minimizing expression (2) do not perform so well as models derived by minimizing errors of the tristimulus encoding.

0740-3232/92/111905-09$05.00

In the computer-graphics example, there is an interdependency between the surface and illuminant collections, as well. If the collection of illuminants used in the simulations has no energy in some spectral range, resources devoted to representing the surface-reflectance functions in that range are wasted. Accurate representations of the surface where there is no illuminant energy reduce the error in expression (2), but they do not improve the quality of the graphics simulation. In some cases, then, we wish to define linear models for surface and illuminant collections simultaneously.

In this paper we describe how to build linear surface and illuminant models that simultaneously take into account the properties of the surface collection, the illuminant collection, and the sensor responsivities. In Section 3 we describe the general principles of our analysis. In Section 4 we show how to derive linear models for a collection of surfaces. In Section 6 we show how to derive linear models simultaneously for collections of surfaces and illuminants. The techniques that we introduce are adapted from numerical methods used in the statistical literature, where they are referred to as one-mode, two-mode, or n-mode analyses;[5-8] we retain the terminology here.

3. BACKGROUND

We use matrix algebra to describe the relationships among surface-reflectance functions, illuminant spectral-power distributions, and sensor responses. The matrix products relating these quantities are illustrated in tableau form in Fig. 1. In our calculations we represent functions of wavelength at $N_w = 31$ sample points, ranging from 400 to 700 nm in 10-nm steps. The formulas that we use apply to materials without phosphorescence or fluorescence.

The entries of the surface-reflectance-function vector, $\mathbf{s}$, are the reflectance values at the N_w sample wavelengths. We assume that the geometric properties of the image (the angle with respect to the illuminant, specularity, etc.) are incorporated within the spectral-reflectance function. We represent the illuminant by a diagonal matrix, $\mathbf{E}$, whose entries contain the illuminant's spectral-power distribution at the sample wavelengths. The sensor responsivities at the sample wavelengths, $\mathbf{X}_i(\lambda)$, are defined by the three columns of the matrix $\mathbf{X}$.

We compute the sensor responses $\mathbf{r}$ from the matrix product $\mathbf{r} = \mathbf{X}^t\mathbf{E}\mathbf{s}$. Generally we compute the sensor responses for many surfaces using a single illuminant. It is convenient to define the system's surface transfer matrix, $\mathbf{T}_E = \mathbf{X}^t\mathbf{E}$:

$$\mathbf{r} = \mathbf{T}_E\mathbf{s}. \tag{3}$$

The sensor responses are the projections of the surface-reflectance function onto the rows of the surface-transfer matrix, $\mathbf{T}_E$. From standard theorems in linear algebra, when the dimension of $\mathbf{r}$ is smaller than the dimension of $\mathbf{s}$ we can express $\mathbf{s}$ in two orthogonal parts, $\mathbf{s} = \hat{\mathbf{s}} + \mathbf{s}^\perp$. The vector $\mathbf{s}^\perp$ is orthogonal to the rows of $\mathbf{T}_E$:

$$0 = \mathbf{T}_E\mathbf{s}^\perp. \tag{4}$$

The vector $\hat{\mathbf{s}}$ falls within the row space of the transfer matrix. We can write $\hat{\mathbf{s}}$ as the weighted sum of the rows of $\mathbf{T}_E$. We denote the vector of weights as $\mathbf{w}$:

$$\hat{\mathbf{s}} = \mathbf{T}_E{}^t\mathbf{w}. \tag{5}$$

The three-dimensional vector $\mathbf{w}$ is an efficient description of the part of the surface-reflectance function that falls within the linear subspace defined by the columns of $\mathbf{T}_E{}^t$. This is the only part of the surface-reflectance function that influences the device response. For example, when a human observer views a surface under a D_{65} illuminant, the photopigment absorption rates measure the projection of the surface-reflectance function onto the three-dimensional subspace defined by the XYZ_{D65} functions. If we know the tristimulus values associated with the surface, we can use the pseudoinverse to recover this component of the surface-reflectance function exactly[9]:

$$\begin{aligned}\mathbf{r} &= \mathbf{T}_E\mathbf{s},\\ \hat{\mathbf{s}} &= \mathbf{T}_E{}^t(\mathbf{T}_E\mathbf{T}_E{}^t)^{-1}\mathbf{r}.\end{aligned} \tag{6}$$

The sensor data from a single illuminant tell us nothing about the orthogonal component, $\mathbf{s}^\perp$. But we may have *a priori* information about the distribution of surface-reflectance functions, or we may know more by having had experience with the object under several illuminants.[10,11]

The sensor responses inform us only about the value of the surface-reflectance function that is within the linear subspace defined by the rows of the surface-transfer matrix. It follows that, to predict the sensor responses, a linear model must represent only the portion of the surface-reflectance function that is in the row space of the transfer matrix.

One way to see this result is to recall that the surfaces $\mathbf{s}$ and $\hat{\mathbf{s}}$ differ only by a term that is orthogonal to the row space. From Eqs. (4) and (5) we see that pairs of surfaces that differ only by this orthogonal term are surface metamers with respect to a device with surface-transfer

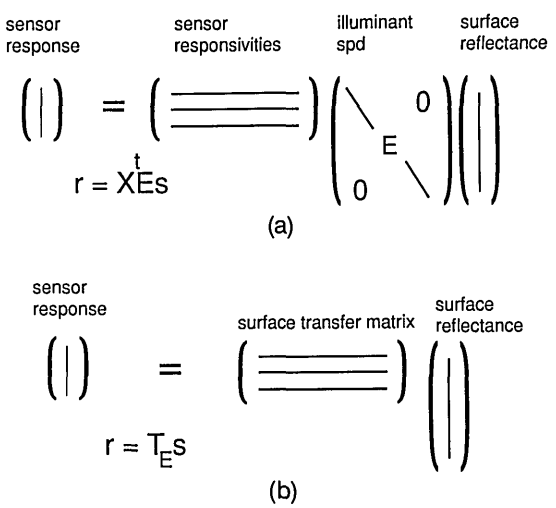

Fig. 1. Device sensor responses. We group the surface-reflectance vectors $\mathbf{s}$ in the columns of a matrix $\mathbf{S}$. Similarly, we group the sensor-response vectors $\mathbf{r}$ in the columns of a matrix $\mathbf{R}$. (a) The sensor responses are determined by the product of the surface-reflectance function (columns of the rightmost matrix) and a diagonal matrix containing the illuminant spectral power distribution and a matrix whose rows contain the sensor responsivities. (b) We group the sensor matrix and the illuminant matrix to define a surface-transfer matrix.

function $\mathbf{T_E}$:

$$\mathbf{T_E s} = \mathbf{r} = \mathbf{T_E}(\hat{\mathbf{s}} + \mathbf{s}^{\perp}) = \mathbf{T_E}\hat{\mathbf{s}}. \quad (7)$$

Only the portion of the surface reflectance in the linear subspace influences the sensor response:

$$\mathbf{r} = \mathbf{T_E}\hat{\mathbf{s}} = (\mathbf{T_E}\mathbf{T_E}^t)\mathbf{w}. \quad (8)$$

The methods in this paper are elaborations of the observation that the sensor responses depend only on the low-dimensional vector, $\mathbf{w}$ [Eq. (8)]. Hence, to predict the sensor responses for a device with surface-transfer function $\mathbf{T_E}$, we can use any linear model that spans the same subspace as the rows of $\mathbf{T_E}$.

Several investigators have used related methods. Takahama and Nayatani[12] used the pseudoinverse to discover metamers. Cohen and Kappauf developed a formal argument that comes to the same conclusion as the previous few paragraphs. Their analysis treats the system input as light at the cornea rather than the surface reflectance. They call the value returned by the pseudoinverse the fundamental metamer.[13–15] Trussell[16] reviews the use of linear algebraic methods in a variety of color-systems applications.

4. SPECTRAL REPRESENTATIONS FOR SURFACES

A. Introduction

Many applications require that we represent the surface-reflectance functions with respect to several different transfer functions. For example, print samples are measured by scanners with different transfer functions; computer-graphics applications render the same surfaces under various illuminants. In this section we extend the basic logic developed in Section 3 to the present application.

Suppose that we want to build a d-dimensional linear model of the N_s surface reflectances in a collection of print samples. In this section we use an example of finding a representation of the print samples that permits us to predict a flatbed-scanner response and the human tristimulus values XYZ_{D65}.

Call the surface-transfer matrix of the scanner $\mathbf{T_E}$ and the surface-transfer matrix for the human eye under illuminant D_{65} $\mathbf{T_H}$. From the discussion in Section 3, we know that only the portion of the reflectance functions that is in the row space of $\mathbf{T_E}$ influences the scanner response. Similarly, the row space of $\mathbf{T_H}$ defines the relevant portion of the reflectance in determining the tristimulus values. It follows that a linear model to predict the responses of both scanners exactly requires, at most, six dimensions to span the rows of the two transfer matrices. See Takahama and Nayatani[12] and Burns *et al.*[17] for related analyses.

It may be possible, however, to use fewer than six dimensions. For example, if the scanner has the same responsivities as the human eye under illuminant D_{65}, we do not need to increase the dimension of the linear model at all. Somewhat more generally, if the scanner responsivities are within a linear transformation of XYZ_{D65}, we need not increase the dimension of the linear model. Finally, if the scanner responsivities are nearly within a linear transformation of one another, there may be little advantage in increasing the dimension of the linear model.

B. Linear Models as Projections

Selecting a linear model for surface reflectances defines a map from an arbitrary surface reflectance to an approximation; the approximation must fall within a linear subspace. One way to conceive of the construction of a linear model is as follows.

We define a d-dimensional model [see expression (1)] by selecting a set of basis functions. Place the basis functions in the columns of a matrix, $\mathbf{L_b}$. Given a spectral reflectance $\mathbf{s}$, we choose the linear model weights by minimizing the least-squared error in

$$\mathbf{s} \approx \mathbf{L_b w}. \quad (9)$$

The vector $\mathbf{w}$ that minimizes the vector length $\|\mathbf{s} - \mathbf{L_b w}\|$ is

$$\mathbf{w} = \mathbf{L_b}^{+}\mathbf{s}, \quad (10)$$

where $\mathbf{L_b}^{+}$ is the pseudoinverse of $\mathbf{L_b}$. We introduce the notation $\mathbf{L_s} = \mathbf{L_b}^{+}$, and we call $\mathbf{L_s}$ the sampling functions of the linear model. Figure 2 shows the mapping from the original surface-reflectance function to its approximation, $\hat{\mathbf{s}} = \mathbf{L_b L_s s}$, in a matrix tableau.

Notice that only the subspace spanned by the basis functions, and not the basis functions themselves, determines the precision of the approximation. Suppose that $\mathbf{A}$ is a $d \times d$ invertible linear transformation. Then the sampling and basis functions $\mathbf{AL_s}$ and $\mathbf{L_b A^{-1}}$ define the same linear model as $\mathbf{L_s}$ and $\mathbf{L_b}$. Since the choice of sampling and basis functions is arbitrary up to a linear transformation, we can select an orthonormal set of basis functions, $\mathbf{L_b}$. By choosing the basis functions to be orthonormal, we obtain so-called self-inverting sampling and basis functions, i.e., $\mathbf{L_b}^t = \mathbf{L_s}$ (see, e.g., Ref. 18). The search for self-inverting sampling and basis functions has played an important role in the work of some of our colleagues in biological spatial vision; they argue that self-inversion makes it easier to interpret the meaning of the weights and to relate theory to the performance of retinal neurons.

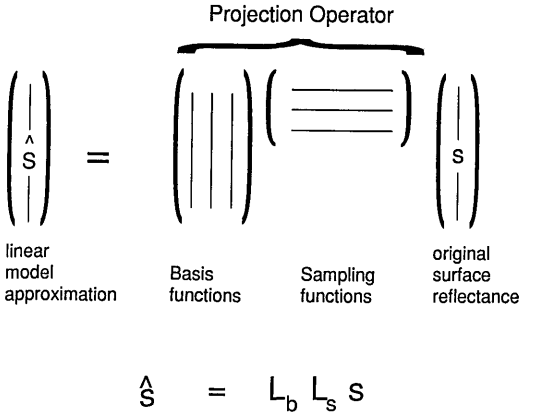

Fig. 2. Linear models as projection operators. These models define a mapping from the original surface reflectance to an approximation that falls in a subspace. We can conceive of the projection as a linear sampling, $\mathbf{L_s}$, followed by a basis reconstruction, $\mathbf{L_b}$. The product, $\mathbf{P_D} = \mathbf{L_b L_s}$, is a projection, i.e., $\mathbf{P_D} = \mathbf{P_D}^2$.

Finally, notice that the mapping from the original surface-reflectance vector to its approximation, $\mathbf{P_D} = \mathbf{L_b L_s}$, is a projection, $\mathbf{P_D} = \mathbf{P_D}^2$.

C. Selecting the Minimization Equation

If we replace the continuous functions of wavelength with sampled functions stored as vectors, then minimizing expression 2 is equivalent to selecting a projection operator to minimize

$$E_{pc} = \|\mathbf{S} - \mathbf{P_D S}\|, \tag{11}$$

where $\mathbf{S}$ is the matrix whose columns contain the surface-reflectance functions in our sample set and the norm operator, $\| \ \|$, is the sum of the squared entries of the matrix.

In the applications that we are considering, however, we seek to minimize a different quantity: the error in predicting the sensor responses:

$$E_{om} = \|\mathbf{TS} - \mathbf{TP_D S}\|. \tag{12}$$

The matrix $\mathbf{T}$ in Eq. (12) includes the sensor responsivities of all the input devices. For example, if we are designing a linear model with respect to two color devices, then $\mathbf{T}$ consists of six rows consisting of all the sensor responsivities. If the linear model is designed with respect to two color devices and one monochrome device, then $\mathbf{T}$ has seven rows, as is illustrated in Fig. 3.

The linear models that minimize the quantities in Eqs. (11) and (12) can be quite different. But the numerical procedures for deriving the d-dimensional projection to minimize either quantity are the same. Equation 11 is ordinarily solved by using some variant of the singular-value decomposition, such as principal components or the eigenvectors of the covariance matrix. We can use the singular-value decomposition to minimize the quantity in Eq. (12), as well. To see this, notice that we can re-express Eq. (12) as

$$\mathbf{TS} \approx \mathbf{T\hat{S}} = (\mathbf{TL_b})(\mathbf{L_s S}), \tag{13}$$

where $\hat{\mathbf{S}}$ is the matrix containing all the linear model approximations $\hat{\mathbf{s}}$ in its columns.

Expression (13) shows that we are seeking to approximate the sensor data $\mathbf{R} = \mathbf{TS}$ by the product of two rectangular matrices. From standard theorems we know that we can use the singular-value decomposition, applied to the sensor responses, to obtain the best least-squared error approximation. We factor the sensor responses, $\mathbf{R}$, into three matrices, $\mathbf{R} = \mathbf{UDV}^t$. Suppose that the matrix $\mathbf{R}$ is $r \times c$ and that $m = min(r, c)$. Then $\mathbf{D}$ is square $(m \times m)$ and diagonal with entries d_i such that $d_i \geq d_{i+1}$. The matrices $\mathbf{U}$ and $\mathbf{V}$ are $r \times m$ and $c \times m$, and their columns are orthonormal. We obtain a d-dimensional approximation of $\mathbf{R}$ by zeroing all but the first d-diagonal values to create $\mathbf{D}_d$. This yields the approximation

$$\mathbf{R} \approx (\mathbf{U}_d \ \ 0)\begin{bmatrix}\mathbf{D}_d & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\mathbf{V}_d^t\\ 0\end{bmatrix}$$

$$= (\mathbf{U}_d\mathbf{D}_d)(\mathbf{V}_d^t). \tag{14}$$

Once we zero these entries, only the first d columns of $\mathbf{U}$ and $\mathbf{V}$ are relevant. So we can approximate the data by $\mathbf{U}_d\mathbf{D}_d\mathbf{V}_d^t$ (e.g., Refs. 9 and 19).

We set $\mathbf{U}_d\mathbf{D}_d$ equal to $\mathbf{TL}_b$, and $\mathbf{V}_d^t$ equal to $\mathbf{L_s S}$. Thus the first d rows of $\mathbf{V}^t$ define the weights of the surfaces in the linear model. We call this rectangular matrix $\mathbf{W}$.

Finally, we recover the matrices $\mathbf{L_s}$ and $\mathbf{L_b}$ as follows. Knowing the surface reflectances $\mathbf{S}$ and the weights $\mathbf{W}$, we can solve for the sampling functions from

$$\mathbf{L_S S} = \mathbf{W}. \tag{15}$$

D. Example Calculation

We have created linear models for the surface-reflectance functions of the Macbeth color checker.[20] We created the one-mode linear model with respect to a commercially available scanner and a human observing the surfaces under a diffuse D_{65} illuminant. The principal-components model does not depend on specifying device sensors.

The Macbeth color checker consists of 24 uniform patches separated by a black border. The scanner measurements consist of three color-sensor responses for the 24 patches, ranging from 0 to 255. Within each color patch there is a distribution of scanner responses, owing to nonuniformities in the sample and sensors as well as to noise. The standard deviations within a single Macbeth color-checker patch range from approximately 0.5 for the darker surfaces to 2.0 for the lighter surfaces.

We have computed the XYZ_{D65} values for these color patches and scaled them to fall within the same range. Because these data are synthesized, the mapping from surface-reflectance function of the Macbeth surface to the scanner responses is not precisely linear; we will quantify the departures from linearity after describing our analysis in more detail below.

The scanner spectral responsivity differs from the XYZ_{D65} functions. We can infer this from the fact that the best linear regression between the scanner rgb values to the XYZ_{D65} values differs by a root-mean-squared error of approximately (7.84, 5.84, 3.43) units for the r, g, and b values, respectively. This error is considerably larger than the measurement error of between 0.5 and 2.0 units.

To create the principal components, we build the matrix $\mathbf{S}$, whose columns contain the surface-reflectance functions of the color checker. The principal components of $\mathbf{S}$ provide the best approximation to the surface-reflectance functions relative to the minimization in Eq. (11). The first three principal components of the Macbeth color checker are plotted in Fig. 4(a).

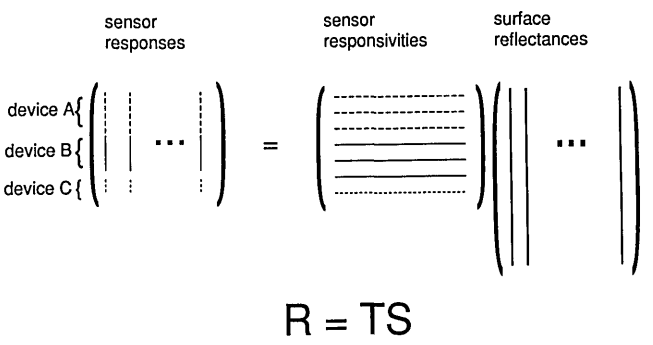

Fig. 3. Pooled sensor responses grouped across devices and surfaces. We group the sensor responsivities into a single matrix, $\mathbf{T}$. We group the surface-reflectance functions in the collection into the columns of a matrix, $\mathbf{S}$. The sensor responses are equal to $\mathbf{TS}$. We analyze the sensor responses to derive the one-mode representation of the surface-reflectance functions.

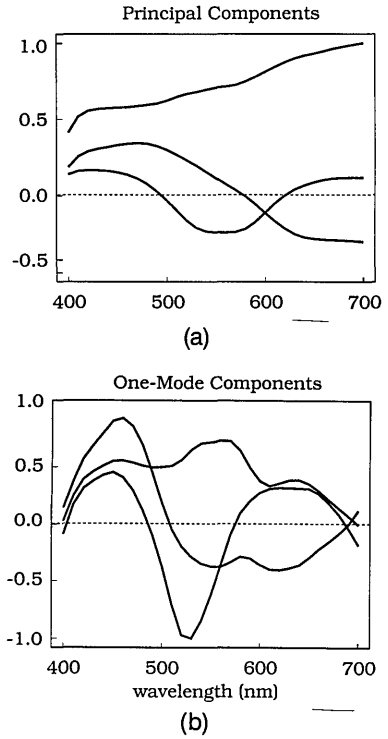

Fig. 4. Comparison of the sampling functions for (a) the principal-components linear model and (b) the one-mode linear model that were used to describe the Macbeth color-checker surfaces. The principal-components representation is independent of the sensors. The one-mode representation is chosen with respect to the sensors described in the text.

To perform the one-mode analysis we create the 6×24 matrix $\mathbf{R}$, containing the scanner responses to the 24 surfaces and the human visual system XYZ_{D65} values. We calculate the one-mode linear model weights from the singular-value decomposition of $\mathbf{R}$, as described above.

We derive the linear model sample functions $\mathbf{L}_s$ by using the known surface-reflectance functions. We have measured the surface-reflectance functions of our Macbeth color checker, using an instrument that is accurate to ~1%. We can approximate the measurements to within the precision of our instrument by factoring the matrix $\mathbf{S}$, using the singular-value decomposition and replacing the diagonal elements beyond the eighth with zero. This creates an approximation to the data, $\hat{\mathbf{S}} \approx \mathbf{S}$. Hence, then, to within the precision of our measurements there are eight independent surfaces in the Macbeth color checker. Following the common practice to reduce the effects of noise, we use $\hat{\mathbf{S}}$ in the matrix-inversion steps below. We plot the one-mode sampling functions in Fig. 4(b).

Notice that the principal-components linear model has large values at the spectral extremes even though the scanner and the eye are insensitive at these wavelengths. The one-mode basis functions allocate their variance in the visible part of the spectrum.

Figure 5 compares how well the two linear models predict the device responses. We used the linear models with dimensions $d = 2, \ldots, 6$, and we calculated the best linear regression between the linear-model surface weights and the observed scanner data. As the linear-model dimension increases, the quality of the fit improves, converging to the best value at a dimension of six (the number of sensors).

Were the sensor responses of both devices linear, a six-dimensional linear model could predict the data perfectly. For the six-dimensional one-mode representation, all the error is due to the scanner nonlinearity. The error in the principal-components representation exceeds the one-mode error in every comparison. Moreover, for the six-dimensional model, the one-mode model predicts the XYZ values accurately, while the principal-components representation is worse than the four-dimensional one-mode model.

The reason that the principal-components model fares so much worse than the one-mode model at low dimensions is that the scanner is quite sensitive to the third and fourth principal components. We can demonstrate this as follows. We have estimated the scanner surface-transfer function by solving the equation $\mathbf{R} = \mathbf{T}_E\hat{\mathbf{S}}$; now we can use $\mathbf{T}_E$ to predict the vector length of the scanner rgb responses to the principal components of the surface reflectances. The vector length of the scanner response to the third principal component exceeds the response to the first. Although the third principal component does not play a significant role in minimizing the quantity in expression (2), it does play a significant role in minimizing the quantity in Eq. (12) (see also Ref. 21).

These calculations illustrate some of the trade-offs involved in using the one-mode representation. The system designer may feel that accurate representation of the scanner responses is much less important than accurate representation of the human observer under XYZ_{D65}. In that case the designer may wish to use the linear model defined entirely by the rows of $\mathbf{T}_H$. In the selection of the XYZ_{D65} functions as the linear model, the best global linear transformation to predict the scanner responses has an rms error of 6.00, worse than either the principal-components or the one-mode representations. The one-mode representation balances the errors more nearly.

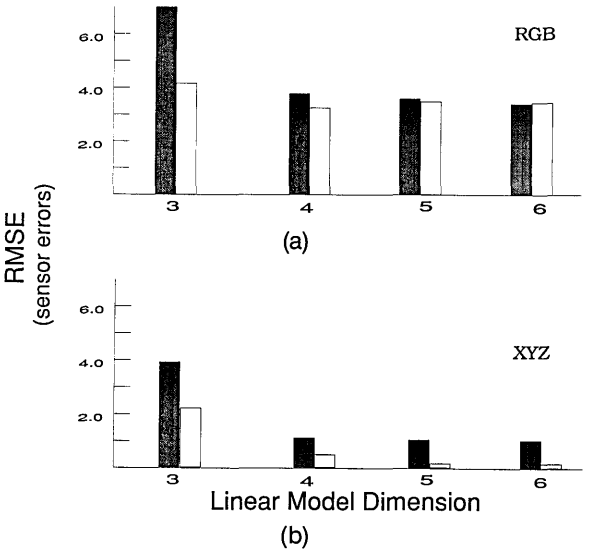

Fig. 5. Root mean-squared error when the sensor responses are predicted by using the principal-components model (dark bars) and the one-mode model (light bars) for different linear model dimensions: (a) errors for the scanner data, (b) errors for the XYZ values.

The designer may compromise between these extremes by adapting the one-mode method to use conventional weighted least-squares fitting procedures to emphasize one device or another.

We can generalize our procedure by using other error measures. For example, the designer may wish to minimize the scanner responses with respect to mean-squared error but to minimize the visual responses with respect to a CIE metric. Such minimizations are possible with the use of iterative search procedures.

5. INTERMEDIATE DISCUSSION

A. Geometric Comparison of One-Mode and Principal Components

The one-mode linear model approximates sensor responses better than the principal-components approximation because the one-mode model is designed to minimize the error in sensor responses. Equations (11) and (12) provide an algebraic comparison of the one-mode and the principal-components minimizations.

Figure 6 illustrates the difference between the two linear-model approximations geometrically. We represent surface vectors by their end points in the plane. We represent the sensor vector as a line in the same plane. We can calculate the sensor response to a surface by drawing the perpendicular between the surface vector end point and the sensor line. The distance from the origin to the point where the perpendicular intersects the sensor line is the size of the sensor response to the surface.

The first principal component is the vector with the smallest average distance from all the surface-reflectance end points; it will pass through the data cloud. Suppose that we approximate a surface, $\mathbf{s}$, by its projection onto the principal-component vector. As Fig. 6 shows, the principal-component approximation does not have the same sensor response as $\mathbf{s}$. The principal-components vector is chosen without reference to the sensor vector. This is the source of error when the principal-components linear model is used to predict the sensor responses.

When there is only a single sensor, the one-mode sampling function is the sensor line. The one-mode approximation is $\hat{\mathbf{s}}$, which falls along the sensor line. The vector $\mathbf{s}^{\perp}$, which is perpendicular to the sensor line, joins the end points of $\hat{\mathbf{s}}$ and $\mathbf{s}$. As Fig. 6 shows, when there is only a single sensor a one-dimensional one-mode linear model predicts the sensor response without error. When there are multiple sensors the one-mode analysis finds a linear-model vector that compromises between the best $\hat{\mathbf{s}}$ associated with all the surface and sensor combinations.

B. Related Work

We draw the reader's attention to a few papers that are closely related to our analysis. First, Drew and Funt[22] perform an analysis that complements ours. They use sensor responses to obtain least-squares estimates of the surface-reflectance function, with the error measured in the wavelength domain. They describe how to use the sensor responses to measure the portion of the surface-reflectance function falling within the span of the first three principal components. Vrhel and Trussell[23] use linear models of reflectance functions to correct for illumination changes. Brainard *et al.*[24] and Maloney[21] analyze the design of sensor responsivities to reduce the effect of illuminant changes.

6. SPECTRAL REPRESENTATIONS FOR SURFACES AND ILLUMINANTS

A. Introduction

In Section 4 we built spectral representations of the surface reflectances by using the one-mode method; we did not model the other spectral components of the system. In this section we consider how to obtain additional efficiencies by modeling other spectral factors.

For example, consider the problem of rendering a collection of surfaces under a collection of illuminants. To render a surface under an illuminant requires calculating the tristimulus values of each surface–illuminant pair. We can make the calculations more efficient by representing the illuminants with respect to a low-dimensional linear model, just as we did for surfaces. In this section we describe a method for simultaneously estimating linear models for the illuminant and surface terms. The method is adapted from Magnus and Neudecker,[6] who call it two-mode analysis.

We calculate the tristimulus values of a simulated surface by multiplying a surface-reflectance function by a matrix that defines the surface-transfer function (see Fig. 1). In computer-graphics applications the rows of the surface-transfer function are the product of the sensor responsivities and the illuminant spectral-power distribution. If the ith illuminant spectral-power distribution in the illuminant collection is $E_i(\lambda)$, then the first row of the surface-transfer matrix is $\bar{x}(\lambda)E_i(\lambda)$, the second is $\bar{y}(\lambda)E_i(\lambda)$, and the third is $\bar{z}(\lambda)E_i(\lambda)$. The surface-transfer matrix for illuminant E_i is then called $\mathbf{T}_{\mathbf{E}_i}$.

As we reviewed in Section 3, when there is only a single surface-transfer matrix, $\mathbf{T}_{\mathbf{E}_1}$, the rows of the matrix serve as an exact linear model for the surfaces. If the graphics application uses a collection of illuminants, we can stack

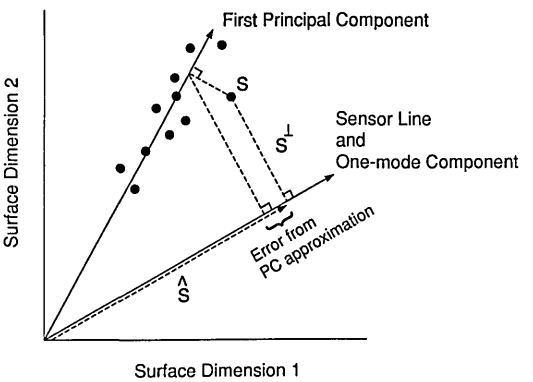

Fig. 6. Geometric interpretation of the one-mode linear model and the principal components. The surfaces in the collection are indicated by their end points; the sensor vector is indicated by a line. The sensor response to a surface, $\mathbf{s}$, is found by dropping a perpendicular from the surface vector end point to the sensor line. The first principal component passes through the data points, minimizing the distance between the vector and the data. The principal-component approximation introduces error in the prediction of the sensor response. The one-mode component is the same as the sensor line. In this example, the one-mode approximation is $\hat{\mathbf{s}}$, a vector on the sensor line. The vector $\mathbf{s}^{\perp}$, which is invisible to the sensor, is perpendicular to the sensor line and is shown added to the vector $\hat{\mathbf{s}}$.

the rows of all the surface-transfer matrices, $\mathbf{T}_{\mathbf{E}_i}$, into a single large matrix, $\mathbf{T}_\mathbf{E}$. We can calculate the tristimulus responses to all the surfaces under all the illuminants from the matrix product:

$$\mathbf{R}_s = \mathbf{T}_\mathbf{E}\mathbf{S}. \tag{16}$$

When we organize the sensor data into the matrix $\mathbf{R}_s$, we can derive a spectral representation for the surface-reflectance functions. We call this organization of the data surface format.

When the data are in surface format, each column contains the sensor responses to a single surface viewed under all the different illuminants. To find a linear model for the illuminants, we reorganize the sensor-data matrix, reversing the roles of the illuminant and surface functions. We transform the data matrix so that each column represents the sensor responses to a single illuminant, reflected from all the different surfaces. To reorganize the data, we perform an operation much like ordinary matrix transposition, except that we transpose the vector of tristimulus values (Fig. 7). The vector transposition operation yields a new data matrix into illuminant format, which we call $\mathbf{R}_e$. We can estimate a spectral representation of the illuminants by using one-mode analysis on the data matrix in illuminant format.

B. Algorithm Definition

To build surface and illuminant linear models simultaneously, we use an iterative algorithm. The algorithm uses the one-mode analysis, alternating between analyzing the data in illuminant and in surface formats. We can select the dimensions of the one-mode approximation for the surface and the illuminant functions independently.

To describe the algorithm, we need a few extra symbols. We use $\mathcal{M}$ to describe the one-mode calculation; we use $\mathcal{V}$ to describe vector transposition (see Fig. 7). We denote the surface weights and the illuminant weights at the ith step of the iterative algorithm as $\mathbf{S}_i$ and $\mathbf{E}_i$, respectively.

First, initialize the estimates for both surface and illuminant weights:

1. $\mathbf{S}_0 = \mathcal{M}(\mathbf{R}_s)$.
2. $\mathbf{E}_0 = \mathcal{M}[\mathcal{V}(\mathbf{R}_\mathbf{S}\mathbf{S}_0^t)]$.

Having obtained initial estimates of the surface and the illuminant, enter the main iteration loop:

1. $\mathbf{S}_i = \mathcal{M}[\mathcal{V}(\mathbf{R}_e\mathbf{E}_{i-1}^t)]$,
2. $\mathbf{E}_i = \mathcal{M}(\mathcal{V}(\mathbf{R}_s\mathbf{S}_i^t)]$,
3. The R^2 value between the observed and the approximated sensor values is guaranteed to be monotonic and nondecreasing as the algorithm iterates.[7] We terminate the iteration loop when the R^2 value increases by less than a criterion amount. Otherwise, we continue the iteration.

Trussell[16] discusses the use of alternating projection techniques in the context of other color-related applications.

C. Example Illuminant and Surface Calculation

We have calculated surface and illuminant two-mode linear models, using the surface-reflectance functions from a collection of 462 Munsell chips measured by Kelly and reported by Nickerson.[25] We have used blackbody radiators (3K, 4K, 5K, 6K, 9K) and CIE standard illuminants a, b, and c as illuminants. The illuminants and surfaces were represented as 31 dimensional vectors representing lights from 400 to 700 nm at 10-nm intervals. The illuminant vectors were normalized to unit length; they are plotted in Fig. 8. We selected the linear models by minimizing the error in the CIE standard observer's tristimulus values XYZ.

Figure 9 compares the principal-components linear models for the illuminants and surfaces with the two-mode linear models. Figures 9(a) and 9(b) contain the first three principal-components sampling functions of the surface and the illuminant functions, respectively; Fig. 9(c) and 9(d) contain the first three two-mode sampling functions. The differences between the two-mode sampling functions and the principal-components functions arise mainly because the human eye is insensitive in the short- and long-wavelength regions of the spectrum.

The two-mode analysis necessarily performs better at minimizing the squared error in the prediction of the sensor responses. To evaluate whether the improvement is perceptually salient, we plot the error by using a

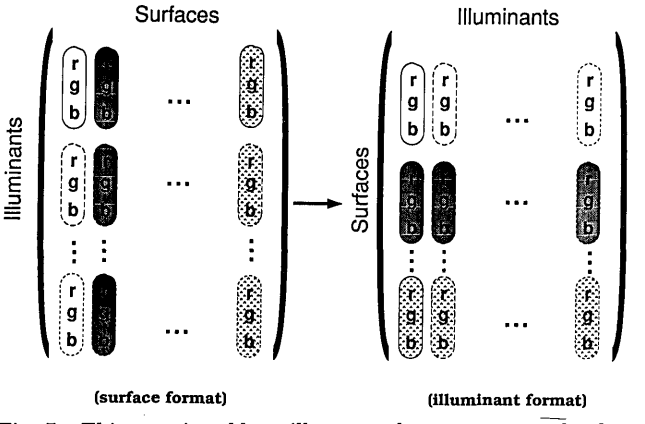

Fig. 7. This matrix tableau illustrates how to convert the data from the format used for the one-mode analysis of surfaces (left) to the one-mode analysis of illuminants (right). The operation is essentially a transposition, but it is applied to the rgb vectors of data rather than to the individual elements.

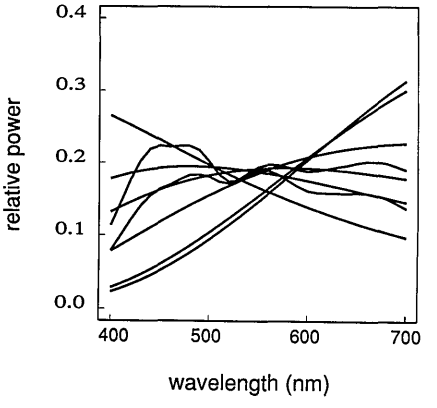

Fig. 8. This figure contains the spectral-power distributions of the illuminants used in our calculation. There are five blackbody radiators (3K, 4K, 5K, 6K, 9K) and three CIE standard illuminants (a, b, and c). The vectors representing the illuminants were normalized to unit length.

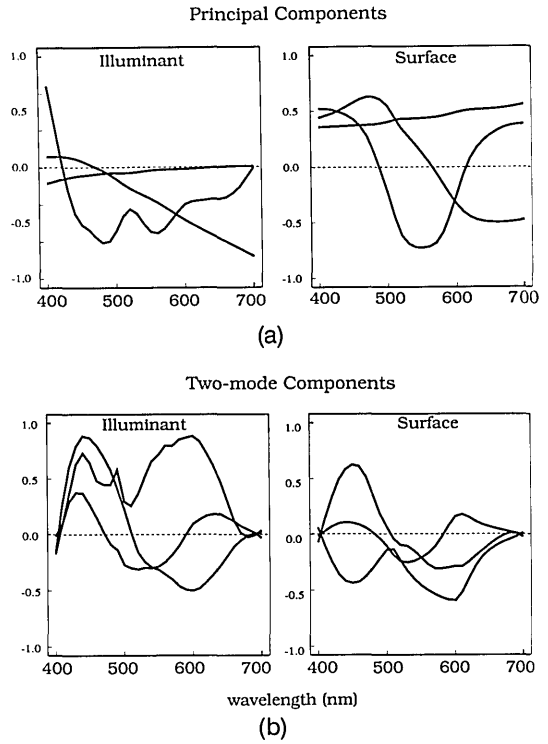

Fig. 9. (a) Sampling functions for three-dimensional surface and illuminant models, respectively, with the use of principal-components methods. (b) Sampling functions for the surface and illuminant models, respectively, with the use of two-mode methods. The linear models were built for a collection of 462 Munsell chips. The collection of illuminants is described in the caption for Fig. 8.

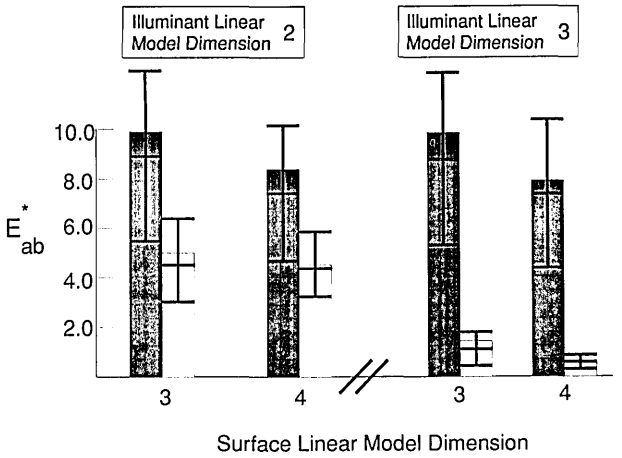

Fig. 10. Comparison of ΔE_{ab} values for different dimensions of the two-mode and the principal-components linear models. The sensor data are the XYZ values of Munsell chips rendered under the illuminants plotted in Fig. 8. The filled bars are the principal-component errrors, and the open bars are the two-mode errors. The height of the bars defines the mean error from the $462 \times 8 = 3696$ illuminant–surface pairs. The horizontal lines define the twenty-fifth-, fiftieth- (mode), and seventy-fifth-percentile errors.

perceptual-error measure, the CIE ΔE_{ab}. The bar graph in Fig. 10 is grouped into two parts. At the left we show the mean ΔE_{ab} error (bar height) and the quartiles (horizontal lines) for a two-dimensional illuminant model and for three- and four-dimensional surface models. At the right we show the error for a three-dimensional illuminant model, again for three- and four-dimensional surface models. Errors less than three ΔE_{ab} units are not visually significant. The two-mode representation performs significantly better than the principal-components representation. The two-mode representation performs nearly perfectly for three-dimensional models of the surface and illuminant functions and better than the principal-components model everywhere.

7. DISCUSSION

A. Linear Models for Devices

The two-mode analysis obtains efficiencies beyond the one-mode method by discovering structure in the matrices $\mathbf{T}_{\mathbf{E}_i}$. The two-mode procedure approximates these matrices as the weighted sum of a few matrices, a basis set. The set of basis matrices defines a linear model for the observed surface-transfer matrices. Each matrix in the basis set is associated with a hypothetical device. The surface-transfer matrix of the hypothetical device combines the sensor responsivities with one illuminant. The output of the real devices is the weighted sum of the outputs of the hypothetical devices.

When all the $\mathbf{T}_{\mathbf{E}_i}$ matrices are derived from one set of sensors, with only the illumination varying, we can interpret the illuminant format weights of the two-mode analysis as a linear model for the illuminants. When the surface-transfer matrices include more than one set of sensor responsivities, we can still apply the two-mode analysis. For example, suppose that the surface-transfer matrices describe a collection of flatbed scanners. We can then build a linear model that describes the outputs of all the flatbed scanners as the weighted sum of outputs from a few hypothetical scanners. In that type of application we cannot interpret the two-mode weights as an illumination model. Rather, the two-mode weights describe the conjunction of illumination and sensor variation.

Related Work

A number of investigators have explored spectral representations of surface reflectances and illuminants, using Gaussian quadrature (GQ) approximations. GQ representations approximate surface reflectance by using functions whose nonzero values are limited to fall at a small number of sample wavelengths.[26–28]

For simple renderings it appears that GQ approximations are significantly less accurate than the two-mode models.[28] But GQ approximations may be much easier to use for computer-graphics calculations of interreflections. The proper architecture for including interreflection effects and spectral representations based on linear models remains open. It may be necessary to build an ordered series of linear models, applying them in turn for each interreflection calculation. Or it may be possible to build a single linear model that incorporates all the interreflection functions.

8. CONCLUSIONS

Ordinarily, tristimulus coordinates serve as the input for the psychological phenomena of color appearance. Color-appearance models begin at the sensor encoding and flow forward toward psychological phenomena. In this

paper we reverse the direction of analysis. We conceive of the tristimulus coordinates as the output of the image-formation process. Spectral representations begin at the sensor encoding and flow backward to the image-formation process.

If color perception serves to estimate the physical factors of image formation, then the two directions for analyzing and representing color may be similar to each other.

The key to our analysis is the observation that tristimulus coordinates, or indeed the sensor responses of any linear device, provide a measure of the color signal. Linear sensor responses measure that part of the incident color signal that falls within the subspace defined by the span of the sensors' color-matching functions. When the illuminant is known, we can also use the sensor responses to measure that part of the surface-reflectance function that falls within a subspace that is defined by the row space of the surface-transfer matrix (i.e., the product of the color-matching functions and the illuminant spectral-power distribution). In this paper we emphasize that the sensor responses estimate the physical signal, because we wish to develop closer ties between perceptual color representations and the physical factors in image formation.

We have described two applications for spectral representations of color information. First, we used the one-mode method to construct linear models for surface reflectances. We described an application of these representations to printer calibration. When we wish to approximate sensor responses, the one-mode representation performs better than the principal-components representation. Second, we used the two-mode method to derive surface and illuminant functions simultaneously. Again, for predicting sensor responses, the two-mode models perform better than the principal-components models.

ACKNOWLEDGMENTS

We thank E. J. Chichilnisky, J. Farrell, K. Fishkin, D. Heeger, D. Kimber, and A. Poirson. The research reported in this paper was partially supported by National Eye Institute contract RO1 EY03164 and NASA contract 2-307.

B. A. Wandell is on leave from the Department of Psychology, Stanford University, Stanford, California.

REFERENCES

1. P. S. Parkkinen, J. Hallikainen, and T. Jaaskelainen, "Characteristic spectra of Munsell colors," J. Opt. Soc. Am. A **6,** 318–322 (1989).
2. J. Cohen, "Dependency of the spectral reflectance curves of the Munsell color chips," Psychonom. Sci. **1,** 369–370 (1964).
3. L. T. Maloney, "Evaluation of linear models of surface spectral reflectance with small numbers of parameters," J. Opt. Soc. Am. A **3,** 1673–1683 (1986).
4. D. B. Judd, D. L. MacAdam, and G. W. Wyszecki, "Spectral distribution of typical daylight as a function of correlated color temperature," J. Opt. Soc. Am. **54,** 1031 (1964).
5. L. R. Tucker, "Some mathematical notes on three-mode factor analysis," Psychometrika **31,** 279–311 (1966).
6. J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics* (Wiley, New York, 1988).
7. A. Kapteyn, H. Neudecker, and T. Wansbeek, "An approach to n-mode components analysis," Psychometrika **51,** 269–275 (1986).
8. P. M. Kroonenberg and J. de Leeuw, "Principal component analyses of three-mode data by means of alternating least squares algorithms," Psychometrika **45,** 69–97 (1980).
9. G. H. Golub and C. F. van Loan, *Matrix Computations* (Johns Hopkins U. Press, Baltimore, Md., 1983).
10. M. Tsukada and Y. Ohta, "An approach to color constancy using multiple images," presented at the Third International Conference on Computer Vision, December 4–7, 1990, Osaka, Japan.
11. M. D'Zmura, "Color constancy: surface color from changing illumination," J. Opt. Soc. Am. A **9,** 490–493 (1992).
12. K. Takahama and Y. Nayatani, "New method for generating metameric stimuli of object colors," J. Opt. Soc. Am. **62,** 1516–1520 (1972).
13. J. B. Cohen and W. E. Kappauf, "Metameric color stimuli, fundamental metamers, and Wyszecki's metameric blacks: theory, algebra, geometry, application," Am. J. Psychol. **95,** 537–564 (1982).
14. J. B. Cohen and W. E. Kappauf, "Color mixture and fundamental metamers: theory, algebra, geometry, application," Am. J. Psychol. **98,** 171–259 (1985).
15. J. B. Cohen, "Color and color mixture: scalar and vector fundamentals," Color Res. Appl. **13,** 4–39 (1988).
16. H. J. Trussell, "Applications of set theoretic methods to color systems," Color Res. Appl. **16,** 31–41 (1991).
17. S. A. Burns, J. B. Cohen, and E. N. Kuznetsov, "Multiple metamers: preserving color matches under diverse illuminants," Color Res. Appl. **14,** 16–22 (1989).
18. E. H. Adelson, E. Simoncelli, and R. Hingorani, "Orthogonal pyramid transforms for image coding," in *Visual Communications and Image Processing II,* T. R. Hsing, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **845,** 50–57 (1987).
19. J. Dongarra, J. R. Bunch, C. B. Moler, and G. W. Stewart, *LINPACK Users Guide* (Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1978).
20. C. S. McCamy, H. Marcus, and J. G. Davidson, "A color-rendition chart," J. Appl. Photog. **48,** 777–784 (1976).
21. L. T. Maloney, "Photoreceptor spectral sensitivities and color correction," in *Perceiving, Measuring, and Using Color,* M. H. Brill, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **1250,** 103–110 (1990).
22. M. S. Drew and B. V. Funt, "Natural metamers," Computer Vision Graphics Image Process. (to be published).
23. M. J. Vrhel and H. J. Trussell, "Color correction using principal components," Color Res. Appl. (to be published).
24. D. H. Brainard, B. A. Wandell, and W. B. Cowan, "Black light: how sensors filter spectral variation of the illuminant," IEEE Trans. Biomed. Eng. **36,** 140–149 (1989).
25. D. Nickerson, *Spectrophotometric Data for a Collection of Munsell Samples* (U.S. Department of Agriculture, Washington, D.C., 1957).
26. C. F. Borges, "Trichromatic approximation method for surface illumination," J. Opt. Soc. Am. A **8,** 1319–1323 (1991).
27. R. Wallis, "Fast computation of tristimulus values by use of Gaussian quadrature," J. Opt. Soc. Am. **65,** 91–94 (1975).
28. G. W. Meyer, "Wavelength selection for synthetic image generation," Comput. Vision Graphics Image Process. **41,** 57–79 (1988).