Optimization of Sensor Response Functions for Colorimetry of Reflective and Emissive Objects

Mark Wolski, Charles A. Bouman, *Member, IEEE*, Jan P. Allebach, *Fellow, IEEE*, and Eric Walowit

***Abstract*—This paper describes the design of color filters for a surface color measurement device. The function of the device is to return the XYZ tristimulus vector characterizing the color of the surface. The device is designed to measure emissive as well as reflective surfaces. It uses an internal set of LED's to illuminate reflective surfaces while characterizing their color under assumed standard illuminants.**

In the design of the filters, we formulate a nonlinear optimization problem with the goal of minimizing error in the uniform color space CIE $L^*a^*b^*$. Our optimization criteria employs a technique to retain a linear structure while approximating the true $L^*a^*b^*$ error. In addition, our solution is regularized to account for system noise, filter roughness, and filter implementation errors. Experimental results indicate average and worst-case device accuracy of 0.27 $L^*a^*b^*$ ΔE units and 1.56 $L^*a^*b^*$ ΔE units for a "system tolerance" of 0.0005.

I. Introduction

A. Overview

COLORIMETRY is the science of characterizing spectral distributions of wavelength by a set of three numbers called tristimulus values. This can be thought of as mapping the infinite-dimensional space of all spectral distributions into the 3-D space of tristimulus values. Lately, there is considerable interest in colorimetric matching of surfaces of different types such as video monitors (soft copy) and color prints (hard copy). The fundamental difference between these two is that hard copy characterization requires specification of the spectral distribution of the illuminant, whereas soft copy characterization does not. An additional complication for hard copy characterization is that tristimulus values describing the surface under one illuminant cannot be transformed to the tristimulus values under a different illuminant, even if the spectral distributions of both illuminants are known. To overcome this, hard copy characterizations are normally made for several different illuminants with the hope that they represent the majority of viewing situations.

In this paper, we discuss the design of a set of color filters for a device which measures surface colors. The function of the device is to return the XYZ tristimulus vector (3-tuple) describing the surface. A feature of the device is its ability to measure both emissive surfaces and reflective surfaces. To measure reflective surfaces, the device uses a set of light emitting diodes (LED's) as its internal illumination source but returns tristimulus values describing the surface under either an equal energy (EE) illuminant or a D65 illuminant [1]. Overall there are three modes of device operation: emissive, reflective/EE, and reflective/D65.

For each of the three operating modes, it is possible to specify a set of three color filters that alter the spectral response of three detector channels so that the channel responses match those of the XYZ color matching functions under the illuminant of interest. Thus, with nine separate filters, no restrictions on the shapes of the filter transmittance functions, and noise-free components, we could design an error-free device.

In order to limit the cost and size of the device, we limit the number of device filters to four. Roughly, we can consider the design problem as one of approximating the nine ideal filters with linear combinations of four filters. Our approach to solving the problem is based on linear vector space models where spectral distributions are represented as 31-element vectors [2], [3].

The filter design problem has been studied before [4]–[10]. Davies and Wyszecki present a method for optimally combining a fixed set of known filters to minimize the integral of the spectral error between the desired function and its approximation [4]. Engelhardt and Seitz use a steepest descent algorithm in $Lu'v'$ [1] space to select coefficients for combining fixed sets of filters. They choose the filter sets heuristically and use device noise models to pick the best set after the corresponding coefficients have been found [5]. Kollarits and Gibbon also choose filter coefficients for fixed sets of bandpass filters [6]. They perform a linear optimization using CIE XYZ as the error metric and then use perturbation techniques to improve the error performance in $Lu'v'$ space. Vrhel and Trussell formulate an iterative approach for optimizing sensing functions in the presence of noise for given ensembles of spectral stimuli [7], [8]. They formulate an error metric based on vector norms in a subspace representing the projection of the desired space into the complement of the sensor space. Vora and Trussell formulate a different iterative optimization using a similar error metric [9], [10]. They introduce the idea of parameterizing the filter space to reduce the number of degrees of freedom of the problem. This makes the optimization more tractable

Manuscript received September 11, 1994; revised February 8, 1995. The associate editor coordinating the review of this paper and approving it for publication was Dr. Frederick Mintzer.

M. Wolski was with the School of Electrical Engineering, Purdue University, West Lafayette, IN 47907 USA. He is now with General Motors Research and Development Center, Warren, MI 48090-9055 USA.

C. A. Bouman and J. P. Allebach are with the School of Electrical Engineering, Purdue University, West Lafayette, IN 47907 USA.

E. Walowit is with the Color Savvy Systems Inc., Springboro, OH 45066 USA..

Publisher Item Identifier S 1057-7149(96)01794-0.

1057–7149/96$05.00 © 1996 IEEE

and forces the resulting filter functions to be smooth, thus increasing the likelihood that they can be manufactured to within reasonable tolerances [11], [12]. However, it precludes some solutions that may not be expressible under the particular parameterization.

Our approach to filter selection is also based on an iterative optimization technique. However, our design criteria include the minimization of error in CIE $L^*a^*b^*$ space [1] over a standard ensemble of stimuli. The ensemble consists of 752 spectral reflectance samples of which 462 are evenly spaced throughout the Munsell color solid, 120 are taken from a collection of Dupont paint chips, and 170 represent various man-made and natural materials [13], [14]. The significance of the $L^*a^*b^*$ error norm is that it is directly related to the perceptibility of color shifts. In addition to minimizing perceptibility of color shifts, we regularize our solution to take into account system noise and variations and filter smoothness.

The transformation from XYZ tristimulus values to $L^*a^*b^*$ values is nonlinear. This means that, given true tristimulus values and device output values, we must perform two sets of nonlinear transformations to find the $L^*a^*b^*$ error vector. However, the error vector can be locally approximated by linearization of the $L^*a^*b^*$ transformations. Since the error vectors are assumed small, the local linearization is a very accurate representation of true $L^*a^*b^*$ errors.

In carrying out the linearizations, we show that the norm of the mean square $L^*a^*b^*$ error over the ensemble of samples can be expressed as a linear transformation of the difference between the XYZ color matching functions and their approximations from device components. This allows us to express a complex design problem, over a large gamut of colors, as a much simpler problem.

There are two parts to our filter optimization problem. The first is to select an optimal set of filters. The second is to select an optimal set of coefficients to combine the outputs of the filter channels. Our solution method uses the fact that once the filters are known, the coefficients can be determined using standard least-squares techniques. The overall solution is based on an iterative gradient descent technique [15]. To solve the iterative optimization, we use the MATLAB® optimization toolbox [16] running on a Hewlett Packard 735 workstation. At each step of the iteration, we use standard MATLAB software routines to solve for the optimal filter coefficients.

After a brief section on notation, the remainder of this paper is organized as follows. In Section II, we develop the error models for the three modes of operation. In Section III, we add regularization constraints to the models and formulate the overall error metric. In Section IV, we present our solution methodology. In Section V, we give experimental results.

B. Notation

We use finite dimensional representations of all continuous spectral functions. All spectral distributions are sampled at 10-nm intervals from 400–700 nm and represented as 31-element vectors. Vectors are represented as boldface lowercase letters, matrices as boldface upper case letters, and the individual elements of vectors or matrices as subscripted normal face letters. All vectors are assumed column vectors unless given the superscript ^{T} representing vector (or matrix) transpose. For a vector $\boldsymbol{v}$, the symbol $[\boldsymbol{v}]$ represents a diagonal matrix with v as the main diagonal. A matrix made up of the side-by-side placement of several column vectors $\boldsymbol{v}_i$ will be represented as $[\boldsymbol{v}_1|\boldsymbol{v}_2|\cdots|\boldsymbol{v}_n]$. Appendix A contains a list of symbols used throughout this paper.

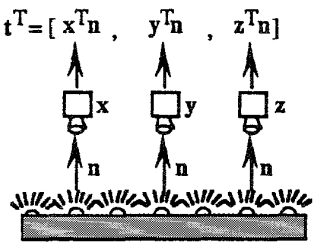


Fig. 1. Representation of an ideal sensor acting on an emissive stimulus to produce its XYZ tristimulus vector.

II. System Models

In this section, we develop error models for the three modes of operation. In each mode, the device output is an estimate of the XYZ tristimulus vector $\boldsymbol{t}$ describing a surface. We formulate a scalar error metric representing the perceived magnitude of color shift due to errors in this tristimulus estimate.

A. Emissive Mode

An emissive device produces its stimulus without regard to the ambient illumination. Fig. 1 represents an ideal set of sensing functions operating on the emissive stimulus $\boldsymbol{n}$ to produce its tristimulus characterization $\boldsymbol{t}$. Fig. 2 represents the device model, which also operates on the stimulus $\boldsymbol{n}$ to produce an estimate of the tristimulus characterization. The device consists of four channels, each of which has a unique filter $\boldsymbol{f}_k$ and a common detector $\boldsymbol{d}$. The detector spectral response for this model is shown in Fig. 3. The four channels, which are inner products of the stimulus with the filter-detector combination, are combined linearly to produce estimates of X, Y, and Z, which are the elements of $\boldsymbol{t}$. Let $\boldsymbol{A}$ be the (3×31) operator matrix whose rows are the color matching functions $\boldsymbol{x}^T$, $\boldsymbol{y}^T$, and $\boldsymbol{z}^T$. Let $\boldsymbol{F}_a$ be the (4×31) operator matrix defined as

$$\boldsymbol{F}_a = \{[\boldsymbol{f}_1]\boldsymbol{d}|[\boldsymbol{f}_2]\boldsymbol{d}|[\boldsymbol{f}_3]\boldsymbol{d}|[\boldsymbol{f}_4]\boldsymbol{d}\}^T. \tag{1}$$

In the above, $\boldsymbol{d}$ is the device detector spectral response vector, and $\boldsymbol{f}_k$ is a filter transmittance vector. Finally, let $\boldsymbol{T}_a$ be a (3×4) coefficient matrix represented by the matrix multiply stage of Fig. 2. Then, the tristimulus error vector can be represented as

$$\Delta\boldsymbol{t} = (\boldsymbol{A} - \boldsymbol{T}_a\boldsymbol{F}_a)\boldsymbol{n}. \tag{2}$$

Our design goal will be to determine both $\boldsymbol{F}_a$ and $\boldsymbol{T}_a$, but for now, we assume that these matrices are given so that we can concentrate on developing the error metric. As a first thought, we might take $||\Delta\boldsymbol{t}||^2$ to be the error metric, where

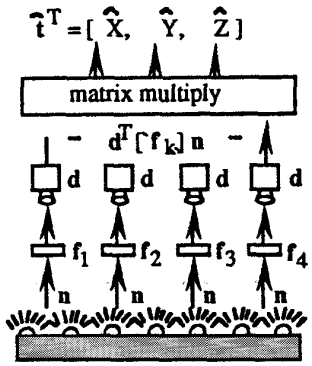


Fig. 2. Representation of the device components in emissive mode as they produce an estimate of the XYZ tristimulus vector.

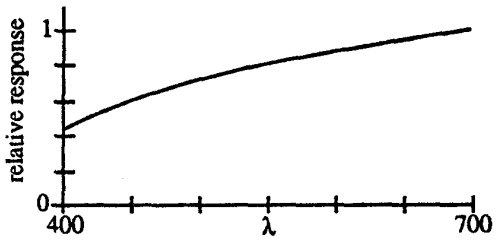


Fig. 3. Spectral response function of the detector $\boldsymbol{d}$ used in the model.

$||\cdot||$ represents the usual vector 2-norm. However, we are interested in quantifying perceived color shifts, and this metric is inappropriate for that task. A better metric would be the norm of the tristimulus error expressed in uniform color space coordinates such as CIE $L^*a^*b^*$. We will call the uniform tristimulus vector $\boldsymbol{u}$. It is determined from the elements X, Y, and Z of t as

$$\boldsymbol{u} = \begin{bmatrix} L^* \\ a^* \\ b^* \end{bmatrix} = \begin{bmatrix} 116\left(\frac{Y}{Y_w}\right)^{1/3} - 16 \\ 500\left[\left(\frac{X}{X_w}\right)^{1/3} - \left(\frac{Y}{Y_w}\right)^{1/3}\right] \\ 200\left[\left(\frac{Y}{Y_w}\right)^{1/3} - \left(\frac{Z}{Z_w}\right)^{1/3}\right] \end{bmatrix}. \tag{3}$$

The constants X_w, Y_w, and Z_w in (3) represent the XYZ tristimulus coordinates of a reference white stimulus. For this emissive mode, we take the reference white to be a vector of all 1's so that the tristimulus coordinates are simply the sums of the elements of the x, y, and z color matching vectors. For our 31-element spectral vector case, $(X_w\, Y_w,\, Z_w) = (10.7,\, 10.7,\, 10.7)$.

The error vector $\Delta\boldsymbol{u}$ is the difference between a nonlinear transformation of $\boldsymbol{t}$ and a nonlinear transformation of its estimate. We can approximate this difference by taking the derivative of (3) near a fixed $\boldsymbol{n}_o$no (having an associated tristimulus vector $\boldsymbol{t}_o = [X_o,\, Y_o,\, Z_o]^T$) and multiplying it by the XYZ tristimulus error $\Delta\boldsymbol{t}$. In other words, $\Delta\boldsymbol{u_o} \approx \boldsymbol{J_o}\Delta\boldsymbol{t_o}$, where

$$\boldsymbol{J}_o = \frac{1}{3}\begin{bmatrix} 0 & 116 & 0 \\ 500 & -500 & 0 \\ 0 & 200 & -200 \end{bmatrix} \cdot \begin{bmatrix} X_w^{-1/3}X_w^{-2/3} & 0 & 0 \\ 0 & T_w^{-1/3}Y_o^{-2/3} & 0 \\ 0 & 0 & Z_w^{-1/3}Z_o^{-2/3} \end{bmatrix}. \tag{4}$$

It is customary to use the symbol ΔE to represent error distances (2-norms) in $L^*a^*b^*$ space. In terms of our device components, the linearized error norm is then expressible as

$$\begin{aligned} \Delta E_o &= ||\Delta\boldsymbol{u}_o||_2 \\ &\approx ||\boldsymbol{J}_o(\boldsymbol{A} - \boldsymbol{T}_a\boldsymbol{F}_a)\boldsymbol{n}_o||_2. \end{aligned} \tag{5}$$

A problem with (5) is that it is only accurate for stimuli near n_o. To overcome this limitation, we used an ensemble $\{n_k\}$ of 752 stimuli, taken from Munsell chips, Dupont paint samples, and various real-world materials [13], [14]. Letting N represent the number of sample stimuli, we form the expression for their mean square error:

$$\Delta E^2_{emis} = \frac{1}{N}\sum_k ||\boldsymbol{J}_k(\boldsymbol{A} - \boldsymbol{T}_a\boldsymbol{F}_a)\boldsymbol{n}_k||_2^2. \tag{6}$$

Our objective will be to minimize (6) over all possible $(\boldsymbol{T}_a\boldsymbol{F}_a)$. As a first step in doing this, we will write an equivalent expression for (6) that will allow us to perform the summation before attempting to solve for $(\boldsymbol{T}_a\boldsymbol{F}_a)$. Notational details and proofs for the following are found in Appendix C. The operator vec( ) creates a vector out of a matrix by stacking its columns (C5). The symbol $\otimes$ represents the Kronecker product (C1). Using these operators and (C7), we can rewrite (6) as

$$\Delta E^2_{emis} = \frac{1}{N}\sum_k ||(\boldsymbol{n}_k^T \otimes \boldsymbol{J}_k)\,\mathrm{vec}\,(\boldsymbol{A} - \boldsymbol{T}_a\boldsymbol{F}_a)||_2^2. \tag{7}$$

For notational convenience, let $\Delta\boldsymbol{a} = \mathrm{vec}(\boldsymbol{A} - \boldsymbol{T}_a\boldsymbol{F}_a)$, and let $\boldsymbol{B}_k = (\boldsymbol{n}_k^T \otimes \boldsymbol{J}_k)$. Then, using the inner product representation of the 2-norm, we can rewrite (7) as

$$\Delta E^2_{emis} = \Delta\boldsymbol{a}^T\left(\frac{1}{N}\sum_k \boldsymbol{B}_k^T\boldsymbol{B}_k\right)\Delta\boldsymbol{a}. \tag{8}$$

Note that in (8), each $\boldsymbol{B}_k$ is completely determined by $\boldsymbol{n}_k$ since $\boldsymbol{J}_k$ is completely determined by $\boldsymbol{n}_k$. Thus, given the ensemble $\{\boldsymbol{n}_k\}$, we can precompute the symmetric matrix $(1/N)\Sigma_k\boldsymbol{B}_k^T\boldsymbol{B}_k$. Moreover, we can find a single matrix $\boldsymbol{B}_{eq}$ (whose rows are the eigenvectors of $(1/N)\Sigma_k\boldsymbol{B}_k^T\boldsymbol{B}_k$ scaled by the square roots of the eigenvalues) such that $\boldsymbol{B}_{eq}^T\boldsymbol{B}_{eq} = (1/N)\Sigma_k\boldsymbol{B}_k^T\boldsymbol{B}_k$. We should note that the size of $\boldsymbol{B}_k$ is (3×93), but the size of $\boldsymbol{B}_{eq}$ may be as large as (93×93). Using $\boldsymbol{B}_{eq}$, we write the emissive mode error expression as

$$\Delta E^2_{emis} = ||\boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{A}) - \boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{T}_a\boldsymbol{F}_a)||_2^2. \tag{9}$$

A few comments about (9) are in order. First, the complexity of the expression is fixed regardless of the number of sample stimuli in the ensemble $\{\boldsymbol{n}_k\}$. We can increase the size of the

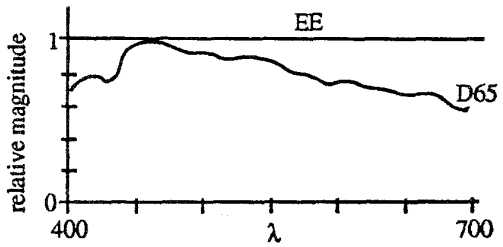


Fig. 4. Standard illuminants—Equal energy i_{EE} and D65 i_{D65}.

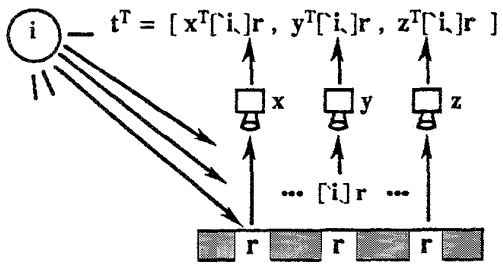


Fig. 5. Representation of an ideal sensor acting on a reflective surface under a fixed illuminant to produce the XYZ tristimulus vector.

ensemble without incurring any time penalty in the optimization. Second, B_{eq} can be set equal to the identity matrix, in which case, (9) becomes the sum of the squared differences between the elements of the color matching functions (A) and their approximations ($T_a F_a$). This is a useful optimization criteria for the case where we do not have an ensemble of sample spectra.

B. Reflective Modes

Tristimulus values describing a reflective surface will vary, depending on the spectral distribution of light illuminating that surface. Furthermore, there is no exact transformation of tristimulus coordinates from those taken under one illuminant to those taken under a different illuminant. For this reason, we must specify the spectral distribution of the illuminant when characterizing reflective surfaces by tristimulus values. It is customary to specify reflective surface tristimulus values under a small number of standard illuminants and to assume that one of these will be close to the desired viewing illuminant. For our model, we use the two standard illuminants illustrated in Fig. 4. The equal energy illuminant i_{EE} is a vector of all 1's. The D65 illuminant i_{D65} is meant to resemble daylight [1].

The ideal measurement situation is illustrated in Fig. 5. The stimulus is the point-by-point product of the illumination and the reflectance. Otherwise, the model is the same as Fig. 1. The corresponding device model is shown in Fig. 6. The main difference between this and the emissive device model of Fig. 2 is the inclusion of illumination sources. We use five different LED's as illumination sources. This keeps the device rugged and compact. The spectral distributions of the LED's are shown in Fig. 7.

The standard illuminants of the ideal measurement model (Figs. 4 and 5) are not the same as the LED's used in the device model (Figs. 6 and 7). We could have attempted to

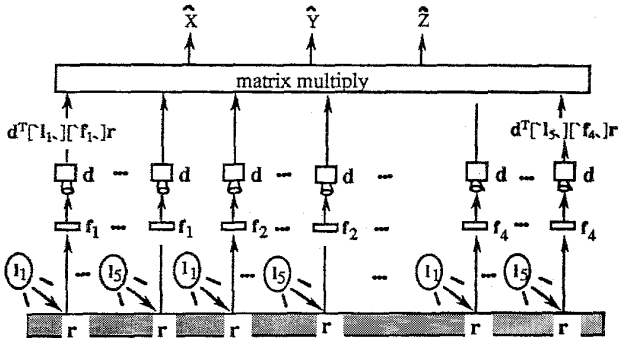


Fig. 6. Representation of the device components in reflective mode as they use a combination of LED's to illuminate a surface and produce an estimate of the XYZ tristimulus vector.

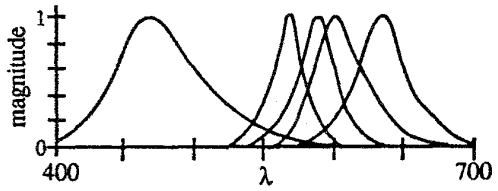


Fig. 7. Normalized spectral distributions of five LED's used as device illuminants.

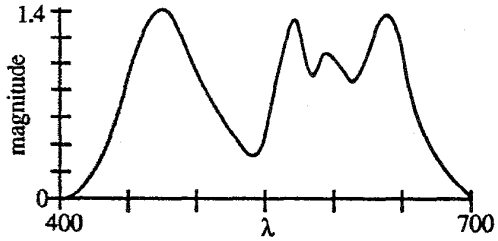


Fig. 8. Linear combination of LED's from Fig. 7 that best approximates an equal energy illuminant.

simulate the standard illuminants by combining the LED's, but this would be very inaccurate. For example, if we were to try to approximate the EE illuminant of Fig. 4 using a linear combination of LED's from Fig. 7 (by minimizing the norm of the spectral error vector), we would obtain the distribution shown in Fig. 8. We make up for the inability to approximate the standard illuminants by using each of the LED's independently. In that case, we can combine each of the five LED's with each of the four filters to give us the 20 sensing channels shown in Fig. 6. The 20 channels greatly increase the design flexibility over that of the four-channel device that would result if we used a fixed linear combination of the five LED's.

We now have all the tools necessary to define the two reflective mode error metrics. Similar to the emissive mode, define A_{EE} and A_{D65} as

$$\begin{aligned} A_{EE} &= A[i_{EE}] \\ &= \{[i_{EE}]x\,|[i_{EE}]y|\,[i_{EE}]z\}^T \end{aligned} \tag{10}$$

and

$$\begin{aligned}\boldsymbol{A}_{D65} &= \boldsymbol{A}[\boldsymbol{i}_{D65}] \\ &= \{[\boldsymbol{i}_{D65}]\boldsymbol{x}\,|[\boldsymbol{i}_{D65}]\boldsymbol{y}|\,[\boldsymbol{i}_{D65}]\boldsymbol{z}\}^T. \end{aligned} \tag{11}$$

Let $\boldsymbol{l}_i$, $i = 1, 2, \cdots 5$, be the spectral distributions of the LED's and $\boldsymbol{f}_k$, $k = 1, \cdots 4$ be the spectral transmittances of the filters. Use these to define a (20×31) channel matrix $\boldsymbol{F}_b$ as

$$\begin{aligned}\boldsymbol{F}_b = \{&[\boldsymbol{f}_1][\boldsymbol{l}_1]\boldsymbol{d}\,|\cdots|\,[\boldsymbol{f}_1][\boldsymbol{l}_5]\boldsymbol{d}|\,[\boldsymbol{f}_2][\boldsymbol{l}_1]\boldsymbol{d}\,|\cdots| \\ &[\boldsymbol{f}_2][\boldsymbol{l}_5]\boldsymbol{d}\,|\cdots|\,[\boldsymbol{f}_4][\boldsymbol{l}_1]\boldsymbol{d}\,|\cdots|\,[\boldsymbol{f}_4][\boldsymbol{l}_5]\boldsymbol{d}\}^T. \end{aligned} \tag{12}$$

Analogous to (9), we wish to approximate $\boldsymbol{A}_{EE}$ or $\boldsymbol{A}_{D65}$ using linear combinations of the rows of $\boldsymbol{F}_b$. Let $\boldsymbol{T}_{EE}$ and $\boldsymbol{T}_{D65}$ be (4×20) coefficient matrices. Then, the error expressions for the reflective modes are

$$\Delta E_{2EE} = ||\boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{A}_{EE}) - \boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{T}_{EE}\boldsymbol{F}_b)||_2^2 \tag{13}$$

and

$$\Delta E_{D65}^2 = ||\boldsymbol{B}_{eq}\mathrm{vec}(\boldsymbol{A}_{D65}) - \boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{T}_{D65}\boldsymbol{F}_b)||_2^2. \tag{14}$$

III. Regularization

In the previous sections, we developed three error metrics that quantify the perceptibility of color shifts for the three modes of operation. Before going on to solve for the unknowns, we need to consider two other factors: filter realizability and design robustness.

The solution filter set that we specify must be implementable. This means that its elements must be nonnegative. However, nonnegativity is not a sufficient constraint for generating realizable filters. Each filter must also be a reasonably smooth function.

The robustness consideration has to do with device noise as well as component manufacturing tolerances. We have no control over these two items, but we can control their effect on output errors. Specifically, we want to be careful so that small amounts of component noise do not get unduly magnified.

A. Filter Feasibility

In order to have a reasonable assurance that our filter design can be built, we attempt to make the filter functions as smooth as possible. As our measure of filter roughness, we use the integral of the square of the second derivative of the filter function (i.e., the sum of the squares of the second difference of the filter vector elements). We define the (31×31) second difference operator $\boldsymbol{D^2}$ as

$$\boldsymbol{D}^2 = \begin{bmatrix} 1 & -1 & 0 & & \cdots & 0 \\ -1 & 2 & -1 & 0 & & \\ 0 & -1 & 2 & \ddots & \ddots & \vdots \\ & 0 & -1 & \ddots & -1 & 0 \\ \vdots & & \ddots & \ddots & 2 & -1 \\ 0 & & \cdots & -1 & 1 & \end{bmatrix}. \tag{15}$$

For a filter vector $\boldsymbol{f}$, $\|\boldsymbol{D^2 f}\|_2^2$ is its roughness cost. For the filter matrix $\boldsymbol{F}$ with four filter functions, the overall roughness cost is a scaled sum of the individual costs. The scaling constant K_s is used to adjust the relative importance of this cost in relation to other cost factors (color shift perceptibility and noise sensitivity). In equation form, the roughness cost is given by

$$\varepsilon_s = K_s(||\boldsymbol{D}^2\boldsymbol{f}_1||_2^2 + ||\boldsymbol{D}^2\boldsymbol{f}_2||_2^2 + ||\boldsymbol{D}^2\boldsymbol{f}_3||_2^2 + ||\boldsymbol{D}^2\boldsymbol{f}_4||_2^2). \tag{16}$$

B. Design Robustness

In this section, we consider device sensitivity to noise and/or variations in the physical components. We want the solution to be robust in the presence of variations in the output spectra of the LED illuminants, variations in the spectral transmittances of the filters, and sensor noise. The combined action of all these components is represented in the matrix $\boldsymbol{F}_a$ of (1) or $\boldsymbol{F}_b$ of (12). We represent component variations and noise as perturbations of individual elements of this matrix. Thus, we replace $\boldsymbol{F}_a$ by $(\boldsymbol{F}_a + \Delta\boldsymbol{F}_a)$ and $\boldsymbol{F}_b$ by $(\boldsymbol{F}_b + \Delta\boldsymbol{F}_b)$. We will detail the development of a new (augmented) error metric for the emissive mode. The reflective mode development is exactly the same. With the substitution of $\boldsymbol{F}_a$, the emissive mode error expression (9) becomes

$$\begin{aligned}\Delta E_{emis}^2 = ||&\boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{A}) - \boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{T}_a\boldsymbol{F}_a) \\ &- \boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{T}_a\Delta\boldsymbol{F}_a)||_2^2. \end{aligned} \tag{17}$$

This expression can be bounded by the sum of two squared norms (see (B5) of Appendix B).

$$\begin{aligned}\Delta E_{emis}^2 \le\ &2||\boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{A}) - \boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{T}_a\boldsymbol{F}_a)||_2^2 \\ &+ 2||\boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{T}_a\Delta\boldsymbol{F}_a)||_2^2. \end{aligned} \tag{18}$$

In sections that follow, we will have reason to consider (18) as a function of $\mathrm{vec}(\boldsymbol{T}_a)$. Let $\boldsymbol{I}_3$ be a (3×3) identity matrix. Then, using (C7) from Appendix C, we can rewrite (18) as

$$\begin{aligned}\Delta E_{emis}^2 \le\ &2||\boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{A}) - \boldsymbol{B}_{eq}(\boldsymbol{F}_a^T \otimes \boldsymbol{I}_3)\,\mathrm{vec}\,(\boldsymbol{T}_a)||_2^2 \\ &+ 2||\boldsymbol{B}_{eq}(\Delta\boldsymbol{F}_a^T \otimes \boldsymbol{I}_3)\,\mathrm{vec}\,(\boldsymbol{T}_a)||_2^2. \end{aligned} \tag{19}$$

From the definition of the operator norm ((B2) of Appendix B), we have that

$$\begin{aligned}||\boldsymbol{B}_{eq}(\Delta\boldsymbol{F}_a^T \otimes \boldsymbol{I}_3)\,\mathrm{vec}\,(\boldsymbol{T}_a)||_2 \le\ &||\boldsymbol{B}_{eq}(\Delta\boldsymbol{F}_a^T \otimes \boldsymbol{I}_3)||_2 \\ &\cdot ||\,\mathrm{vec}\,(\boldsymbol{T}_a)||_2. \end{aligned} \tag{20}$$

Thus, letting $K_t = ||\boldsymbol{B}_{eq}(\Delta\boldsymbol{F}_a^T \otimes \boldsymbol{I}_3)\,||_2^2$, and dropping the constant 2, we define an error metric ε_a so that $\Delta E_{emis}^2 \le \varepsilon_a$.

$$\begin{aligned}||\varepsilon_a = &\boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{A}) - \boldsymbol{B}_{eq}(\boldsymbol{F}_a^T \otimes \boldsymbol{I}_3)\,\mathrm{vec}\,(\boldsymbol{T}_a)||_2^2 \\ &+ K_t^2||\,\mathrm{vec}\,(\boldsymbol{T}_a)||_2^2. \end{aligned} \tag{21}$$

The corresponding error metrics for the reflective modes are

$$\begin{aligned}\varepsilon_{EE} = &||\boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{A}_{EE}) - \boldsymbol{B}_{eq}(\boldsymbol{F}_b^T \otimes \boldsymbol{I}_3)\,\mathrm{vec}\,(\boldsymbol{T}_{EE})||_2^2 \\ &+ K_t^2\,||\mathrm{vec}\,(\boldsymbol{T}_{EE})||_2^2, \end{aligned} \tag{22}$$

and

$$\begin{aligned}\varepsilon_{D65} = &||\boldsymbol{B}_{eq}\,\mathrm{vec}\,(\boldsymbol{A}_{D65}) - \boldsymbol{B}_{eq}(\boldsymbol{F}_b^T \otimes \boldsymbol{I}_3)\,\mathrm{vec}\,(\boldsymbol{T}_{D65})||_2^2 \\ &+ K_t^2\,||\mathrm{vec}\,(\boldsymbol{T}_{D65})||_2^2. \end{aligned} \tag{23}$$

C. The Overall Cost Function

In this section, we combine the error metrics of the previous sections to formulate a scalar cost function that we symbolize as $h(\cdot)$. This scalar cost is a function of the four error terms ε_a, ε_{EE}, ε_{D65}, and ε_s, which, in turn, are functions of the four unknown filters $\boldsymbol{f}_1$–$\boldsymbol{f}_4$, the unknown coefficient matrices $\boldsymbol{T}_a$, $\boldsymbol{T}_{EE}$, and $\boldsymbol{T}_{D65}$, and the scale factors K_s and K_t.

When combining the error terms, we must keep in mind that the device will operate in only a single mode at a time. We are interested in the worst-case performance of the three modes. In other words, we want max $(\varepsilon_a, \varepsilon_{EE}, \varepsilon_{D65})$. However, it is advantageous to have an error metric that is a continuous function of its arguments. To allow for this, we use a high index p-norm (Appendix B) to approximate the maximum. The combined scalar cost function is then defined as

$$h(\boldsymbol{F}, \boldsymbol{T}_a, \boldsymbol{T}_{EE}, \boldsymbol{T}_{D65}, K_s, K_t) = \varepsilon_p + \varepsilon_s, \tag{24}$$

$$\varepsilon_p = ||[\varepsilon_a, \varepsilon_{EE}, \varepsilon_{D65}]||_{20}. \tag{25}$$

The terms ε_s, ε_a, ε_{EE}, and ε_{D65} are defined in (16) and (21)–(23), respectively.

IV. Optimization

In the following several sections, we present a method for solving for a set of filters that minimizes the cost function $h(\cdot)$. We show that the minimization can be thought of as a combination of two problems—one that has a closed-form solution and one that does not. On the highest level, the solution is an iterative optimization based on gradient descent techniques [15]. The variables in this iterative optimization are the four filter vectors. At each step of the iteration, we take advantage of the closed-form portion of the solution to solve for optimal coefficient matrices.

A. Separation of Problem into Two Parts

The variables of $h(\cdot)$ can be grouped into three categories:

1) the unknown filters $\boldsymbol{F} = [\boldsymbol{f}_1|\boldsymbol{f}_2|\boldsymbol{f}_3|\boldsymbol{f}_4]$
2) the coefficient matrices $\boldsymbol{T}_a$, $\boldsymbol{T}_{EE}$, and $\boldsymbol{T}_{D65}$
3) the scaling constants K_s and K_t.

The scaling constants are used to adjust the relative importance of the components of the error metric. Large values of K_s tend to make the solution filters smooth. Large values of K_t tend to make the solution robust. On the other hand, large values of these parameters tend to degrade the solution with respect to color shifts. We will show the effect of varying these parameters in a later section.

The first two categories of the aforementioned list represent the true unknowns of our problem. The first category—the unknown filter function matrix $\boldsymbol{F}$—occurs in each of the separate error terms ε_s, ε_a, ε_{EE}, and ε_{D65}. This means that we must minimize these error terms simultaneously. Conversely, the items of the second category, the coefficient matrices $\boldsymbol{T}_a$, $\boldsymbol{T}_{EE}$, and $\boldsymbol{T}_{D65}$ occur alone in ε_a, ε_{EE}, and ε_{D65}, respectively. Therefore, we can minimize each of these error expressions independently with respect to the coefficient matrices. This independence is the key that allows us to solve for the coefficient matrices in closed form.

B. Closed-Form Portion of the Solution

In this section, we assume that we are dealing with a fixed filter matrix $\boldsymbol{F}$. We concentrate on solving for the unknowns $\boldsymbol{T}_a$, $\boldsymbol{T}_{EE}$, and $\boldsymbol{T}_{D65}$. We will give a detailed explanation for obtaining the optimal $\boldsymbol{T}_a$. The other two coefficient matrices are obtained similarly.

The only error term containing $\boldsymbol{T}_a$ is ε_a (21). Letting $\boldsymbol{I}_{12}$ be the (12×12) identity matrix, (21) can be rewritten as

$$\varepsilon_a = \left\| \begin{bmatrix} \boldsymbol{B}_{eq} \operatorname{vec}(\boldsymbol{A}) \\ 0 \end{bmatrix} - \begin{bmatrix} \boldsymbol{B}_{eq}(\boldsymbol{F}_a^T \otimes \boldsymbol{I}_3) \\ K_t \boldsymbol{I}_{12} \end{bmatrix} \operatorname{vec}(\boldsymbol{T}_a) \right\|_2^2. \tag{26}$$

Equation (26) is of the form $||\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}||_2^2$, where $\boldsymbol{x}$ is the unknown, and $\boldsymbol{A}$ has independent columns. The minimum error norm solution to (26) is given by (see [17])

$$\operatorname{vec}(\boldsymbol{T}_a) = [(\boldsymbol{F}_a^T \otimes \boldsymbol{I}_3)^T \boldsymbol{B}_{eq}^T \boldsymbol{B}_{eq} (\boldsymbol{F}_a^T \otimes \boldsymbol{I}_3) + K_t^2 \boldsymbol{I}_{12}]^{-1} \cdot (\boldsymbol{F}_a^T \otimes \boldsymbol{I}_3)^T \boldsymbol{B}_{eq}^T \boldsymbol{B}_{eq} \operatorname{vec}(\boldsymbol{A}). \tag{27}$$

Analogous solutions can be found for (22) and (23).

C. Iterative Portion of the Solution-Starting Sets

Since, for a fixed $\boldsymbol{F}$, we know how to solve for the coefficient matrices in closed form, we can represent our high-level error metric as a function of only three parameters.

$$h^*(\boldsymbol{F}, K_s, K_t) = \min_{(\boldsymbol{T}_a, \boldsymbol{T}_{EE}, \boldsymbol{T}_{D65})} h(\boldsymbol{F}, \boldsymbol{T}_a, \boldsymbol{T}_{EE}, \boldsymbol{T}_{D65}, K_s, K_t) \tag{28}$$

In the iterative portion of the solution, we fix K_s and K_t and optimize over $\boldsymbol{F}$. At each step of the iterative process, we fix $\boldsymbol{F}$ and solve the closed-form portion of the problem to determine the value of $h^*(\boldsymbol{F}, K_s, K_t)$. We used the commercial routine ("constr") from the MATLAB® optimization toolbox [16] to carry out the optimization. We constrained the solution space to filter vectors having elements in the range 0–1, and we constrained the peak transmittance (largest element) of each filter vector to be approximately 1.

For fixed K_s and K_t, the function $h^*(\boldsymbol{F}, K_s, K_t)$ is continuous so that the constrained minimization routine has no trouble converging. However, there is no guarantee that it converges to the global minimum. The value of $\boldsymbol{F}$ at the convergence point depends on the starting guess for $\boldsymbol{F}$.

For each choice of K_s and K_t, we used 25 different starting sets ($\boldsymbol{F}$ matrices) and always picked the best solution of the 25. The filter vectors used in these starting sets were generated randomly using some constraints. The reason for the random choice was to avoid biases due to preconceived notions of "good filter functions." The reason for the constraints was to provide starting sets that will converge in a reasonable amount of time. The choice of constraints was arbitrary on our part. Other constraints would probably work as well. Fig. 9

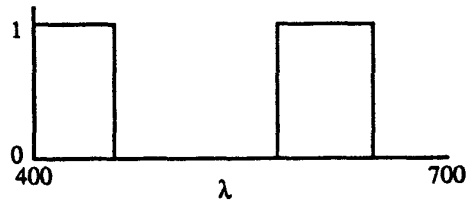


Fig. 9. Example of a starting filter. There are four different starting filters in any starting set.

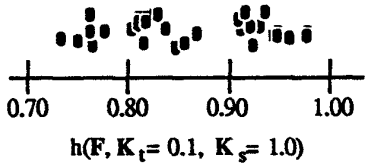


Fig. 10. Scatter plot of 25 final values of the cost function $h(F, K_t, K_s)$ for a typical run of 25 starting filter sets. (Vertical separation is used to help distinguish overlapping dots.)

is an example of a starting filter. The constraints used in its generation follow.

1) Each starting filter vector takes on values of only 0 or 1 for each of its 31 elements.
2) The first element is chosen with equal probability to be 0 or 1.
3) The number of transitions (0–1 or 1–0) is a uniform random variable on the set $\{1, 2, 3, 4\}$.
4) The positions of the transitions are uniform random variables subject to the constraint that there can be at most one transition in any consecutive group of three elements.

V. Experimental Results

A single run of 25 starting sets resulted in as many as 25 unique solutions. However, in any given run, the 25 final values of the cost function $h^*(\boldsymbol{F}, K_s, K_t)$ generally clustered around only a few values (Fig. 10). Thus, even though there were several local minima, they could all be categorized performance-wise into a small number of groups. This gives us confidence that the best performing group of any given run is near the global minimum. In the sections that follow, we report values of the best of 25 solution sets for any particular run.

We report errors in terms of color shifts in $L^*a^*b^*$ space. Recall that (9), (13), and (14) represent the noise-free mean square $L^*a^*b^*$ errors over our ensemble of spectra for the emissive, reflective/EE and reflective/$D65$ modes, respectively. We are interested in the maximum of these three errors; therefore, we define the following error term for plots of filter set performance.

$$\Delta E_{rms} = \max(\Delta E_{emis}, \Delta E_{EE}, \Delta E_{D65}). \quad (29)$$

Equation (29) is the worst-case root-mean square color shift error for the ensemble of spectral stimuli $\{n_k\}$. It is an upper bound on the average ΔE of the ensemble and a lower bound on the maximum ΔE (see (B3) and (B4) of Appendix B).

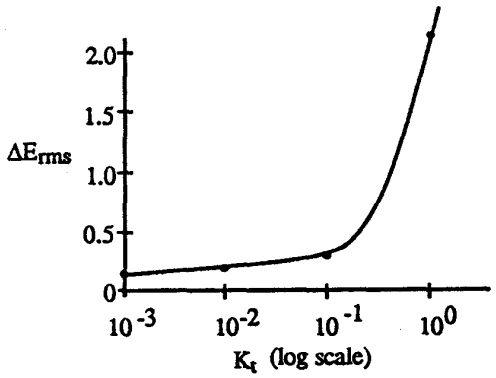


Fig. 11. Error performance for filter sets designed with varying values of K_t. K_s is fixed a 1.0.

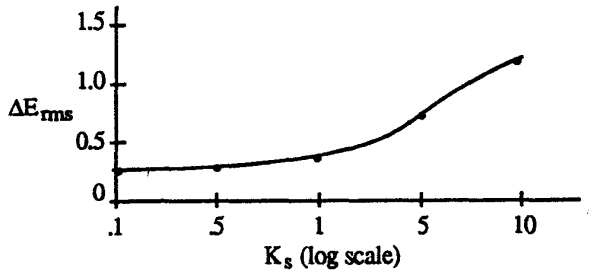


Fig. 12. Error performance for filter sets designed with varying values of K_s. K_t is fixed at 0.1.

A. Effect of Varying the Parameters K_t and K_s

K_t is the robustness parameter of $h^*(\boldsymbol{F}, K_s, K_t)$. Large values of K_t correspond to greater design robustness in the presence of device noise. Fig. 11 shows the trend in ΔE_{rms} (29) for filter sets designed with varying values of K_t. The figure is based on optimization runs using a fixed K_s of 1.0. The errors stay within a narrow range for small values of K_t and increase exponentially when K_t becomes greater than about 0.1. At about this value, the robustness error term is approximately equal to the color shift error term.

K_s is the smoothness parameter in the optimization cost function. Large values of K_s correspond to smoother solution filters. Fig. 12 shows the trend in ΔE_{rms} (29) for filter sets designed with varying values of K_s. The figure is based on a fixed K_t of 0.1. The effect of varying K_s is smaller than that of K_t. The value of K_s at which the roughness cost becomes about equal to the color shift and noise sensitivity costs is about 1. As a practical matter, we obtained fairly smooth solution sets for all nonzero values of K_s.

Fig. 13 shows our best solution set for optimizations using $K_s = 1$ and $K_t = 0.1$.

B. Effect of Device Noise

We simulate device noise by adding quantities δ_{ik} to the (i, k) elements of $\boldsymbol{F}_a$ and $\boldsymbol{F}_b$ of (9), (13), and (14). Each δ_{ik} is a random variable uniformly distributed between the limits $[-\Omega, \Omega]$. The system tolerance Ω accounts for manufacturing, long-term, and short-term variations in the LED spectral

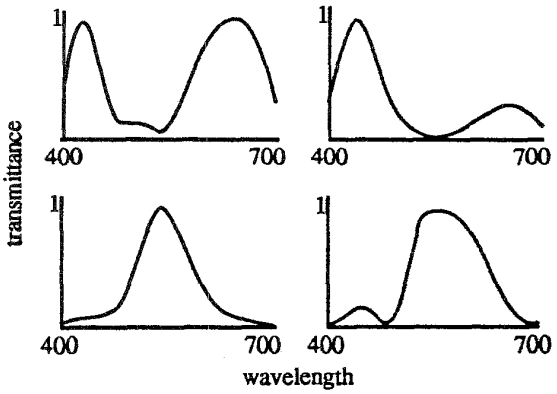


Fig. 13. Solution set obtained using the design parameters $K_t = 0.1$ and $K_s = 1$.

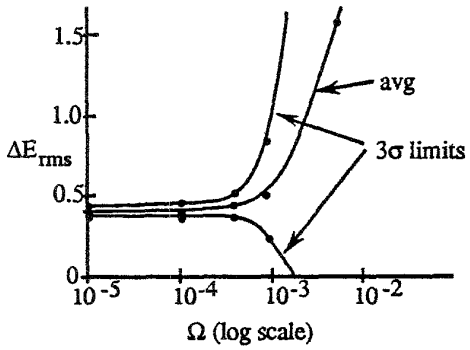


Fig. 14. Effect of system tolerance Ω on mean-squared error for the filter set shown in Fig. 13. Average and 3σ performance were calculated from 100 simulations of random noise.

outputs, the filters, and the detector sensitivities. We use ΔE_{rms} (29) to plot errors since it is a function of (9), (13), and (14), which are the terms affected by the noise.

For the simulations, we used the filter set shown in Fig. 13 and the corresponding optimal coefficient matrices $\boldsymbol{T}_a$, $\boldsymbol{T}_{EE}$, and $\boldsymbol{T}_{D65}$. We varied the system tolerance over several orders of magnitude. For each order of magnitude, we ran 100 simulations and recorded the average ΔE_{rms} and its variation. In Fig. 14, we plot ΔE_{rms} as a function of system tolerance showing its average value and 3σ limits.

C. Average and Worst-Case Performance in True $L^*a^*b^*$ Space

As a validation check, we calculated the true $L^*a^*b^*$ ΔE errors in device outputs for the 752 spectral samples used in our model. We used the filter set shown in Fig. 13 and the corresponding optimal coefficient matrices. We did not simulate any noise in these calculations. However, since the filter set was designed to tolerate noise, the results should hold for a system tolerance of up to 0.0005 (see the previous section). Note that in the noise-free case, the robustness criteria force a more accurate solution for the reflective modes than for the emissive mode. This is because the reflective mode has five times as many potential noise channels. Thus, the worst-case noise-free performance generally occurs in the emissive mode.

TABLE I
ERROR PERFORMANCE IN TRUE $L^*a^*b^*$ SPACE FOR THE FILTER SET SHOWN IN FIG. 13 AVERAGED OVER 752 SPECTRAL SAMPLES

mode	avg L*a*b* ΔE	max L*a*b* ΔE	% ΔE's > 1
emissive	0.27	1.56	1.5
reflective / EE	0.18	1.00	0.2
reflective / D65	0.15	0.91	0

This is confirmed in Table I, which gives the $L^*a^*b^*\Delta E$ errors for the various modes.

In Fig. 15, we plot the error surface for the emissive mode. This plot was obtained by projecting all 752 ΔE values onto a single a^*b^* plane (i.e., ignoring the L^* coordinate) and averaging over 10×10 cells on the a^*b^* plane. While this figure shows that the errors averaged with respect to L^* tend to be largest at higher chromaticities, the maximum errors with respect to L^* are more uniformly distributed over the a^*b^* plane.

D. Approximation of Ideal Sensing Channels

One way in which the filter design problem can be viewed is as approximating the ideal sensing channels of Figs. 1 and 5, by linear combinations of the device channels of Figs. 2 and 6. The ideal channels are the XYZ color matching functions (rows of the matrices $\boldsymbol{A}$ or $\boldsymbol{A}_{EE}$) or the color matching functions multiplied by the $D65$ illuminant (rows of $\boldsymbol{A}_{D65}$). Fig. 16 shows the ideal functions for the three cases (solid lines) with the channel approximations (dashed lines) resulting from the filter set of Fig. 13.

VI. CONCLUSIONS

Our results show that it is possible to design a filter set giving excellent performance in terms of both $L^*a^*b^*$ error and robustness to noise. In our experiments, we minimized the $L^*a^*b^*$ error over a set of 752 spectral stimuli. Our technique would not be any more complex if we had used a larger or smaller set of spectral stimuli. Moreover, the technique is applicable even if we have no ensemble of sample spectra. In that case, we simply set $\boldsymbol{B}_{eq}$ of (9), (13), and (14) equal to the identity matrix. This forces the optimization to be over the error norms of the spectral approximations of the color matching functions.

Our regularization technique resulted in a set of smooth filters (Fig. 13) that appear to be reasonable to manufacture. Except for 11 of the 752 points, all color shift errors were less than 1 $L^*a^*b^*\Delta E$ unit, which is an imperceptible shift to most human observers.

APPENDIX A
TABLE OF SYMBOLS

The following symbols are used throughout this paper.

$\boldsymbol{A}$ (3×31) matrix whose rows are $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$.
$\boldsymbol{d}$ Spectral response of the detector.

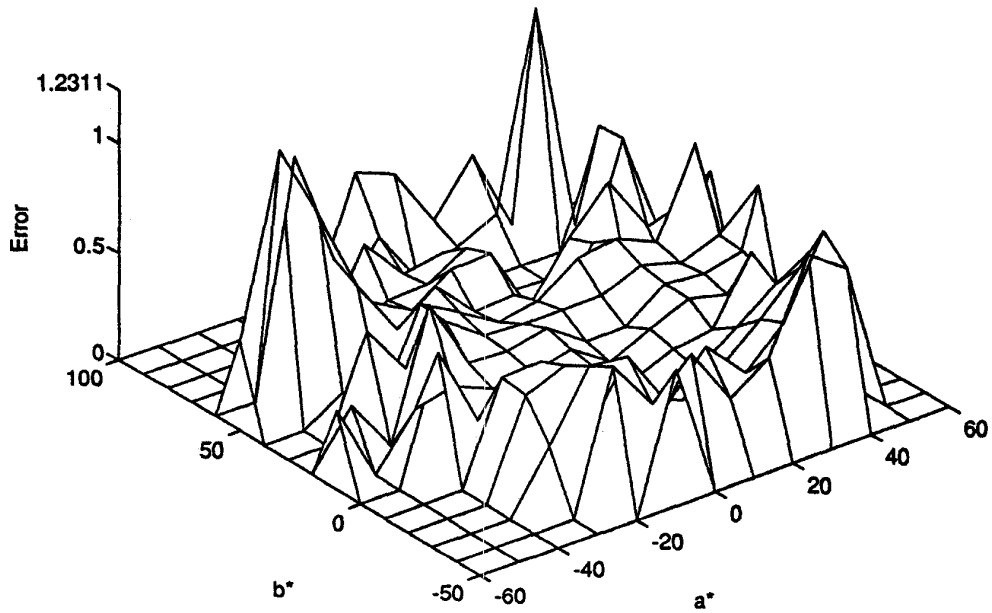


Fig. 15. Emissive mode $L^*a^*b^*$ error surface for the filter set of Fig. 13 (noise-free case). All points in $L^*a^*b^*$ space are projected onto the a^*b^* chromaticity plane. The height of each grid point above the a^*b^* plane is the ΔE value obtained by averaging within a 10 × 10 cell centered at that grid point.

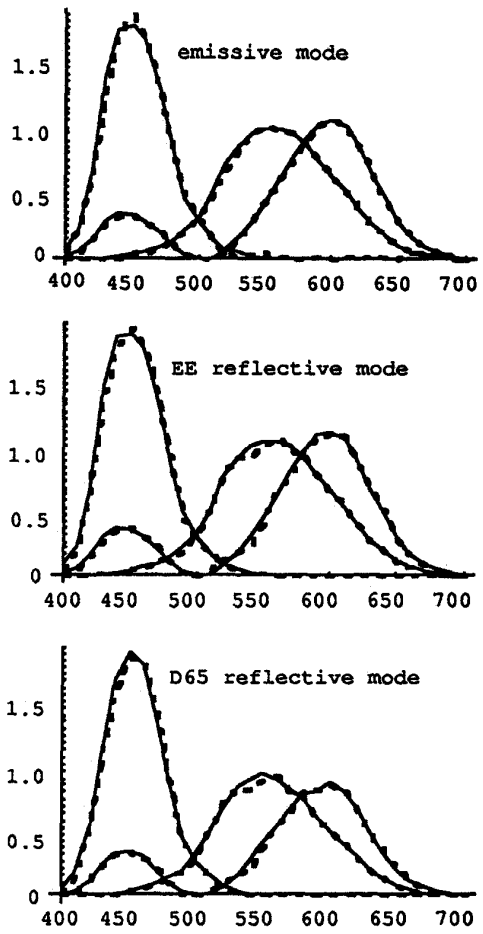


Fig. 16. Approximations of ideal sensing channels (solid lines) by linear combinations of the device channels (dashed lines) for the 3 modes of operation. Device channels were obtained usign the filter set of Fig. 13.

ΔE	Error norm in L^*a^*b space.
$\boldsymbol{f}$	Color filter transmittance function.
$\boldsymbol{i}$	Standard illuminant (EE or $D65$).
$\boldsymbol{J}$	Derivative of the transformation from XYZ to $L^*a^*b^*$ tristimulus coordinates.
$\boldsymbol{l}$	Spectral distribution a single light emitting diode.
$\boldsymbol{n}$	Spectral stimulus.
$\boldsymbol{r}$	Surface reflectance.
$\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$	CIE 2° XYZ color matching functions.
$\boldsymbol{t}$	an XYZ tristimulus vector $[X, Y, Z]^T$.
vec	See Appendix C.
$\otimes$	See Appendix C.

APPENDIX B
VECTOR AND MATRIX NORMS AND INEQUALITIES

Here, we review some definitions and properties of vector and matrix norms used in this paper.

A comprehensive review of normed spaces is given in [15]. Additional properties of vector and matrix norms can be found in [17].

The p-norm of a vector $\boldsymbol{x}$ is defined as

$$||\boldsymbol{x}||_p = \left(\sum |x_k|^p\right)^{(1/p)}. \tag{B1}$$

The (operator) p-norm of a matrix is defined as

$$||\boldsymbol{A}||_p = \sup_{\boldsymbol{x}\neq 0} \frac{||\boldsymbol{A}\boldsymbol{x}||_p}{||\boldsymbol{x}||_p}. \tag{B2}$$

We can make use of the Cauchy–Schwartz inequality [15] to show that the average of m elements of a vectors $\boldsymbol{x}$ is less than the square root of the average of the squares of the elements.

Let $\boldsymbol{l}$ be an m element vector of all 1's.

$$\begin{aligned}\frac{1}{m}\sum_k x_k &\le \frac{1}{m}\sum_k |x_k \cdot 1| \\ &\le \frac{1}{m}||\boldsymbol{x}||_2\,||\boldsymbol{l}||_2 \\ &= \left(\frac{1}{m}\sum_k x_k^2\right)^{0.5}. \end{aligned} \tag{B3}$$

We also have a relationship between the average of the squares of the elements of $\boldsymbol{x}$ and the element with the greatest magnitude. Let M be $\max(|x_k|)$:

$$\begin{aligned}\frac{1}{m}\sum_k x_k^2 &\le \frac{1}{m}\sum_k M^2 \\ &= M^2. \end{aligned} \tag{B4}$$

For any two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, we have the following property [15]. It can be proven by expressing the 2-norms as inner products (i.e., $||\boldsymbol{a}||_2^2 = \boldsymbol{a}^T\boldsymbol{a}$) and expanding terms.

$$||\boldsymbol{a}+\boldsymbol{b}||_2^2 + ||\boldsymbol{a}-\boldsymbol{b}||_2^2 = 2||\boldsymbol{a}||_2^2 + 2||\boldsymbol{b}||_2^2. \tag{B5}$$

Appendix C
The vec and Kronecker Product Operators

The theory presented in this Appendix is adopted from ch. 2 of [18].

For any matrices, $\boldsymbol{A}$ and $\boldsymbol{B}$, their Kronecker product is defined as follows:

$$\boldsymbol{A}\otimes\boldsymbol{B} \equiv \begin{bmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & & & \\ \vdots & & & \vdots \\ a_{m1}\boldsymbol{B} & & \cdots & a_{mn}\boldsymbol{B} \end{bmatrix}. \tag{C1}$$

For a scalar α and for vectors $\boldsymbol{b}$ and $\boldsymbol{c}$, the following are readily apparent:

$$\begin{aligned}\alpha\boldsymbol{A} &= \alpha\otimes\boldsymbol{A} \\ &= \boldsymbol{A}\otimes\alpha \end{aligned} \tag{C2}$$

$$\boldsymbol{a}\boldsymbol{b}^T = \boldsymbol{a}\otimes\boldsymbol{b}. \tag{C3}$$

Whenever $\boldsymbol{AC}$ and $\boldsymbol{BD}$ exist, we have the following:

$$\begin{aligned}&(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D}) \\ &= \begin{bmatrix} a_{11}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ \vdots & & \vdots \\ a_{m1}\boldsymbol{B} & \cdots & a_{mn}\boldsymbol{B} \end{bmatrix} \cdot \begin{bmatrix} c_{11}\boldsymbol{D} & \cdots & c_{1n}\boldsymbol{D} \\ \vdots & & \vdots \\ c_{m1}\boldsymbol{D} & \cdots & c_{mn}\boldsymbol{D} \end{bmatrix} \\ &= \begin{bmatrix} \sum_k a_{1k}c_{k1}\boldsymbol{BD} & \cdots & \sum_k a_{1k}c_{km}\boldsymbol{BD} \\ \vdots & & \vdots \\ \sum_k a_{mk}c_{k1}\boldsymbol{BD} & \cdots & \sum_k a_{mk}c_{km}\boldsymbol{BD} \end{bmatrix} \\ &= \boldsymbol{AC}\otimes\boldsymbol{BD}. \end{aligned} \tag{C4}$$

The vec operator stacks the columns $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots$ of the matrix $\boldsymbol{A}$.

$$\text{vec}\,(\boldsymbol{A}) = \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_n \end{bmatrix}. \tag{C5}$$

We use the following property of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ in proving (C7):

$$\begin{aligned}\text{vec}(\boldsymbol{a}\boldsymbol{b}^T) &= \text{vec}\begin{bmatrix} a_1b_1 & \cdots & a_1b_n \\ \vdots & & \vdots \\ a_mb_1 & \cdots & a_mb_n \end{bmatrix} \\ &= \boldsymbol{b}\otimes\boldsymbol{a}. \end{aligned} \tag{C6}$$

The next equality is a key relationship between the vec operator and the Kronecker product. Let $\boldsymbol{e}_i$ be the ith column of an identity matrix. Then

$$\begin{aligned}\text{vec}\,(\boldsymbol{ABC}) &= \text{vec}\left[\boldsymbol{A}\left(\sum \boldsymbol{b}_i\boldsymbol{e}_i^T\right)\boldsymbol{C}\right] \\ &= \sum \text{vec}\left[(\boldsymbol{A}\boldsymbol{b}_i)\,(\boldsymbol{C}^T\boldsymbol{e}_i)^T\right] \\ &= \sum (\boldsymbol{C}^T\boldsymbol{e}_i)\otimes(\boldsymbol{A}\boldsymbol{b}_i) \qquad \text{by (C6)} \\ &= \sum (\boldsymbol{C}^T\otimes\boldsymbol{A})(\boldsymbol{e}_i\otimes\boldsymbol{b}_i) \qquad \text{by (C4)} \\ &= (\boldsymbol{C}^T\otimes\boldsymbol{A})\sum \text{vec}\,(\boldsymbol{b}_i\boldsymbol{e}_i^T) \qquad \text{by (C6)} \\ &= (\boldsymbol{C}^T\otimes\boldsymbol{A})\,\text{vec}\,(\boldsymbol{B}). \end{aligned} \tag{C7}$$

References

[1] G. Wyszecki and W. S. Stiles, *Color Science*. New York: Wiley, 1982.
[2] J. B. Cohen, "Color and color mixture: Scalar and vector fundamentals," *Color Res. Appl.*, vol. 13, pp. 5–39, 1988.
[3] H. J. Trussell, "Applications of set theoretic methods to color systems," *Color Res. Appl.*, vol. 16, pp. 31–41, 1991.
[4] W. E. R. Davies and G. Wyszecki, "Physical approximation of color mixture functions," *J. Opt. Soc. Amer.*, vol. 52, pp. 679–685, 1962.
[5] R. V. Kollarits and D. C. Gibbon, "Improving the color fidelity of cameras for advanced television systems," in *SPIE Proc.*, vol. 1656, 1992, pp. 19–29.
[6] K. Engelhardt and P. Seitz, "Optimum color filters for CCD digital cameras," *Appl. Opt.*, vol. 32, pp. 3015–3023, 1993.
[7] M. J. Vrhel and H. J. Trussell, "Filter considerations in color correction," *IEEE Trans. Image Processing*, vol. 3, pp. 147–161, 1994.
[8] ——, "Color filter selection for color correction in the presence of noise," in *Proc. IEEE. Conf. Acoust., Speech, Signal Processing*, vol. 5, 1993, pp. v313–v316.
[9] P. L. Vora and H. J. Trussell, "A mathematical method for designing a set of color scanning filters," *Color Hard Copy Graphic Arts II, SPIE*, vol. 1912, pp. 322–332, 1993.
[10] ——, "Measure of goodness of a set of color-scanning filters," *J. Opt. Soc. Amer. A*, vol. 10, pp. 1499–1508, 1993.
[11] V. Falletti, A. Premoli, and M. L. Rastello, "Optical glass filters: An algorithm for optimum design," *Appl. Opt.*, vol. 9, pp. 1396–1402, 1970.
[12] Y. Liu, R. S. Berns, and Y. Shu, "Optimization algorithm for designing colored glass filters to simulate CIE illuminant D65," *Color Res. Appl.*, vol. 16, pp. 89–96, 1991.
[13] K. L. Kelly, K S. Gibson, and D. Nickerson, "Tristmulus specification of the munsell book of color from spectrophotometric measurements," *J. Opt. Soc. Amer.*, vol. 33, pp. 355–376, 1943.
[14] M. J. Vrhel, R. Gershon, and L. S. Iwan, "Measurement and analysis of object reflectance spectra," *Color Res. Appl.*, vol. 19, pp. 4–9, 1994.

[15] D. G. Luenberger, *Optimization by Vector Space Methods.* New York: Wiley, 1969.
[16] A. Grace, *Matlab Optimization Toolbox User's Guide.* Boston, MA: The Math Works, 1992.
[17] G. H. Golub and C. F. VanLoan, *Matrix Computations.* Baltimore, MD: Johns Hopkins University Press, 1989.
[18] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus.* New York: Wiley, 1988.



Mark Wolski received the B.S. degree in 1978 from GMI Institute, the M.S. degree from the Massachusetts Institute of Technology, Cambridge, in 1980, and the Ph.D. degree from Purdue University, West Lafayette, IN, in 1994, all in electrical engineering.

He has served in various capacities for General Motors since 1973 and is currently doing work on Intelligent Highway Systems at the GM Research Laboratories, Warren, MI.

Charles A. Bouman (M'89), for photograph and biography, please see p. 492.



Jan P. Allebach (F'91) received the B.S. degree from the University of Delaware, Newark, in 1972 and the M.S. and Ph.D. degrees from Princeton University, Princeton, NJ, in 1975 and 1976, respectively, all in electrical engineering.

He was with the Department of Electrical Engineering at the University of Delaware from 1976 to 1983. Since 1983, he has been with the School of Electrical Engineering at Purdue University, West Lafayette, IN, where he holds the rank of Professor. His research has primarily been in the areas of printing and display of images, scanning and sampling of multidimensional signals, and synthesis of digital diffractive elements. He has been a visiting summer faculty member at IBM Watson Research Center in Yorktown Heights, NY, Sandia Laboratories, Livermore, CA, and Hewlett-Packard Labs, Palo Alto, CA. In addition, he has consulted extensively for industry and government laboratories.

From 1972 to 1975, Prof. Allebach was a National Science Foundation Graduate Fellow. He is a member of the Optical Society of America and the IS&T. In 1987, he received the Senior Award from the IEEE Signal Processing Society for his paper on time-sequential sampling published in the February 1984 issue of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. During 1991–1992, he served as the IS&T Visiting Lecturer. During 1994–1995 he was a Distinguished Lecturer for the IEEE Signal Processing Society. He chaired the Image and Multidimensional Signal Processing Technical Committee of the IEEE Signal Processing Society from 1991 to 1992. He is a past Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and currently serves on the Editorial Board for the *Journal of Electronic Imaging*. Presently, he is Chapters Vice-President of IS&T and Secretary for the IEEE Signal Processing Society.



Eric Walowit graduated from the Image Science Program at the Rochester Institute of Technology, Rochester, NY, where he concentrated in color science. His thesis was on nonlinear and statistical methods of computer color matching.

He is the President of Color Savvy Systems, Inc., Springboro, OH. He has worked in the color imaging field for more than 10 years and has authored over 30 patents, publications, and presentations.

Mr. Walowit is a member of IS&T, TAGA, and ISCC.