Diagonal versus affine transformations for color correction

Brian V. Funt and Benjamin C. Lewis

School of Computing Science, Simon Fraser University, Vancouver V5A 1S6, Canada

Received October 21, 1999; accepted June 6, 2000; revised manuscript received June 30, 2000

Standard methods for color correction involve the use of a diagonal-matrix transformation. Zaidi proposes the use of a two-parameter affine model; we show that this offers no improvement in terms of accuracy over the diagonal model, especially if a sharpening transformation is also used. © 2000 Optical Society of America [S0740-3232(00)01411-3]

OCIS codes: 150.0150, 100.0100, 330.0330, 330.1720, 330.4060.

1. INTRODUCTION

In Ref. 1, Zaidi proposes the use of a two-parameter affine model as a vehicle for transforming sensor outputs from a scene into illumination-independent color descriptors. The idea of mapping raw cone signals to color descriptors is motivated in part by the problem of object recognition in varying illuminations. An algorithm that uses the color of an object as a recognition aid is not useful in different illuminations unless these descriptors can be found with some degree of accuracy, since the raw sensor data may be very different in these various conditions.

In Zaidi's model, the sensor responses are combined into a luminance channel and two color channels in MacLeod–Boynton[2] chromaticity space. These are referred to as the *rg* and *yv* coordinates. To account for changes in illumination, these coordinates are transformed into illumination-independent color descriptors with an affine transformation that involves a scaling of one coordinate and a translation of the other.

This is an interesting approach, which we will examine to see whether it offers any improvements over the performance of the more standard diagonal-matrix transform (DMT). Zaidi's model accounts for illumination change by a two-parameter affine transformation. Using two parameters, instead of the three parameters implicit in a typical von Kries scaling of the cone signals, is an interesting approach, but the affine model is not the only two-parameter model available. We consider as an alternative a two-parameter diagonal model and find it to model illumination change more accurately than Zaidi's affine model, especially when used in conjunction with a technique known as spectral sharpening.

A. Color Constancy

Color constancy can be defined as the operation of mapping responses from surfaces under an arbitrary illuminant to illumination-independent color descriptors. These descriptors are generally the response vectors of the sensors, such as the cones, under a canonical illumination.

Most existing color-constancy algorithms attempt to determine chromaticity information about the illumination in an input scene for which the only known information is the sensor responses for each point (or color patch) in that scene. The problem of finding this chromaticity information will not be dealt with here. Instead, we will be concerned with the task of transforming these sensor response vectors into descriptors once this information is already known. A standard method for doing this has been the use of a DMT; in other words, the response of each sensor is adjusted independently by a multiplicative scaling as suggested by von Kries (see Ref. 3).

B. Spectral Sharpening

Spectral sharpening[4] is the construction of new sensor sensitivities as linear combinations of the original sensors, such as the cones, by applying a 3×3 "sharpening" matrix. Many existing color-constancy algorithms use a DMT to map sensor response vectors to surface descriptors; however, the ability of a DMT to map these vectors to descriptors accurately is dependent on the spectral sensitivity functions of the sensors used. The aim behind sharpening is to produce sensors more optimal for use with a DMT. The process is called sharpening, from the intuition that narrow-band sensors will produce better results.

There are several ways to derive the sharpening matrix. In this paper we use the data-based method, where an optimal matrix T is found for use with a DMT over a data set of reflectances and test illuminants. This method is outlined in detail in Ref. 4 and is summarized as follows.

Let W^c be the $3 \times n$ matrix of descriptors from n surfaces observed under a canonical illuminant, and let W^e be the observations of the same surfaces under a test illuminant. We would like W^c and W^e to be approximately equal under a DMT:

$$W^e \approx DW^e. \tag{1}$$

The idea of sharpening is to improve the error by transforming W^c and W^e by a 3×3 sharpening matrix T:

$$TW^c \approx DTW^e. \tag{2}$$

In Ref. 4 it is shown that to optimize D and T in the least-squares sense, T is calculated as

0740-3232/2000/112108-05$15.00

$$T = U^{-1}, \tag{3}$$

where

$$UDU^{-1} = W^c[W^e]^+ \tag{4}$$

is the eigenvector decomposition of $W^c[W^e]^+$ and + denotes the Moore–Penrose inverse; i.e., $A^+ = A^t[AA^t]^{-1}$.

For data-based sharpening we choose a set of n surface reflectances and a pair of test and canonical illuminants i and j and then solve for T_{ij}. This value of T_{ij} is of course optimized for these particular illuminants; however, it often gives good results for other illuminant pairs as well.

In this communication we generate sharpening matrices for each pair of illuminants and then choose one of these sharpening matrices that gives good global results for all pairs of illuminants in our database.

2. ILLUMINANT DATABASE

Six different test illuminants were used for this experiment. Illuminants were chosen that exhibited a wide chromaticity variation from one another while still being "reasonable" in the sense that similar illuminants might be found in the real world. The six illuminants are

1. CIE Illuminant D50
2. CIE Illuminant D65
3. CIE Illuminant D250
4. CIE Illuminant A
5. Sylvania Cool White/No Filter
6. Room lighting (fluorescent)

Table 1. Illuminant 5 Spectral Information

λ	$E(\lambda)$	λ	$E(\lambda)$	λ	$E(\lambda)$	λ	$E(\lambda)$
380	0.2973	480	3.328	580	10.03	680	0.8206
384	0.3367	484	3.326	584	9.416	684	0.7390
388	0.4028	488	3.317	588	8.979	688	0.6723
392	0.4870	492	3.303	592	8.631	692	0.6223
396	0.6660	496	3.271	596	8.215	696	0.5742
400	1.518	500	3.219	600	7.747	700	0.5088
404	2.717	504	3.175	604	7.260	704	0.4840
408	2.130	508	3.135	608	6.654	708	0.4440
412	1.431	512	3.125	612	6.128	712	0.3731
416	1.466	516	3.167	616	5.593	716	0.3376
420	1.640	520	3.267	620	5.075	720	0.3088
424	1.815	524	3.443	624	4.542	724	0.3010
428	2.558	528	3.698	628	4.012	728	0.2863
432	6.414	532	4.016	632	3.571	732	0.2990
436	9.050	536	4.448	636	3.174	736	0.3292
440	5.088	540	5.580	640	2.823	740	0.2879
444	2.829	544	8.650	644	2.469	744	0.2708
448	2.753	548	8.566	648	2.182	748	0.3719
452	2.856	552	7.079	652	1.949	752	0.3478
456	2.965	556	7.333	656	1.743	756	0.2752
460	3.066	560	7.887	660	1.546	760	0.4224
464	3.159	564	8.404	664	1.354	764	0.4520
468	3.244	568	8.884	668	1.196	768	0.3520
472	3.307	572	9.453	672	1.063	772	0.3319
476	3.324	576	10.16	676	0.9408	776	0.2441
						780	0.1988

Table 2. Illuminant 6 Spectral Information

λ	$E(\lambda)$	λ	$E(\lambda)$	λ	$E(\lambda)$	λ	$E(\lambda)$
380	0.9173	480	3.049	580	8.540	680	2.662
384	1.143	484	2.850	584	13.46	684	2.362
388	1.924	488	2.632	588	34.01	688	2.240
392	2.405	492	2.585	592	31.83	692	2.468
396	2.415	496	2.913	596	17.14	696	3.333
400	3.896	500	3.436	600	9.539	700	3.450
404	4.920	504	5.972	604	6.762	704	2.748
408	3.812	508	8.045	608	6.629	708	2.053
412	2.825	512	5.181	612	8.857	712	1.722
416	2.546	516	2.814	616	11.78	716	1.716
420	2.579	520	2.619	620	11.66	720	1.612
424	2.993	524	2.464	624	8.875	724	1.372
428	3.356	528	2.561	628	7.359	728	1.213
432	5.099	532	3.598	632	6.121	732	1.249
436	6.165	536	4.262	636	4.949	736	1.349
440	4.492	540	5.683	640	4.458	740	1.400
444	3.024	544	11.54	644	4.179	744	1.291
448	2.455	548	10.89	648	4.223	748	1.327
452	2.245	552	5.772	652	4.713	752	1.349
456	2.128	556	3.638	656	5.440	756	1.358
460	1.943	560	3.226	660	5.660	760	1.701
464	2.108	564	6.363	664	5.992	764	2.857
468	3.351	568	11.17	668	8.206	768	3.098
472	5.528	572	10.52	672	7.313	772	2.197
476	4.715	576	10.50	676	3.913	776	1.907
						780	2.048

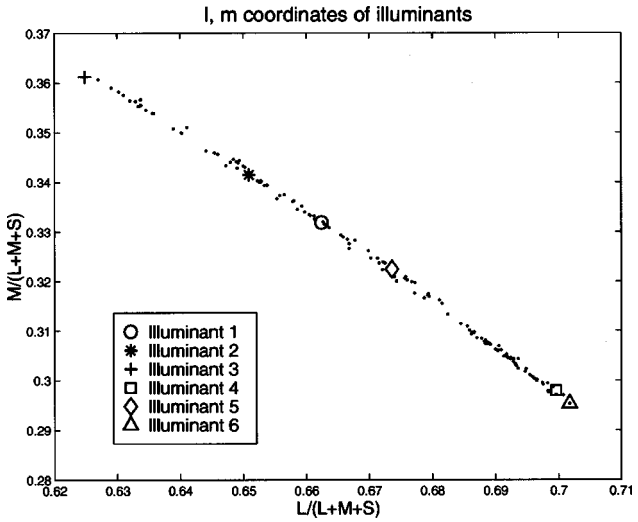

Fig. 1. Coordinates of the six illuminants plotted in (l, m) space. 140 other illuminants are plotted as small dots.

Spectral information for illuminants 5 and 6 is listed in Tables 1 and 2. The chromaticities of the illuminants are plotted in Figs. 1–3. In these figures are also shown 140 other illuminants, which are plotted as small dots and include various measurements taken both indoors and outdoors.

3. REFLECTANCE DATABASE

The database of reflectances used contains 1995 different entries, each consisting of 101 data points. These data

give the amount of light of a particular wavelength reflected from some reflectant surface, expressed as a fraction of the incident light. The surfaces are sampled at 4-nm intervals throughout the visible spectrum (from 380 to 780 nm). The database consists of 120 Dupont paint chips, 170 natural objects, the 24 Macbeth color-checker patches and 1269 Munsell chips,[5] the 350 surfaces in the Krinov data set,[6] and 57 additional surfaces measured by us.

4. EXPERIMENT

In this experiment we calculate the long-, middle-, and short-wavelength-sensitive (*L*-, *M*-, and *S*-) cone responses for each of the reflectances in our database under a given test illuminant. We then try to predict what the responses would be for the same set of reflectances under a different target illuminant, given the chromaticities of the two illuminants. We also calculate what the actual cone responses would be under the target illuminant and compare these values with our predictions in order to evaluate the method of prediction. The process is repeated each pair of illuminants in our database.

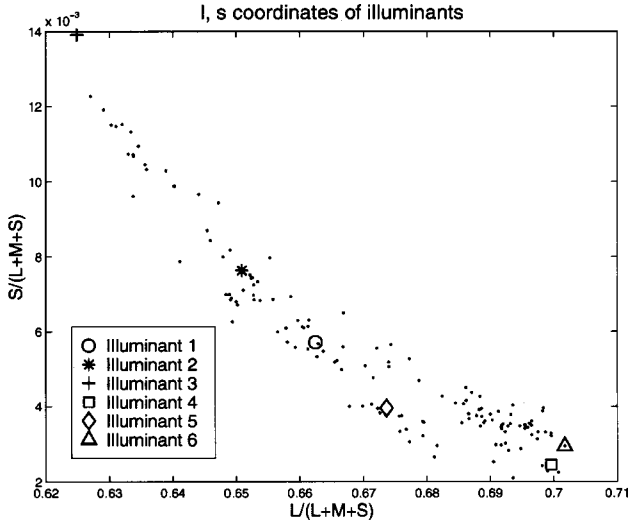

Fig. 2. Coordinates of the six illuminants plotted in (*l, s*) space. 140 other illuminants are plotted as small dots.

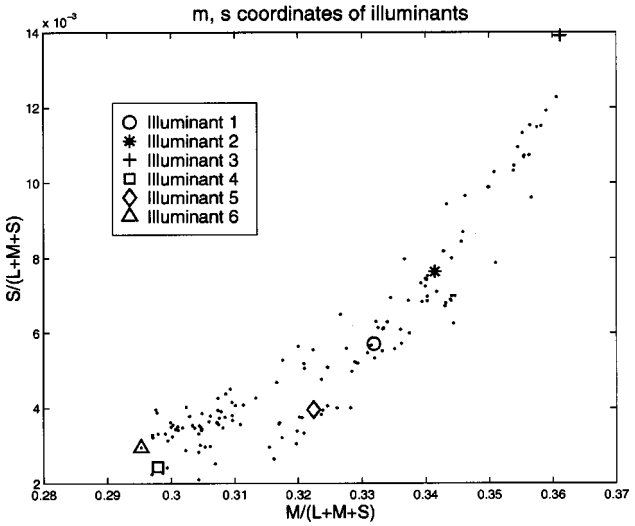

Fig. 3. Coordinates of the six illuminants plotted in (*m, s*) space. 140 other illuminants are plotted as small dots.

For each source illuminant, the sensor responses for the reflectant patches are given by

$$\mathbf{p}_i = \int E(\lambda)S(\lambda)R_i(\lambda)\mathrm{d}\lambda, \tag{5}$$

which is approximated by

$$\mathbf{p}_i = \sum E(\lambda)S(\lambda)R_i(\lambda), \tag{6}$$

where $E(\lambda)$, $S(\lambda)$, and $R_i(\lambda)$ are the illumination, the surface reflectance, and the sensor sensitivities, respectively, and the sum is taken over the visible spectrum, 380–780 nm, at 4-nm intervals. $\mathbf{p}$ is the three-dimensional response vector with one value for each cone; i.e.,

$$\mathbf{p} = [L \quad M \quad S]. \tag{7}$$

The sensor sensitivities used were the cone fundamentals derived by Vos and Walraven.[7]

The cone response vectors for the two illuminants are found by

$$\mathbf{p}_i = \sum E(\lambda)R_i(\lambda), \tag{8}$$

which is then used to find the chromaticities of the illuminants in (l, m) space, where l and m are given by

$$l = \frac{L}{L + M + S}, \tag{9}$$

$$m = \frac{M}{L + M + S}. \tag{10}$$

Using only this information, we predict the canonical chromaticities of the patches, using the three following methods.

A. Affine Model

In Zaidi's method, the chromaticities in the MacLeod–Boynton rg–yv space are calculated as follows:

$$s_a = \frac{S}{L + M}, \tag{11}$$

$$l_a = \frac{L}{L + M}. \tag{12}$$

To predict the chromaticities in the target illuminant, the s_a coordinate is scaled by σ, and the l_a coordinate is shifted by τ:

$$s_a' = s_a\sigma, \tag{13}$$

$$l_a' = l_a + \tau. \tag{14}$$

We wish to find σ and τ, which, when applied to each of the chromaticities of the reflectances under the source il-

luminant, yield a good approximation (s'_a, l'_a) to the chromaticities of the same reflectances under the target illuminant. We do not know the actual chromaticities in the target illuminant, but we do know that the chromaticity of a "perfect reflector" or "ideal white" in the target illuminant will be that of the illuminant itself. We calculate the values of σ and τ such that when applied to an ideal white under the source illuminant, the resultant rg–yv chromaticity is the same as that of the ideal white under the target illuminant:

$$\tau = \frac{L_{\text{target}}}{L_{\text{target}} + M_{\text{target}}} - \frac{L_{\text{source}}}{L_{\text{source}} + M_{\text{source}}}, \tag{15}$$

$$\sigma = \left(\frac{S_{\text{target}}}{L_{\text{target}} + M_{\text{target}}}\right) \Big/ \left(\frac{S_{\text{source}}}{L_{\text{source}} + M_{\text{source}}}\right). \tag{16}$$

To apply the transformation to the reflectant patches, we can calculate the equivalent transformation in (L, M, S) space:

$$\frac{S'}{L' + M'} = \frac{\sigma S}{L + M}, \tag{17}$$

$$\frac{L'}{L' + M'} = \frac{L}{L + M} + \tau. \tag{18}$$

Here we take $L + M + S = 1$; the values will be scaled later to account for luminance.

$$L' + M' + S' = L + M + S = 1; \tag{19}$$

solving Eqs. (17)–(19) for S', L', and M' gives

$$S' = \frac{\sigma S}{\sigma S + L + M}, \tag{20}$$

$$L' = \left(\frac{1}{L + M} + \tau\right)(1 - S'), \tag{21}$$

$$M' = 1 - L' - S'. \tag{22}$$

This transformation provides a mapping from an input image to an output image, where the input image can be seen as the cone responses from a collection of surface under a source illuminant and the output image is the set of predicted cone responses of the same surfaces under the target illuminant.

B. Diagonal Model

In contrast to the affine model, the DMT performs a multiplicative scaling on each of the raw *L,M,S* channels independently. In the standard DMT, each of the three sensors is scaled by a constant value:

$$\begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix}. \tag{23}$$

These values are calculated such that when applied to an ideal white under the source illuminant, the resulting chromaticity is the same as that of the ideal white under the target illuminant:

$$a = \frac{L'}{L} = \frac{l'}{l}, \tag{24}$$

$$b = \frac{M'}{M} = \frac{m'}{m}, \tag{25}$$

$$c = \frac{S'}{S} = \frac{1 - l' - m'}{1 - l - m}. \tag{26}$$

Note that although we have three coefficients, we are really working in two dimensions, since $l + m + s = 1$. It can also be seen that *a, b*, and *c* are functions of only two parameters, l and m, since the chromaticity coordinates of the canonical illuminant, l' and m', are constant. Therefore having three coefficients does not in fact give the diagonal model an inherent advantage over the two-parameter affine model, since all three coefficients are functions of the same two parameters.

C. Diagonal Model with Sharpening

Like the diagonal model, the DMT with sharpening involves the independent scaling of three coordinates, but where the standard DMT operates on raw cone response vectors, this model first applies a linear transformation. Sharpening is applied to the (L, M, S) coordinates with a 3×3 sharpening matrix T [Eqs. (3) and (4)]:

$$\begin{bmatrix} L_s \\ M_s \\ S_s \end{bmatrix} = T \times \begin{bmatrix} L \\ M \\ S \end{bmatrix}. \tag{27}$$

We then work in the new (L_s, M_s, S_s) coordinate space. A diagonal transform is applied to the new coordinates as in Eq. (23):

$$\begin{bmatrix} L'_s \\ M'_s \\ S'_s \end{bmatrix} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} L_s \\ M_s \\ S_s \end{bmatrix}. \tag{28}$$

It should be noted that sharpening adds no extra free parameters to the diagonal model. The nine components of the sharpening matrix are a fixed attribute of the model—the same sharpening matrix is used for all illuminants and all images. The sharpening matrix is not something that is changed to suit the circumstances. If the sharpening matrix were allowed to change, then the diagonal model with sharpening would effectively become a nine-parameter least-squares fit, which clearly would be more accurate than Zaidi's two-parameter affine model. It is not possible to perform such a nine-parameter fit, because all that is known in the case of illumination modeling are the two parameters represented by the chromaticity of the illumination. There are insufficient data to solve for nine parameters. Sharpening simply optimizes the effectiveness of the available two parameters; it does not introduce new ones.

D. Error Metric

To determine how well a transformation works, we convert the predicted chromaticities of the patches, as well as the actual values, to $L^*a^*b^*$ coordinates[8] via *XYZ* coordinates. Converting from *LMS* to *XYZ* requires a linear transformation.[9] The actual *LMS* coordinates of the re-

flectances in the target illuminant are calculated by using the spectral information of the illuminant, as in Eq. (6).

In this experiment we are comparing to what degree of accuracy each method can predict chromaticities, and so we ignore luminance information. This is done by scaling the predicted values for each patch to match the actual luminance of that patch under the target illuminant. In other words, we scale each XYZ triple by Y_{actual}/Y, where Y_{actual} is the actual value of Y for that patch under the target illuminant, as calculated with spectral information. We are assuming that each method can predict luminance information perfectly, allowing us to compare their performance at predicting chromaticities independent of luminance.

The XYZ values are then converted to $L^*a^*b^*$ space, and the ΔE (Euclidean distance in $L^*a^*b^*$ space) error is calculated between each predicted/actual pair. Since we have scaled by Y_{actual}/Y, there is no error in the L coordinates, so we are effectively calculating errors in a two-dimensional chromaticity space. The error for a given transformation and source/target illuminant pair is the arithmetic mean of ΔE errors for each surface.

Table 3. Mean ΔE Errors for the Three Transformation Methods under Various Illuminations Changes[a]

Δ Illum.	Affine	Diagonal	Sharp
1 → 2	1.3770	1.3878	**0.6415**
1 → 3	4.6767	4.6066	**2.2696**
1 → 4	3.6417	3.5209	**1.9154**
1 → 5	5.4890	**5.2359**	5.5988
1 → 6	5.0514	**4.2943**	4.5018
2 → 1	1.3471	1.3516	**0.6244**
2 → 3	3.2520	3.1731	**1.6862**
2 → 4	4.8454	4.7062	**2.3222**
2 → 5	5.4898	**5.0146**	5.6132
2 → 6	5.5378	4.7249	**4.5612**
3 → 1	4.2844	4.2056	**2.2446**
3 → 2	3.0395	2.9658	**1.7139**
3 → 4	*	*	*
3 → 5	6.7828	**6.0704**	6.3894
3 → 6	7.4110	6.5305	**5.3664**
4 → 1	4.0076	3.9832	**2.2910**
4 → 2	5.4638	5.4723	**2.8850**
4 → 3	*	*	*
4 → 5	5.8199	6.5076	**5.7592**
4 → 6	4.5408	4.4755	**4.3232**
5 → 1	5.8771	**5.6795**	5.7869
5 → 2	6.3195	5.9182	**5.8582**
5 → 3	9.1361	8.4556	**6.4883**
5 → 4	5.3768	5.7662	**5.1354**
5 → 6	1.9891	2.0129	**1.0505**
6 → 1	6.0512	5.2903	**5.0901**
6 → 2	6.9800	6.1134	**5.1799**
6 → 3	10.4682	9.4186	**5.9129**
6 → 4	4.6511	4.6010	**4.4653**
6 → 5	2.0605	2.1618	**1.1126**
Mean	5.0345	4.7730	**3.8138**
3 → 4	7.4434	7.1846	**3.4056**
4 → 3	8.9822	8.8912	**4.2255**

[a] Sharpening is done with $T_{4,3}$. Results for this pair of illuminants are not included in the global mean but are shown at the bottom of the table.

5. RESULTS

Table 3 shows the mean ΔE error for the three different methods under each of the 30 possible illumination changes. The mean over all the illumination changes is also shown for each method at the bottom of the table. For each change in illumination the method that gives the best results, i.e., the lowest mean ΔE error, is shown by bold text.

The sharpening matrix used was $T_{4,3}$, and so sharpening has an inherent advantage with this pair of illuminants. Because of this, results for this pair of illuminants are not calculated into the global mean and are shown only at the bottom of Table 1.

6. CONCLUSIONS

In this experiment we have examined several methods of predicting cone response vectors under a target illuminant, given the vectors under a source illuminant and the chromaticities of the two illuminants. We assessed the performance of these methods based on how closely the predicted vectors matched the actual vectors. The results in Table 3 clearly show that if accuracy in predicting colors under a different illuminant is the goal, the two-parameter affine transformation offers no advantage over the two-parameter diagonal-matrix transform (DMT) or the DMT with sharpening. In no case did the affine transformation outperform both the DMT and the DMT with sharpening, although in 7 out of 30 cases it outperformed the DMT.

REFERENCES

1. Q. Zaidi, "Identification of illuminant and object colors: heuristic-based algorithms," J. Opt. Soc. Am. A **15**, 1767–1776 (1998).
2. D. I. A. MacLeod and R. M. Boynton, "Chromaticity diagram showing cone excitation by stimuli of equal luminance," J. Opt. Soc. Am. **69**, 1183–1186 (1979).
3. G. West and M. H. Brill, "Necessary and sufficient conditions for von Kries chromatic adaption to give colour constancy," J. Math. Biol. **15**, 249–258 (1982).
4. G. D. Finlayson, M. S. Drew, and B. V. Funt, "Spectral sharpening: sensor transformations for improved color constancy," J. Opt. Soc. Am. A **11**, 1553–1563 (1994).
5. M. J. Vrhel, R. Gershon, and L. S. Iwan, "Measurement and analysis of object reflectance spectra," Color Res. Appl. **19**, 4–9 (1994).
6. E. L. Krinov, *Spectral Reflectance Properties of Natural Formations*, Technical Translation TT-439 (National Research Council of Canada, Ottawa, 1947).
7. J. J. Vos and P. L. Walraven, "On the derivation of the foveal receptor primaries," Vision Res. **11**, 799–818 (1971).
8. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulas*, 2nd ed. (Wiley, New York, 1982), Eq. [5(3.3.9)].
9. See Ref. 7, Eq. (26).