Reconstruction of reflectance data using an interpolation technique

Farhad Moghareh Abed,[1,*] Seyed Hossein Amirshahi,[2] and Mohammad Reza Moghareh Abed[2]

[1]*Department of Color Physics, Institute for Colorants, Paint & Coatings, Tehran, Iran*
[2]*Department of Textile Engineering, Amirkabir University of Technology, Tehran, Iran*
**Corresponding author: abed@icrc.ac.ir*

Received May 6, 2008; revised November 17, 2008; accepted November 23, 2008;
posted January 5, 2009 (Doc. ID 95771); published February 23, 2009

A linear interpolation method is applied for reconstruction of reflectance spectra of Munsell as well as ColorChecker SG color chips from the corresponding colorimetric values under a given set of viewing conditions. Hence, different types of lookup tables (LUTs) have been created to connect the colorimetric and spectrophotometeric data as the source and destination spaces in this approach. To optimize the algorithm, different color spaces and light sources have been used to build different types of LUTs. The effects of applied color datasets as well as employed color spaces are investigated. Results of recovery are evaluated by the mean and the maximum color difference values under other sets of standard light sources. The mean and the maximum values of root mean square (RMS) error between the reconstructed and the actual spectra are also calculated. Since the speed of reflectance reconstruction is a key point in the LUT algorithm, the processing time spent for interpolation of spectral data has also been measured for each model. Finally, the performance of the suggested interpolation technique is compared with that of the common principal component analysis method. According to the results, using the CIEXYZ tristimulus values as a source space shows priority over the CIELAB color space. Besides, the colorimetric position of a desired sample is a key point that indicates the success of the approach. In fact, because of the nature of the interpolation technique, the colorimetric position of the desired samples should be located inside the color gamut of available samples in the dataset. The resultant spectra that have been reconstructed by this technique show considerable improvement in terms of RMS error between the actual and the reconstructed reflectance spectra as well as CIELAB color differences under the other light source in comparison with those obtained from the standard PCA technique. © 2009 Optical Society of America

OCIS codes: 300.6550, 330.1690, 330.1730.

1. INTRODUCTION

Color is usually characterized by three colorimetric values in standard spaces such as CIEXYZ or CIELAB colorimetric systems. Obviously, this type of description includes the effect of several parameters such as the illuminant, the observer, and the material in which the object is not individually measured. On the other hand, the main factor that fully characterizes a color is the spectral behavior of the object. Hence, depending on the optical property of the stimulus, the reflectance or the transmittance behavior of the object is known as the fingerprint of the color, which is independent of the applied light source and the observer for nonfluorescent materials. Thus, color could be instrumentally measured and described in two different ways using colorimetric values or spectral information.

Clearly, the spectral data give the most useful information for color specification under different viewing conditions as well as for color reproduction algorithms. In fact, the reflectance curve of an opaque material could be considered as a principal element that could fully characterize the color of the object. On the other hand, color could be quantitatively defined as a projection of the semi-infinite-dimensional spectral space to a fully defined three-dimensional space, i.e., the product of light source and observer. This operation provides the capability of representing the surface color information by a set of tristimulus values [1,2]. Although the computation of colorimetric information from spectral data can be easily performed, the calculation of spectral reflectance from the colorimetric value is an underdetermined problem and so is not a routine procedure.

Several mathematical techniques such as the simplex method, genetic algorithms, simulated annealing, application of subtractive and additive primaries, neural networks, and the principal component analysis (PCA) technique have been introduced for extracting reflectance data from the relevant color coordinates [3–17]. The most suitable technique implements the well-known linear model based on PCA. The technique linearly reduces the dimensionality of the data while protecting as much as possible the variations present in the dataset. The method works on the strong correlation of the reflectance spectra of natural and man-made surfaces, which are generally smooth functions of wavelength over the range of the visible spectrum. These functions may be represented as the weighted sum of a small numbers of orthogonal basis functions. The functions can be simply obtained from a suitable set of available spectra by applying the PCA technique [11–14].

This paper presents a method for reproducing reflectance data from the tristimulus values by using a linear interpolation method. The method is used for recovery of spectral data of Matt Munsell color chips and the color

samples of ColorChecker SG from the color coordinates of samples under a given set of illuminant–viewing conditions. The performance of the proposed method in the spectral estimation of samples was compared with those obtained with the classical PCA routine. The evaluations were conducted by calculation of color difference values under different illuminants and RMS errors between the actual and the reconstructed spectra.

2. THEORETICAL BACKGROUND

A. Linear Interpolation Method

Interpolation methods are classified as statistical methods and are widely used in the field of engineering such as color engineering. These methods are usually employed when the relationships between two groups of data are so complicated that physical and other global statistical methods such as regression models do not yield accurate results [18].

In general, the problem of scattered data interpolation consists of constructing a continuous function of two or more independent variables interpolating data values that are known only at some discrete points in the two-dimensional (2D) plane or, respectively, in the n-dimensional space. The need for interpolation from irregularly scattered data occurs in many different fields, such as medical imaging, meteorological or geological modeling, and cartography, as well as color reproduction efforts. Interpolation methods usually concern single-valued data, where the underlying function has the form $f:R^2 \rightarrow R$ or $f:R^3 \rightarrow R$. However, other cases may require the interpolation of multivalued data, where the underlying function appears in the form $f:R^2 \rightarrow R^2$ or $f:R^3 \rightarrow R^3$. For example, in the field of electronic color reproduction systems and color management, the interpolation is usually carried out in the case of $f:R^3 \rightarrow R^3$, including calibration and/or characterization of color imaging devices (such as a scanner or printer) in which the device-dependent coordinates (such as RGB or CMYK) map into an independent color space (such as CIEXYZ or CIELAB). In such cases the interpolation is performed in the form of $f:R^3 \rightarrow R^3$ to establish a connection between two 3D spaces usually done using standard color patches. For each color patch the device-independent color values obtained by a spectrophotometer or colorimeter and the corresponding device-dependent values, i.e., RGB or CMYK, form a lookup table (LUT) required for interpolation [19–21].

All interpolation techniques can be classified into two categories, i.e., global and local methods. In global methods, the interpolated point is affected by all elements of the data set. In contrast, in local interpolation methods, the interpolated value is influenced only by the values at "nearby" points. Global methods are usually used for small data sets, and each variation or modification in the data set will modify the final interpolated values. In turn, local methods are capable of treating larger data sets with less sensitivity to data modification [19].

The linear triangular (or tetrahedral) interpolation technique is one of the methods among different approaches to the interpolation of scattered data. In this method, the interpolation process is divided into two steps: first the scattered points in the set are triangulated (in 2D space), tetrahedrized (in 3D space) or tessellated (in n-D space), and then an interpolation scheme is performed within each triangle or tetrahedron or hypertetrahedron (the names "simplex" or "pentatope" are also used).

In such methods a given point set in the 3D space could be shaped or partitioned by many different tetrahedrization procedures. Delaunay tetrahedrization is one of the "nicer" methods in which tetrahedrization is obtained by connecting all the neighboring points in the Voronoi diagram of the given point set. The Voronoi tessellation partitions the plane (or space) into an exhaustive and disjoint set of tiles (polygons, or respectively, polytopes), where each tile (T_i) encloses one point x_i of the given point set. The tile T_i is defined as the area (or volume) that is closer to the scatter point x_i than to any other scatter points:

$$T_i = \{x \in R^2 | d(x,x_j) \leqslant d(x,x_j) \; \forall \, j = 1 \ldots n\}, \tag{1}$$

where $d(a,b)$ denotes the Euclidean distance between the points a and b. Illustrations of Delaunay triangulation and Voronoi partitioning are shown in Fig. 1 [18,19].

In the 3D space, for instance, it is assumed that the scattered points (x_i,y_i,z_i) are first tetrahedrized using the Delaunay tetrahedrization method. Each of these tetrahedrals is used to interpolate any arbitrary point inside it. Given the four vertices $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$, and $\mathbf{P}_4$ of any such tetrahedron, in the locations (x_1,y_1,z_1), (x_2,y_2,z_2), (x_3,y_3,z_3), (x_4,y_4,z_4) of its four vertices that have their values $\mathbf{R}$, the color values at any arbitrary point $\mathbf{P}$ located at (x,y,z) in this tetrahedron can be found by writing the linear interpolation polynomial,

$$P = aP_1 + bP_2 + cP_3 + dP_4. \tag{2}$$

Then, the coefficients a,b,c,d could be found simply by solving the linear system of equations

$$\begin{aligned} ax_1 + bx_2 + cx_3 + dx_4 &= x, \\ ay_1 + by_2 + cy_3 + dy_4 &= y, \\ az_1 + bz_2 + cz_3 + dz_4 &= z, \\ a + b + c + d &= 1. \end{aligned} \tag{3}$$

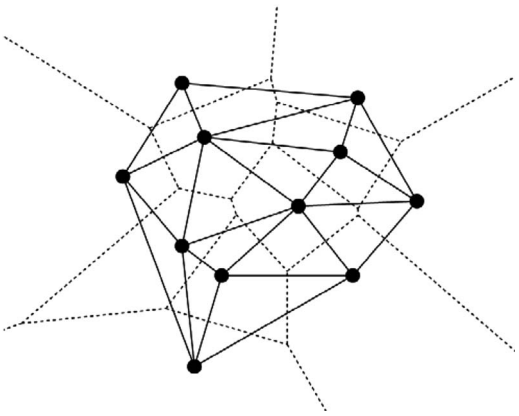

Fig. 1. Voronoi and Delaunay partitioning for scattered points in a plane. Solid lines indicate the Delaunay triangles associated with the Voronoi tessellation, and the dotted lines inside it show the Voronoi tiles of the scattered point set [19].

Alternatively, the value of **R** at any point **P** located at (x,y,z) within the tetrahedral is

$$R_j = aR_{j_{P_1}} + bR_{j_{P_2}} + cR_{j_{P_3}} + dR_{j_{P_4}}, \tag{4}$$

where the weights a,b,c,d are the barycentric coordinates of the point **P** in the tetrahedron $\mathbf{P_1P_2P_3P_4}$. The a,b,c,d coefficients can be calculated by the matrix equation

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ z_1 & z_2 & z_3 & z_4 \\ 1 & 1 & 1 & 1 \end{bmatrix}^{-1} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}. \tag{5}$$

It is also possible simply to generalize this algorithm to the higher 4- or 6-dimensional spaces. For example in 4D space, the scattered points can be located by (x_i, y_i, z_i, t_i) and the color values at any arbitrary point **P** located at (x, y, z, t) inside the 4-simplex $\mathbf{P_1P_2P_3P_4P_5}$ can be calculated by

$$R_j = aR_{j_{P_1}} + bR_{j_{P_2}} + cR_{j_{P_3}} + dR_{j_{P_4}} + eR_{j_{P_5}}, \tag{6}$$

where coefficients a,b,c,d,e are found by extending Eqs. (3) in 4-D space:

$$\begin{pmatrix} a \\ b \\ c \\ d \\ e \end{pmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ y_1 & y_2 & y_3 & y_4 & y_5 \\ z_1 & z_2 & z_3 & z_4 & z_5 \\ t_1 & t_2 & t_3 & t_4 & t_5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}^{-1} \begin{pmatrix} x \\ y \\ z \\ t \\ 1 \end{pmatrix}. \tag{7}$$

Analogous equations could be arranged for n-D space by extending Eqs. (4) and (5) as follows

$$R_j = aR_{j_{P_1}} + bR_{j_{P_2}} + cR_{j_{P_3}} + dR_{j_{P_4}} + \cdots + nR_{j_{P_n}}, \tag{8}$$

$$\begin{pmatrix} a \\ b \\ c \\ d \\ \cdots \\ m \\ p \end{pmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & \cdots & x_n & x_{n+1} \\ y_1 & y_2 & y_3 & y_4 & \cdots & y_n & y_{n+1} \\ z_1 & z_2 & z_3 & z_4 & \cdots & z_n & z_{n+1} \\ t_1 & t_2 & t_3 & t_4 & \cdots & t_n & t_{n+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ m_1 & m_2 & m_3 & m_4 & \cdots & m_n & m_{n+1} \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^{-1} \begin{pmatrix} x \\ y \\ z \\ t \\ \cdots \\ m \\ 1 \end{pmatrix}. \tag{9}$$

It should be emphasized that there are different methods for calculating linear interpolation coefficients within any specific triangle or tetrahedral, leading to identical results [18,19].

B. Principal Component Analysis

The CIEXYZ tristimulus values of surface color under a given set of illuminant and observer are simply calculated by

$$T_j = \int_{400}^{700} S_\lambda r_\lambda q_{j,\lambda} \mathrm{d}_\lambda, \tag{10}$$

where T denotes the tristimulus values of X, Y, and Z, and j varies from 1 to 3 over X, Y, and Z. S, r, and q represent the spectral power distribution of illuminant, the reflectance spectrum of the surface color, and the color matching functions of CIE standard observer, respectively. Equation (10) can be rewritten in the matrix form as

$$\mathbf{T} = \mathbf{A}^{\mathrm{T}}\mathbf{r}, \tag{11}$$

where **A** is the weight product sets of standard illuminant and observer and the superscript T shows the matrix transpose.

The most common approach to estimating the spectral reflectance of surface colors from the CIEXYZ tristimulus values is based upon applying dimensionality reduction techniques. The method originates from the fact that the reflectance spectra of natural and man-made surfaces are generally smooth functions of wavelength over the visible spectrum. So, the spectra are strongly correlated and may be represented as the weighted sum of a small number of orthogonal basis functions called eigenvectors. The algorithm in its classical form has been simply formulated by Fairman and Brill [11]. This method uses linear models that represent each reflectance spectrum through the weighted sum of the small number k of basis functions $\mathbf{v}_j$, as shown in

$$\mathbf{r} \approx \mathbf{v}_0 + \sum_{j=1}^{k} c_j \mathbf{v}_j, \tag{12}$$

where **r** is the spectral reflectance, c_j is the weight of the jth basis function (called principal component) and produces the least-square best fit with that reflectance spectrum, $\mathbf{v}_0$ shows the mean spectral reflectance value of the dataset, and $\mathbf{v}_j$ is the jth eigenvector. Equation (12) can be written in the matrix form as

$$\mathbf{r} \approx \mathbf{v}_0 + \mathbf{V}\mathbf{c}, \tag{13}$$

where **c** is a column vector of **k** elements that contains the principal components, and **V** is the first k eigenvectors, so that $\mathbf{V}=[\mathbf{v}_1,\ldots,\mathbf{v}_k]$. Now, the term of reflectance in Eq. (13) can be replaced in Eq. (11) to give

$$\mathbf{T} = \mathbf{A}^{\mathrm{T}}\mathbf{v}_0 + \mathbf{A}^{\mathrm{T}}\mathbf{V}\mathbf{c}, \tag{14}$$

where **T** is the tristimulus values of a given sample whose spectral reflectance is desired, and $\mathbf{A}^{\mathrm{T}}\mathbf{v}_0$ and $\mathbf{A}^{\mathrm{T}}\mathbf{V}$ are the vector and matrix containing the tristimulus values of the mean spectrum and the first k basis functions, respectively. By denoting them as $\mathbf{T}_{\mathrm{V}_0}$ and $\mathbf{T}_{\mathrm{V}}$, we can form

$$\mathbf{T} = \mathbf{T}_{\mathrm{V}_0} + \mathbf{T}_{\mathrm{V}}\mathbf{c}. \tag{15}$$

Since the values of $\mathbf{T}_{\mathrm{V}_0}$ and $\mathbf{T}_{\mathrm{V}}$ are known, the calculation of column vector **c** is simply

$$\mathbf{c} = \mathbf{T}_{\mathrm{V}}^{-1}(\mathbf{T} - \mathbf{T}_{\mathrm{V}_0}). \tag{16}$$

By applying the column vector **c** in Eq. (13) the spectral reflectance could be generated that is sure to match the

actual reflectance under a given light source and to possess the same tristimulus values **T** of the proposed sample.

According to the literature, the number of basis vectors needed for acceptable recovery of reflectance largely depends on the type of applied datasets. However, reconstruction of reflectance data from the proposed tristimulus values in the case of one illuminant–observer condition limits the number of eigenvectors to three to prepare a fully determined equation system. This limitation could clearly lead to unsatisfactory estimation. To overcome such a problem, Ansari *et al.* suggested an adaptive technique to recover the spectral reflectance of surface colors from their color coordinates [5]. Similarly, to improve the reconstruction results, Ayala and his colleagues divided the Munsell color space into ten different zones; hence, different basis vectors were obtained for each subspace. The idea of dividing Munsell chips into different groups was enhanced by extending the principal components of PCA [13]. Finally, the weighted version of PCA has been recently introduced by Agahian and her coworkers to reconstruct the spectral reflectance of samples from their colorimetric data under one set of illuminant–observer conditions [12].

3. EXPERIMENTAL

A dataset consisting of 1269 reflectance spectra of color chips in the Munsell Book of Color—Matt Finish Collection was borrowed from the color group of the University of Joensuu [22]. The dataset was measured with a Perkin-Elmer Lambda 18 spectrophotometer; the wavelength range was from 380 nm to 800 nm at 1 nm interval. The reflectance values of 137 colored samples of ColorChecker SG from GretagMacbeth were also measured, personally. The color measurements of samples were carried out using a GretagMacbeth Color-Eye 7000A spectrophotometer with d/8 geometry. The reflectance values were measured at 10 nm intervals from 400 nm to 700 nm with specular component excluded. In the current research, the reflectance data of the Munsell collections was fixed between 400 nm to 700 nm at 10 nm intervals. The CIEXYZ as well as CIELAB values of samples were calculated under D65, A, and TL84 illuminants and 1964 standard observer.

A. Interpolation Method

In the present study, we concentrated on the local linear interpolation method to estimate the reflectance curve in n-dimensional space from the corresponding tristimulus values, i.e., CIEXYZ or CIELAB values. As a result, the underlying function has the form $f:R^m \rightarrow R^n$, in which m number of color values are used to reconstruct the reflectance from an m-dimensional space to an n-dimensional space. The linear interpolation method was chosen because of the characteristic of tristimulus values and because of its linear combination of reflectance curves in each wavelength shown in Eq. (10). As a result, the linear interpolation method seemed to yield the most acceptable results.

To interpolate the target points, the scattered color points in the first space were tetrahedrized and then an interpolation scheme was used within each tetrahedral created. The Delaunay tetrahedrization method was used in the first 3D space to tetrahedrize the scattered points. By finding the optimal tetrahedron that enclosed the arbitrary new point **P**, the linear interpolation scheme was used within the tetrahedral to calculate the reflectance curve. The reflectance space is directly related to the wavelength interval for defining of the reflectance over the visible spectrum. For example, if the reflectance curve samples in 31 wavelengths, the reflectance space would be a 31-dimensional space.

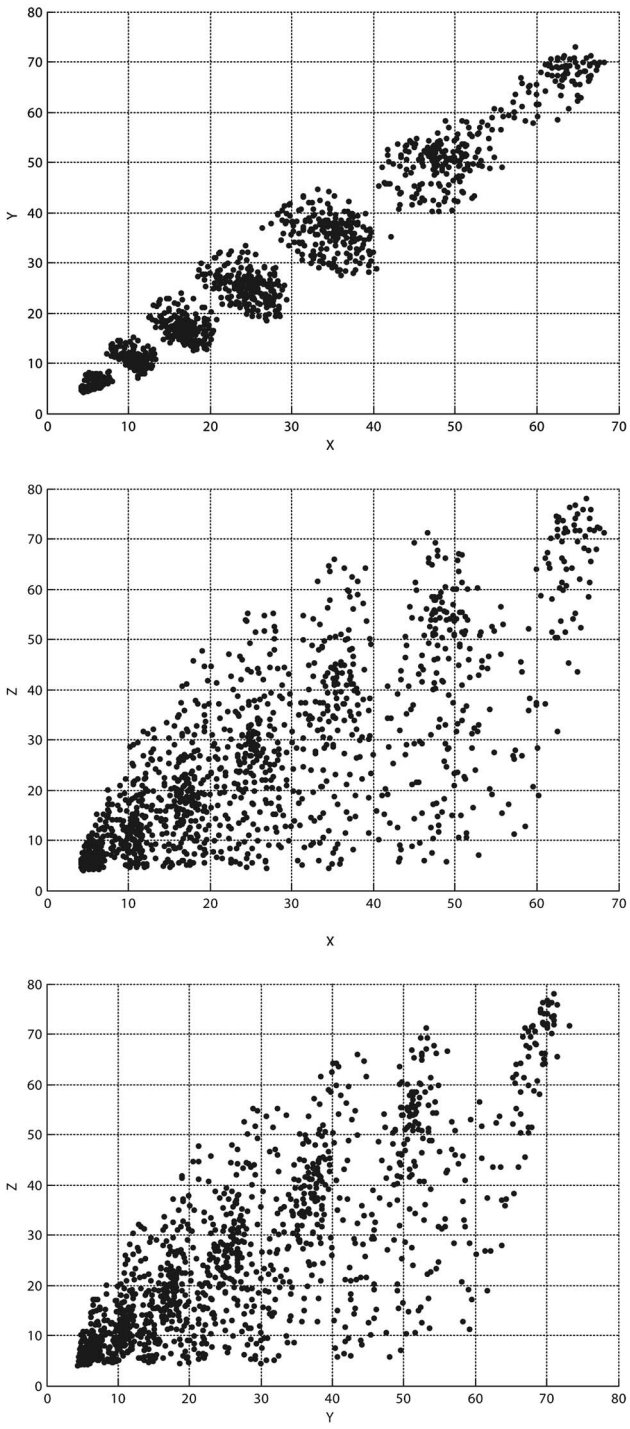

Fig. 2. Tristimulus values of 1269 Matt Munsell color chips in CIEXYZ space under D65 illumination and 10° observer.

First the tristimulus color values of 1269 Matt Munsell color chips were assumed as distinct points $\mathbf{P}_1,\ldots,\mathbf{P}_n$ that formed a 3D space. Figure 2 shows the color values of 1269 Matt Munsell color chips in the 2D subplots of CIEXYZ color space under D65 illumination and 10° viewing condition. In such condition, point $\mathbf{P}_i$ located in the vector position in CIEXYZ 3D space has a numerical value, i.e. $\mathbf{R}_{\lambda,i}$. Here, the task was to find a good interpolation function $\hat{f}(x)$ such that $R_{\lambda,i}=\hat{f}(x)$ for all $\lambda=1\ldots n$ (n is the dimension of reflectance space, that is, 31 in this study). The interpolation method [Eqs. (2) and (3)] was used to estimate the $\mathbf{R}_{\lambda,i}$ in the 31-dimensional space.

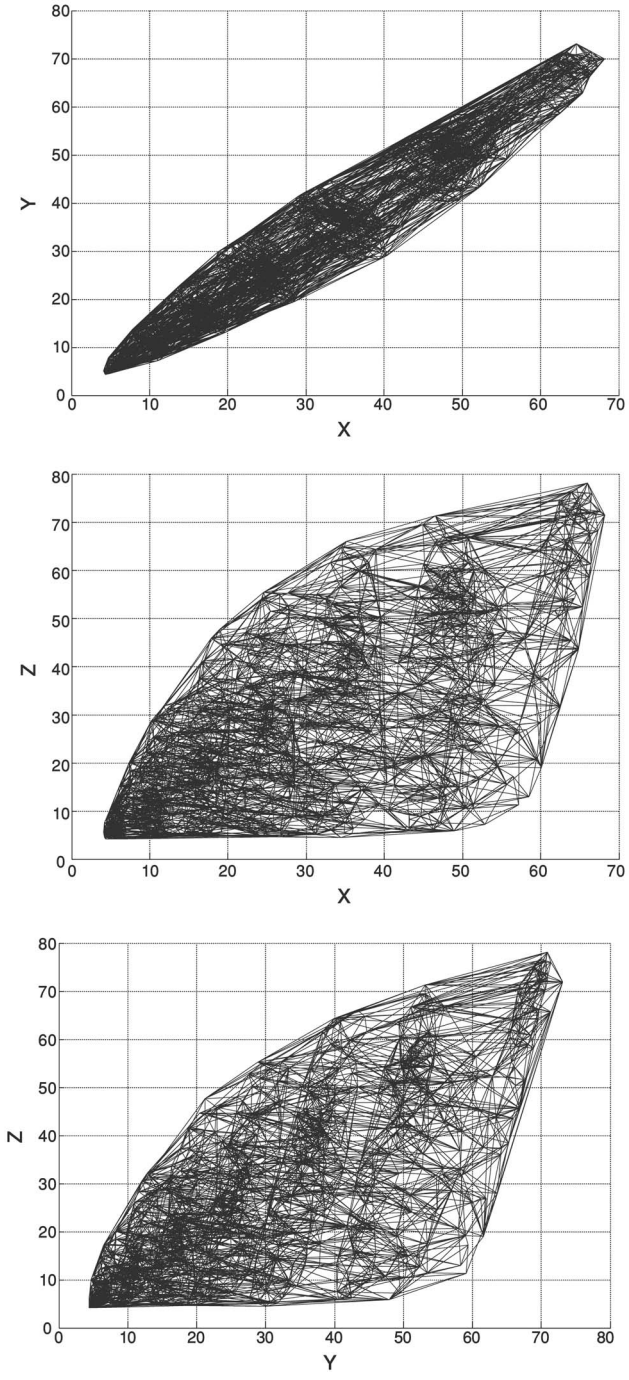

Fig. 3. The 1269 Munsell color points in CIEXYZ tetrahedrized by the Delaunay method.

Figure 3 shows the 1269 Munsell color points in CIEXYZ that is tetrahedrized by the Delaunay method in 3D space.

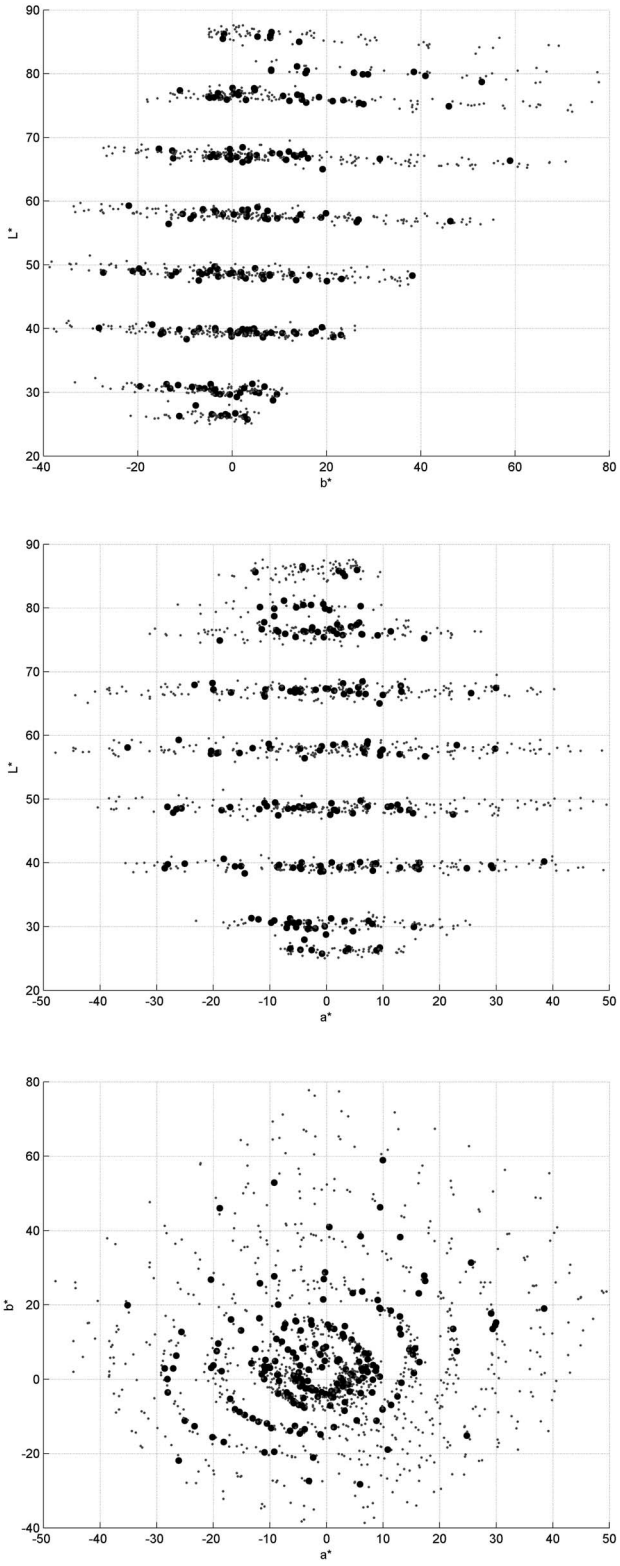

Fig. 4. Distribution of 200 chosen samples and 1069 Munsell color points in CIELAB space using D65 illuminant and 1964 standard observer. 200 Munsell samples are shown by solid circles.

Table 1. Spectral and Colorimetric Accuracy of Spectral Estimation by LUT and PCA Methods

			Interpolation Method		PCA Method	
Color Coordinates in Source LUT			XYZ_{D65}	XYZ_{D65} and XYZ_A	XYZ_{D65}	XYZ_{D65} and XYZ_A
Color Difference (ΔE)	Mean	D65	0	0	0	0
		A	0.5013	0	1.3472	0
		TL84	0.7123	0.2771	1.4715	0.9876
	Maximum	D65	0	0	0	0
		A	2.9086	0	7.0532	0
		TL84	6.9625	1.4703	8.1676	15.3214
	Variance	D65	0	0	0	0
		A	0.2515	0	1.0441	0
		TL84	0.8037	0.0429	1.5455	1.7974
RMS	Mean		0.0079	0.0027	0.0188	0.0082
	Maximum		0.0592	0.0147	0.0951	0.0308
	Variance		0.0000	0.0000	0.0001	0.0000

To investigate the performance of the method, the spectral reflectance curves of 200 color patches were reconstructed and compared with those obtained by the classical PCA method. These patches were chosen from 1269 Munsell patches in order to effect a good coverage in all parts of color space; there are no color points outside the chosen color gamut. Figure 4 shows the color distribution of the Munsell color chips and the selected color samples in CIELAB color space. The mean and the maximum values of RMS error between the actual and the estimated reflectance spectra and the color difference values under A and TL84 illuminants were calculated for both methods. It should be emphasized that the testing samples were not employed in the creation of LUT tables, nor in determination of eigenvectors in the PCA method. Therefore the number of Munsell color samples for creation of the first space in the LUT method was 1069, and these samples were also used for extraction of the proposed eigenvectors. The processing time of the LUT method was also investigated.

4. RESULTS AND DISCUSSION

A. General Achievements

Table 1 shows the RMS error between the actual and reconstructed reflectance values and the color difference values under different illuminants. The XYZ tristimulus value samples under D65 and A illuminants were used by the LUT and PCA methods to recover the reflectance data.

As Table 1 shows, the mean as well as the maximum RMS errors and the color difference values are smaller for the LUT method in comparison with those for the PCA routine. According to Eq. (2), the reflectance curves were reconstructed by using only four neighboring points in the LUT method, while three color values are used for reconstruction. According to the results, RMS and color difference errors were smaller by using higher-dimensional spaces, i.e., employing a second set of XYZ values under different viewing conditions. The smaller RMS error achieved by the LUT method shows the closeness between the actual and the reconstructed curves. Figure 5 shows the selected samples from the LUT table during the interpolation steps that formed the reflectance curve of a randomly selected sample with X=32.4, Y=42.2, and Z=48.8 tristimulus values. As shown in this figure, in the LUT method, the XYZ values of four neighboring points were chosen for the recovery of reflectance data. Figure 6 shows the reconstructed reflectance curves of eight randomly selected samples. The actual and the recovered reflectance curves, which were estimated by the PCA method, are also shown in this figure. Consequently, excellent results could be expected if the first dataset covers a limited area of color space in the manner that the target sample places inside the source color gamut.

As opposed to classical recovery methods, one of the most important advantages of the LUT method is the creation of nonnegative reflection values. In fact, the inter-

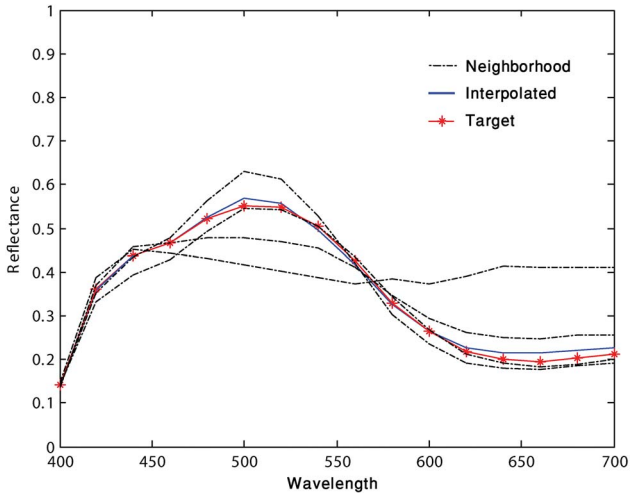

Fig. 5. (Color online) Reconstruction of reflectance data in 3D XYZ space. Asterisk (red online) and solid (blue online) curves show the standard and predicted reflectance curves, respectively, while dashed–dotted curves show the reflectance data of neighboring points that were used for the reconstruction process.

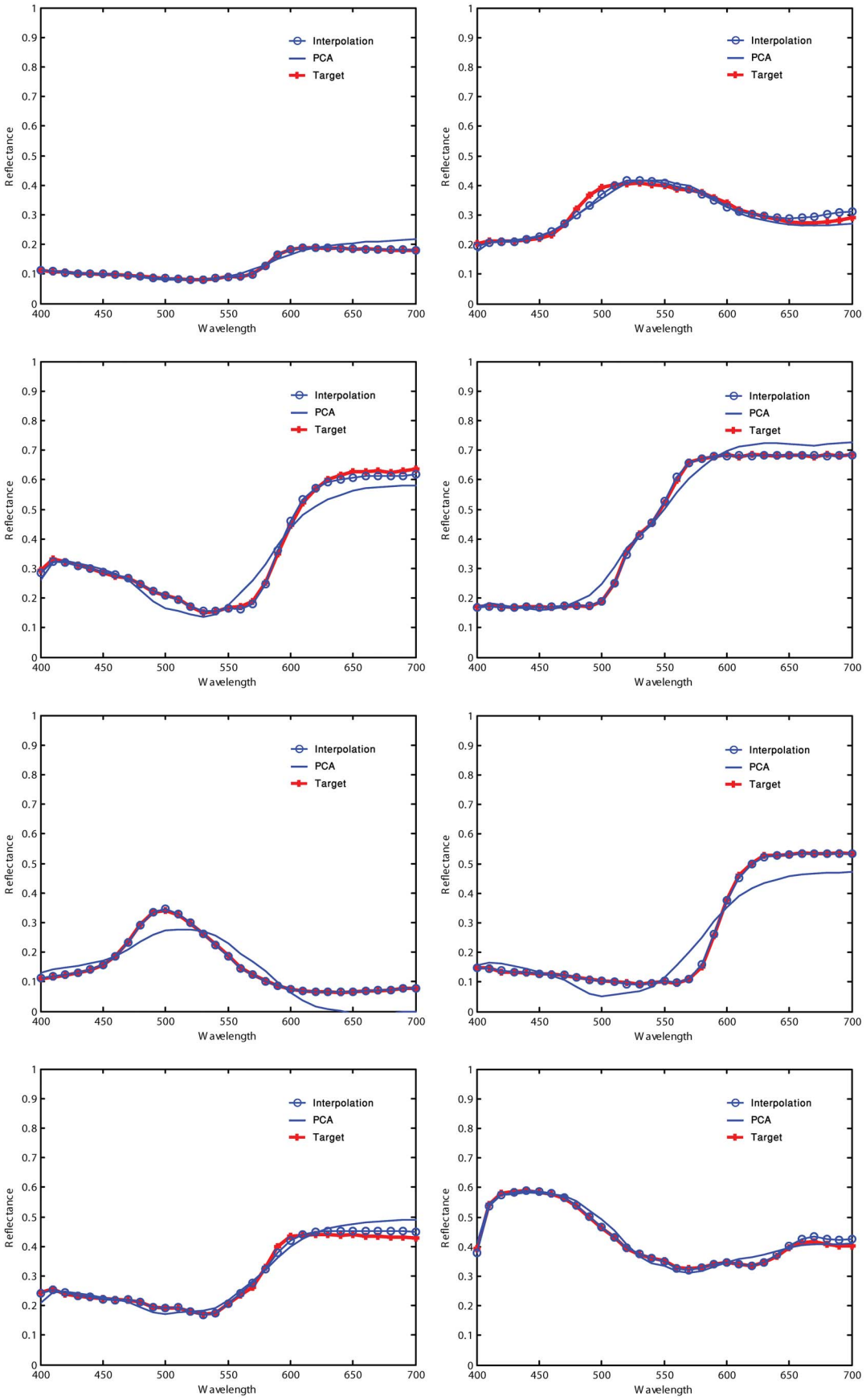

Fig. 6. (Color online) Results of spectral recovery of eight randomly selected samples of Matt Munsell chips from their tristimulus values by using LUT and PCA methods.

polation between positive reflectance data prevents the formation of any spiky or negative values. Such abnormal behaviors sometimes result when using some prior recovery techniques in the reconstruction process, and can be clearly observed in Fig. 6 for those that have been estimated by the PCA method.

RMS and color differences decreased for both PCA and LUT methods when two sets of colorimetric values under two A and D65 illuminants were used, as we expected. Overall, the comparison between PCA and LUT methods in different states proves the superiority of the LUT technique, which is shown by smaller RMS and ΔE values for this method. All corresponding conditions of reconstruction show slightly lower RMS and ΔE for LUT than for the PCA results.

A Z-test was performed to statistically analyze the mean of RMS errors and color difference values of PCA and the interpolation methods in 3D and 6D cases. The Z-test is recommended when the population is far enough apart from the type of distribution of samples. Results of the hypothesis tests performed are shown in Table 2 for comparison of the mean values for PCA and the interpolation method with the same conditions. If calculated Z values $>Z_{critic}$, the null hypothesis is rejected as the source of significant difference in the means of RMS or ΔE. According to Table 2 all calculated Z-values are greater than the upper critical value ($Z_{0.05}=1.645$) with 95% confidence level; consequently the null hypothesis is rejected in both 3D and 6D cases. This means that color difference values as well as RMS errors are significantly smaller for the LUT than for the PCA method.

B. Effect of Color Space Types

The interpolation method is based on the finding of neighboring points, and different color spaces provide different neighborhood points. So, CIEXYZ and CIELAB color spaces were used to study the effect of color space types on the recovery of 200 selected Munsell color samples. In addition, the CIEXYZ color space was used in the CIEYxy form. Table 3 shows the RMS errors and color difference values for three different color spaces. As this table shows, the CIEXYZ and CIELAB color spaces lead to similar results, while the CIEYxy space yields the worst results.

Table 2. Hypothesis Tests for RMS Errors and Color Difference Values for PCA and Interpolation Methods[a]

Hypothesis Test	Matched under Illuminant(s)	Z-Values
H_0: $\Delta E_{PCA}-\Delta E_{Interpolation}=0$ H_1: $\Delta E_{PCA}-\Delta E_{Interpolation}>0$ For illuminant A	One Illuminant: D65	10.5099
H_0: $\Delta E_{PCA}-\Delta E_{Interpolation}=0$ H_1: $\Delta E_{PCA}-\Delta E_{Interpolation}>0$ For illuminant Tl84	One Illuminant: D65	7.0050
H_0: $\Delta E_{PCA}-\Delta E_{Interpolation}=0$ H_1: $\Delta E_{PCA}-\Delta E_{Interpolation}>0$ For illuminant TLt84	Two Illuminants: D65 and A	7.4069
H_0: $RMS_{PCA}-RMS_{Interpolation}=0$ H_1: $RMS_{PCA}-RMS_{Interpolation}>0$	One Illuminant: D65	11.0276
H_0: $RMS_{PCA}-RMS_{Interpolation}=0$ H_1: $RMS_{PCA}-RMS_{Interpolation}>0$	Two Illuminants: D65 and A	15.2017

[a]Upper critical value $Z_{0.05}=1.645$ (95% confidence level). Null hypothesis: H_0: $\mu_{PCA}-\mu_{interpolation}=0$. Alternative hypothesis: H_1: $\mu_{interpolation}-\mu_{PCA}>0$.

As can be observed, a zero color difference value is achieved in the case of using CIEXYZ space. The reason may be found in the following equations:

$$X_t=\sum_{i=1}^{n+1}a_iX_i,$$

$$\hat{R}_\lambda=\sum_{i=1}^{n+1}a_iR_i,$$

$$\begin{aligned}\hat{X}_t&=\sum_{\lambda=400}^{700}\hat{R}_\lambda E_\lambda\bar{x}_\lambda=\sum_{\lambda=400}^{700}\left(\sum_{i=1}^{n+1}a_iR_i\right)E_\lambda\bar{x}_\lambda\\&=\sum_{\lambda=400}^{700}\sum_{i=1}^{n+1}(a_iR_{i,\lambda}E_\lambda\bar{x}_\lambda)=\sum_{i=1}^{n+1}\sum_{\lambda=400}^{700}(a_iR_{i,\lambda}E_\lambda\bar{x}_\lambda)\\&=\sum_{i=1}^{n+1}a_i\sum_{\lambda=400}^{700}(R_{i,\lambda}E_\lambda\bar{x}_\lambda)=\sum_{i=1}^{n+1}a_iX_i=X_t.\end{aligned}\tag{17}$$

In these equations n is the number of dimensions of the first space that interpolation is performed in. Similar equations can be written for the Y and Z values. So, mathematically, implementation of tristimulus values for reproducing the reflectance spectra yields exactly the same XYZ values, i.e., the reconstructed and standard reflectances would be metameric pairs.

C. Effect of Color Coordinates

To compare the results of LUT and PCA methods, the various numbers of color coordinates were also utilized, e.g., using just the Y coordinate as input data to recover the reflectance spectra. In other words, the reflectance data of chosen samples were reconstructed by using one of the X, Y, and Z tristimulus values and also different combinations of them. Results are shown in Table 4. As this table shows, in the case of 1D input, the most effective color coordinate was Y, while Z provided the worst recovery results. According to these results, application of the PCA method clearly yielded smaller errors in 1D space. In other words, in the case of one color coordinate, PCA leads to more accurate results. However, by increasing the dimension of the source space and using extra color coordinates for reconstruction LUTs, the recovery results improved noticeably. In 4D spaces, where the XYZ coordinates under illuminate D65 and tristimulus X under illuminant A have been employed to create the first space LUT, the RMS and mean of ΔE indicate better fitness.

D. Processing Time

The slowest steps in the LUT algorithm are the tessellation and interpolation processes, which verify the processing time. According to the equations discussed, the processing duration is completely dependent on the

Table 3. Spectral and Colorimetric Accuracy of Spectral Estimation by LUT Method Using Different Color Spaces

			Color Coordinates in Source Color Space		
			XYZ_{D65}	LAB_{D65}	xyY_{D65}
Color Difference (ΔE)	Mean	D65	0	0.1247	3.4089
		A	0.5013	0.3461	3.8609
		TL84	0.7123	0.482	4.0554
	Maximum	D65	0	0.5772	20.1946
		A	2.9086	1.9349	21.5153
		TL84	6.9625	2.7296	23.6602
	Variance	D65	0	0.0106	15.8759
		A	0.2515	0.121	17.5312
		TL84	0.8037	0.2535	18.5898
RMS	Mean		0.0079	0.0056	0.0248
	Maximum		0.0592	0.0304	0.1212
	Variance		0.0000	0.0000	0.0004

Table 4. Spectral and Colorimetric Accuracy of Spectral Estimation by LUT and PCA Methods Using Different Color Coordinates

			X_{D65}	Y_{D65}	Z_{D65}	XY_{D65}	XZ_{D65}	YZ_{D65}	XYZ_{D65} and X_A	XYZ_{D65} and XY_A
PCA	ΔE	Mean	15.0573	15.1305	19.0049	13.5947	9.3465	11.2119	0	0
			15.8328	16.245	20.9649	15.0013	8.0739	11.7245	0.1899	0.1412
			15.806	16.0727	20.5863	15.327	8.8375	11.3518	0.9366	0.7077
	RMS	Mean	0.0549	0.0582	0.0893	0.0475	0.0342	0.0507	0.0108	0.0083
		Maximum	0.1852	0.208	0.3298	0.2586	0.1609	0.2749	0.0452	0.0282
		Variance	0.0014	0.0016	0.0052	0.0015	0.0006	0.0019	0.0000	0.0000
LUT	ΔE	Mean	22.576	22.1128	29.2599	11.9537	12.5323	12.8872	0	0
			23.3595	23.3057	31.4674	13.0925	10.6430	13.0216	0.0569	0.0469
			23.3444	23.0808	31.3248	13.4512	11.5146	12.6392	0.3527	0.2872
	RMS	Mean	0.0813	0.0891	0.1387	0.0417	0.0407	0.0565	0.0041	0.0032
		Maximum	0.2939	0.3016	0.4079	0.1828	0.1401	0.2869	0.0210	0.0168
		Variance	0.0043	0.0043	0.0079	0.0012	0.0013	0.0023	1.09E−05	0.0000

dimensions of the first space if the number of samples in the first LUT remains constant. The results are contained in Table 5. In fact, the one-dimensional spaces benefited from the highest speed, while the processing time was longer for higher-dimensional space. This information also is helpful to estimate the process time for each model before the prediction process. In Subsection 4.E it is shown that the number of the samples in the first space affects the tessellation and interpolation process time the same as the dimension of the LUT space.

E. Trial of Different Datasets

In this part of the work, we attempted to examine the role of applied datasets in the performance of the LUT method in reflectance recovery. Hence, two different LUTs were

Table 5. Processing Time of Reflectance Recovery by LUT Method for Different Input Spaces

Dimension of Source Space	1D	2D	3D	4D	5D	6D
Total time (s)	1.8906	2.5313	5.1094	18.00	76.6719	269.5
Tessellation time (s)	0.1875	0.2188	0.375	0.5500	2.2813	13.6875
Mean for each sample (s)	0.0082	0.0113	0.0234	0.0900	0.3719	1.2791

constructed by using different input–output data. First, the tristimulus values of Munsell color chips were used as the first space and the reflectance behaviors of 137 samples of ColorChecker SG were estimated. In the next turn, the color sets were replaced and the reflectance behavior of 200 Munsell color chips were predicted from the 137 GrateMacbeth samples. Figures 7 and 8 show the color specifications of samples in the CIELAB color space. As these figures show, the ColorChecker samples enclose the highly saturated as well as lighter samples. The same strategy was employed in reconstruction of reflectance curves by PCA. In fact, the eigenvectors of the Munsell

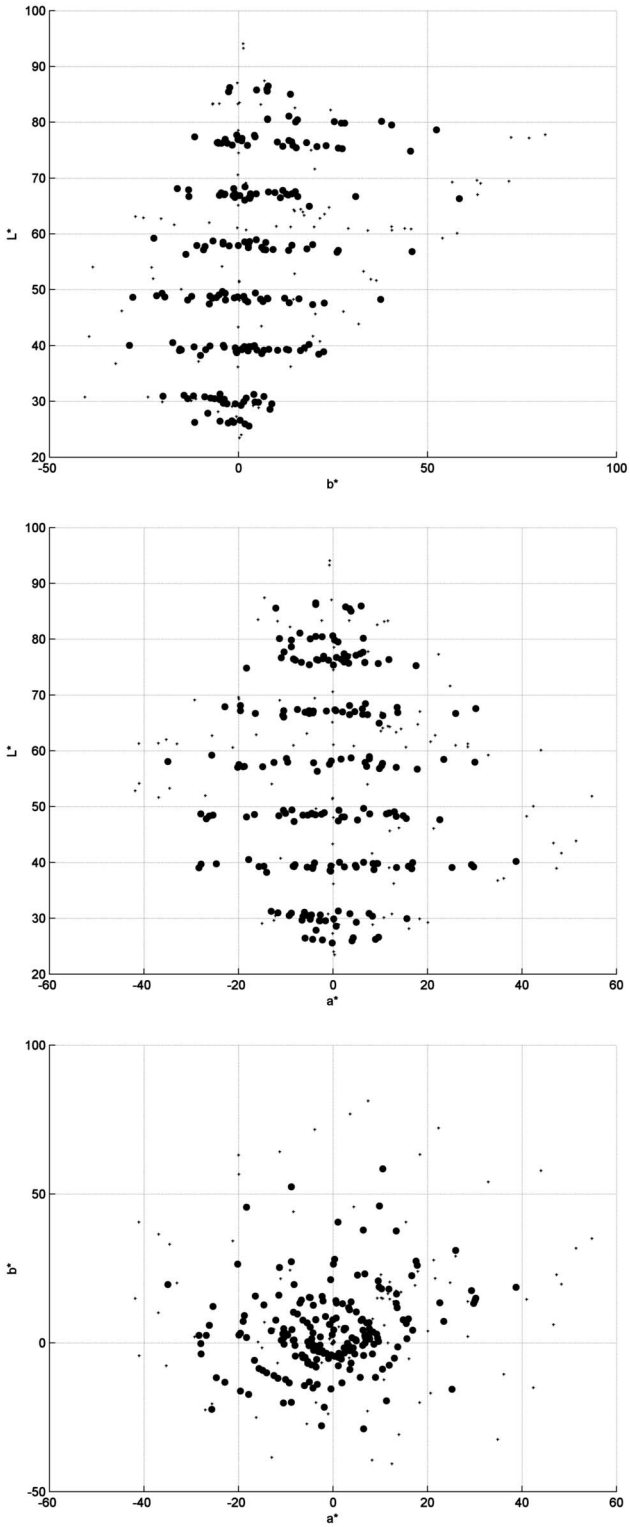

Fig. 7. Distribution of 137 samples of ColorChecker chart and 200 Munsell color samples in L*a*b* color space.

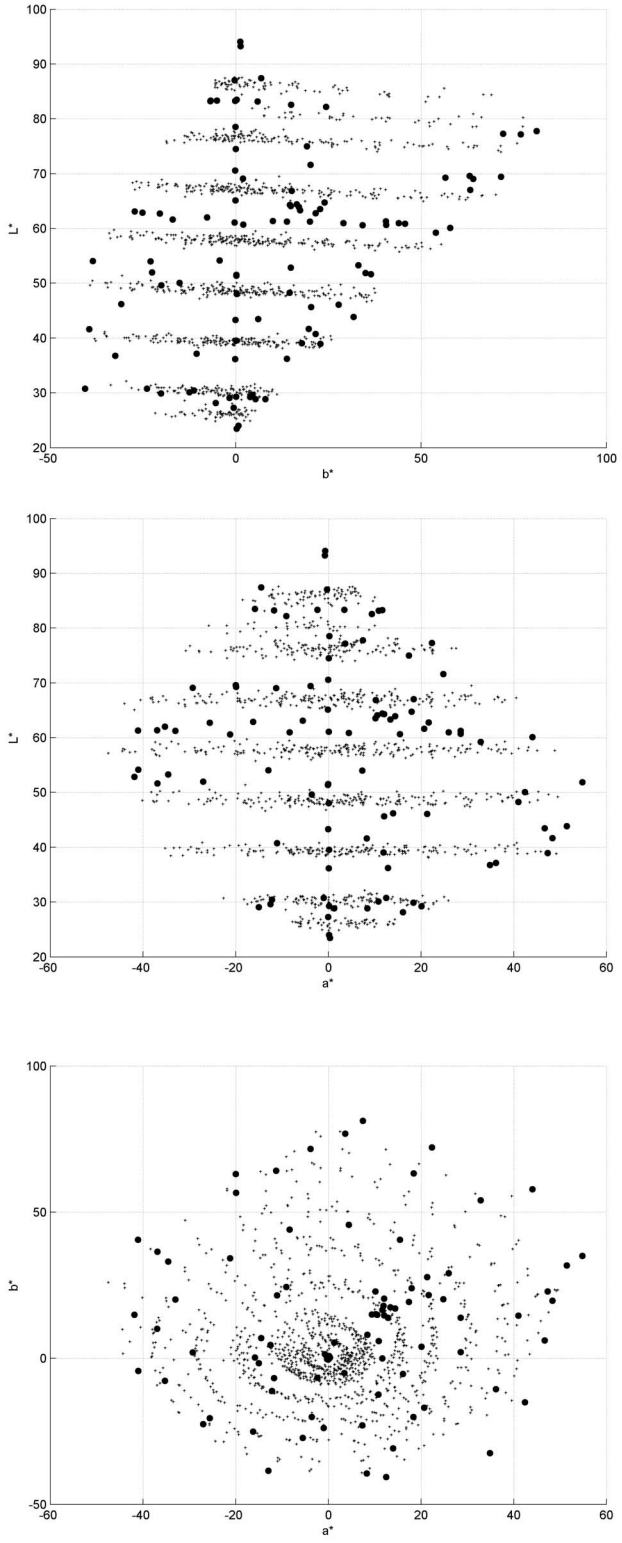

Fig. 8. Distribution of 137 samples of ColorChecker chart and 1269 Munsell color samples in L*a*b* color space.

samples were extracted and used for reconstruction of the spectral behavior of ColorChecker samples, and vice versa.

Table 6 shows the results of recovery in terms of RMS and color difference values. Unlike with PCA method, there are some samples—called out-of-gamut specimens—for which their reflectances could not be reconstructed by interpolation technique. However the recovery results for those samples that lay in the gamut are quite acceptable. The number of out-of-gamut points is related to the source and trial color gamut and the position of each target point in the source LUT space. It would seem reasonable to select a wide color gamut set so as to build a more universal LUT for estimation of the reflectance data of proposed samples to cover the whole color database properly and prevent out-of-gamut points. It is appropriate to mention that ColorChecker SG is usually employed for characterization of digital cameras and scanners, so it contains the most saturated, lightest, and darkest colors to properly cover the color gamut of scanners and digital cameras, as is shown in Figure 8. As a result, scattered samples near the boundaries of the color gamut will not lead to creation of nice tetrahedrals since not enough neighbor points are available.

The number of points in the source database also affects the process time for interpolation. Table 6 shows that the process time decreased when the ColorChecker SG was considered as the source of LUT. In this table, for example, the average time for interpolation of each sample from ColorChecker SG as the source of LUT is $\approx$0.0036 s, while it increased to 0.0223 s when 1269 Munsell samples created the first LUT.

5. CONCLUSION

The linear interpolation method was employed for reconstruction of the reflectance behavior of different groups of color samples, and the results were compared with those from the PCA approach. Effects of several factors such as the color space types, the number of color coordinates, and the applied dataset in the creation of LUTs were investigated. The performance of methods was evaluated by using the average and the maximum values of RMS errors between the reconstructed and actual spectra and the color difference values under different sets of illuminants. The interpolation method yielded more accurate results overall in reconstructed reflectance—especially when the number of inputs was increased—in comparison with the PCA method.

Among three different color spaces, viz., CIEXYZ, CIExyY, and CIELAB, the CIEXYZ was the most suitable space for creation of LUTs. In such conditions the reconstructed and actual samples form a metameric pair. In turn, CIELAB space led to slightly more desirable results with smaller mean color differences for other viewing conditions. Even though color difference values were not zero, this method benefited from more visual uniformity. Inaccurate results in the CIExyY show the importance of scale uniformity of color space in this method. As the x, y,

Table 6. RMS and Color Difference Errors of Recovery Process Using LUT and PCA Methods and Different Databases

	Trial Database		200 Munsell Samples	ColorChecker SG
	Source Database		ColorChecker SG	1269 Munsell Samples
			XYZ_{D65}	XYZ_{D65}
PCA	ΔE		0	0
		Mean	2.2953	2.5479
			1.984	2.8122
	RMS	Mean	0.0337	0.0389
		Maximum	0.0994	0.163
		Variance	0.0002	0.001
	ΔE		0	0
		Mean	1.4333	1.1738
			1.5886	1.7544
LUT	RMS	Mean	0.0253	0.0233
		Maximum	0.1057	0.1133
		Variance	0.0003	0.0005
	Process Time	Total	0.9844	2.5625
		Tessellation	0.2031	0.2969
		Mean	0.0036	0.0223
	Out-of-Gamut Samples		9 from 200	27 from 137

and Y axes in this space are not uniform in scale, the calculated weights do not lead to proper estimation.

Since reconstruction methods that might be used in the field of imaging and multispectral imaging often involve a huge number of color pixels, the processing time was considered as a significant factor. Increasing the dimension and the number of color points in the first space led to longer calculation times.

The effect of the applied dataset in the formation of LUTs was also studied. According to the results, the first space of the LUT should cover all the color space adequately. Although some of the points resulted in out-of-gamut source data, the mean of RMS and ΔE of the predicted points were acceptable.

Finally, interpolation methods capable of combining various color values and color spaces resulted in different accuracies and process time. Therefore, the choice among different types of interpolation models should be based on the final application of the algorithm and process speed.

REFERENCES

1. R. S. Berns, *Billmeyer and Saltzman's Principles of Color Technology*, 3rd ed. (Wiley, 2000).
2. D. Dupont, "Study of the reconstruction of reflectance curves based on tristimulus values: comparison of methods of optimization," Color Res. Appl. **27**, 88–99 (2002).
3. N. Salamati and S. H. Amirshahi, "The comparison between PCA and simplex methods for reflectance recovery," in *Proceedings of AIC Interim Meeting on Color Science for Industry*, Hangzhou, China (2007), pp. 149–152.
4. S. Usui, S. Nakauchi, and M. Nakano, "Reconstruction of Munsell color space by a five-layer neural network," J. Opt. Soc. Am. A **9**, 516–520 (1992).
5. K. Ansari, S. H. Amirshahi, and S. Moradian, "Recovery of reflectance spectra from CIE tristimulus values using a progressive database selection technique," J. Color. Technol. **122**, 128–134 (2006).
6. C. J. Hawkyard, "Synthetic reflectance curves by subtractive color mixing," J. Soc. Dyers Colour. **109**, 246–251 (1993).
7. C. J. Hawkyard, "Synthetic reflectance curves by additive mixing," J. Soc. Dyers Colour. **109**, 323–329 (1993).
8. R. S. Berns and C. J. Hawkyard, "Synthetic reflectance curves," J. Soc. Dyers Colour. **110**, 386–389 (1994).
9. G. Wang, C. Li, and M. R. Luo, "Improving the Hawkyard method for generating reflectance functions," Color Res. Appl. **30**, 283–287 (2005).
10. Y. Zhao and R. S. Berns, "Image-based spectral reflectance reconstruction using the matrix R method," Color Res. Appl. **32**, 343–351 (2007).
11. H. S. Fairman and M. H. Brill, "The principal components of reflectance," Color Res. Appl. **29**, 104–110 (2004).
12. F. Agahian, S. A. Amirshahi, and S. H. Amirshahi, "Reconstruction of reflectance spectra using weighted principal component analysis," Color Res. Appl. **33**, 369–371 (2008).
13. F. Ayala, J. F. Echávarri, P. Renet, and A. I. Negueruela, "Use of three tristimulus values from surface reflectance spectra to calculate the principal components for reconstructing these spectra by using only three eigenvectors," J. Opt. Soc. Am. A **23**, 2020–2026 (2006).
14. N. Attarchi and S. H. Amirshahi, "Reconstruction of reflectance data by modification of Berns' Gaussian method," Color Res. Appl. **34**, 26–32 (2009).
15. A. Shams-Nateri, "Effect of a standard colorimetric observer on the reconstruction of reflectance spectra of coloured fabrics," Coloration Technology **124**, 14–18 (2008).
16. S. Zuffi and R. Schettini, "Reflectance function estimation from tristimulus values," Proc. SPIE **5293**, 222–231 (2003).
17. H.-L. Shen, P.-Q. Cai, S.-J. Shao, and J. H. Xin, "Reflectance reconstruction for multispectral imaging by adaptive Wiener estimation," Opt. Express **15**, 15545–15554 (2007).
18. M. de Berg, M. van Krefeld, M. Overmars, and O. Schwarzkopf, *Computational Geometry: Algorithms and Applications*, 2nd ed. (Springer, 2000).
19. I. Amidror, "Scattered data interpolation methods for electronic imaging systems: A survey," J. Electron. Imaging **11**, 157–176 (2002).
20. P. Green and L. MacDonald, *Color Engineering Achieving Device Independent Colour* (Addison-Wesley, 2002).
21. J. Kasson, W. Plouffe, and S. Nin, "A tetrahedral interpolation technique for color space conversion," Proc. SPIE **1909**, 127–138 (1993).
22. University of Joensuu Color Group, "Spectral Database," http://spectral.joensuu.fi/.