Reconstruction of reflectance spectra using weighted principal component analysis

Article *in* Color Research & Application · October 2008

DOI: 10.1002/col.20431

3 authors:

Farnaz Agahian

Seyed Ali Amirshahi

Norwegian University of Science and Technology

Seyed Hossein Amirshahi

Amirkabir University of Technology

Reconstruction of Reflectance Spectra Using Weighted Principal Component Analysis

Farnaz Agahian,[1] Seyed Ali Amirshahi,[2] Seyed Hossein Amirshahi[1]*

[1] Department of Textile Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran

[2] Department of Electrical Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran

Received 29 August 2007; revised 30 October 2007; accepted 11 December 2007

Abstract: The weighted principal component analysis technique is employed for reconstruction of reflectance spectra of surface colors from the related tristimulus values. A dynamic eigenvector subspace based on applying certain weights to reflectance data of Munsell color chips has been formed for each particular sample and the color difference value between the target, and Munsell dataset is chosen as a criterion for determination of weighting factors. Implementation of this method enables one to increase the influence of samples which are closer to target on extracted principal eigenvectors and subsequently diminish the effect of those samples which benefit from higher amount of color difference. The performance of the suggested method is evaluated in spectral reflectance reconstruction of three different collections of colored samples by the use of the first three Munsell bases. The resulting spectra show considerable improvements in terms of root mean square error between the actual and reconstructed reflectance curves as well as CIELAB color difference under illuminant A in comparison to those obtained from the standard PCA method. © 2008 Wiley Periodicals, Inc. Col Res Appl, 33, 360–371, 2008; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/col.20431

Key words: reflectance; spectrum reconstruction; principal component analysis; weighted principal component analysis; tristimulus values

INTRODUCTION

The spectral data are defined as "fingerprint" of object and provide the most useful information for color specification under different viewing conditions as well as color reproduction algorithms. For example, it has been used as an input of computer color formulation of paints, textiles, plastics, and inks to calculate the concentrations of the required colorants for matching the tristimulus values of the target under a given set of condition. Besides, the reflectance data are critical for prediction of the changes of the appearance of object under different illuminants in computer-aided design applications and also provide suitable input for many general computer graphic applications that require a wavelength-based approach to specify color.

Color spaces such as CIE Yxy or RGB models project the semi infinite-dimensional spectral space to three-dimensional color spaces and allow the surface color information to be represented by a set of tristimulus values. 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.

In the past few years, there has been an enormous interest to extract the spectral data of samples as well as the spectral radiance distribution of light sources from their colorimetric specifications.[1–4] Several researchers have also investigated the modeling of spectral reflectance or spectral power distribution throughout dimensionality reduction techniques.[5–10] The most suitable technique implements the well-known linear model based on principal component analysis (PCA).[11,12] This is a classic technique in data compression, which linearly reduces the dimensionality of the data while

*Correspondence to: Seyed Hossein Amirshahi (e-mail: hamirsha@aut.ac.ir).

preserving as much as possible of the variation present in the dataset. Geometrically, it rotates the coordinate system such that the most variations of the original data can be efficiently represented by using a few numbers of principal components. The method takes advantage of the fact that the reflectance spectra of natural and man-made surfaces are generally smooth function of wavelength over the range which the human visual system is sensitive. Such spectra are strongly correlated and may be represented as the weighted sum of a small number of orthogonal basis functions.[7] These functions can be simply obtained from a suitable set of available spectra by applying the PCA technique. According to literature, the number of basis vectors needed for acceptable recovery of reflectance largely depend on the type of applied datasets and bases that are used.[13,14] Nevertheless, reconstruction of reflectance data from the proposed tristimulus values, in the case of one illuminant-observer condition, limits the numbers of eigenvectors to three to prepare a fully determined equation system. Clearly, this limitation leads to noticeable differences between the actual and estimated spectra. Recently, Ansari *et al*.[15] improved this method to recover the spectral reflectance of surface colors from their color coordinates, more efficiently. In the suggested procedure, by fixing certain criteria based on color difference values between the proposed sample and those in dataset, a series of suitable samples have been preliminary selected from a main dataset. Then, PCA applied to a series of different databases containing the reflectance values of confirmed specimens relating to the particular samples. Ayala *et al*.[16] performed a good reconstruction of surface reflectance spectra using just three eigenvectors. Corresponding to the 10 Munsell hues, they divided the color space into 10 zones and different basis vectors were obtained for each subspace.

This article presents a method that attempts to recover the reflectance data from the tristimulus values by using weighted PCA (wPCA). The method is implemented for recovery of spectral data of 1269 matte Munsell color chips, 237 colored samples of GretagMacbeth ColorChecker DC, and a collection of 200 woolen yarn samples dyed in different shades, from the color coordinates of samples under one illuminant-viewing condition. The performance of the proposed method in the spectral estimation of mentioned samples has been investigated by determination of color difference values under illuminant A and root mean square (RMS) errors between the actual and the reconstructed spectra.

MATHEMATICAL BACKGROUND

Theoretical Basis for Spectral Recovery

The CIE tristimulus values of surface color under a given set of illuminant and observer are simply calculated by Eq. (1):

$$T_j = \int_{400}^{700} S_\lambda r_\lambda q_{j,\lambda} d_\lambda \tag{1}$$

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. Eq. (1) can be rewritten in the matrix form as Eq. (2):

$$\mathbf{T} = \mathbf{A}^{\mathrm{T}}\mathbf{r} \tag{2}$$

where $\mathbf{A}$ is the weight product sets of standard illuminant and observer and the superscript T shows the matrix transpose.

To extract the reflectance data, $\mathbf{r}$, from the color coordinates, a number of approaches have been suggested.[3,17] Estimation of the reflectance functions from the tristimulus values by using the pseudo-inverse algorithm, as shown in Eq. (3), is an immediate and straightforward solution. The performance of this method is poor and in the case of inadequate number on data points yields atypical spiky estimates of spectrum.[4,18]

$$\mathbf{r} = \left(\mathbf{A}^{\mathrm{T}}\right)^{+}\mathrm{T} \tag{3}$$

The "+" sign indicates the pseudo-inverse of the proposed matrix.

The most successful approach to estimation of the spectral reflectance from the CIEXYZ values is based upon applying dimensionality reduction techniques. The algorithm has been formulated by Fairman and Brill.[2] This method exploits 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 Eq. (4):

$$\mathbf{r} \approx \mathbf{v}_0 + \sum_{j=1}^{k} c_j \mathbf{v}_j \tag{4}$$

where $\mathbf{r}$ is the spectral reflectance, c_j is the weight of the jth basis function called as the principal component and produces the least square best fit with that reflectance spectrum, $\mathbf{v}_0$ shows the mean spectral reflectance value of dataset and $\mathbf{v}_j$ is the jth eigenvector. Eq. (4) can be written in the matrix form as follows:

$$\mathbf{r} \approx \mathbf{v}_0 + \mathbf{VC} \tag{5}$$

$\mathbf{C}$ is a column vector of k elements, which contains the principal components, and $\mathbf{V}$ is the first k eigenvectors, so that $\mathbf{V} = [\mathbf{v}_1,\ldots,\mathbf{v}_k]$. Now, the term of reflectance in Eq. (5) can be replaced in Eq. (2) to give Eq. (6):

$$\mathbf{T} = \mathbf{A}^{\mathrm{T}}\mathbf{v}_0 + \mathbf{A}^{\mathrm{T}}\mathbf{VC} \tag{6}$$

where $\mathbf{T}$ is the tristimulus values of a given sample whose reflectance is desired, the $\mathbf{A}^{\mathrm{T}}\mathbf{v}_0$ and the $\mathbf{A}^{\mathrm{T}}\mathbf{V}$ are the vector and matrix containing the tristimulus values of the mean spectrum and the first k (in the case of using one illuminant, $k = 3$) basis functions, respectively. By denoting them as $\mathbf{T}_{v_0}$ and $\mathbf{T}_v$, Eq. (7) can be formed:

$$\mathbf{T} = \mathbf{T}_{v_0} + \mathbf{T}_v C \tag{7}$$

As the values of $\mathbf{T}_{v_0}$ and $\mathbf{T}_v$ are known, the calculation of column vector $\mathbf{C}$ is simply possible:

$$\mathbf{C} = \mathbf{T}_v^{-1}(\mathbf{T} - \mathbf{T}_{v_0}) \tag{8}$$

By applying the column vector $\mathbf{C}$ in Eq. (5) the spectral reflectance could be generated, which assures to match the actual reflectance under a given light source and possess the same tristimulus values $\mathbf{T}$ of the proposed sample.

Weighted Principal Component Analysis

As mentioned earlier, the method of spectrum reconstruction from the tristimulus values has a limitation on the number of principal axes. It means that for the common three-dimensional color spaces the dimension of eigenvector subspace should be kept equal to three and only the first three basis vectors are employed for reconstruction of data. So, it seems that the performance of the recovery process strongly depends on the eigenvector subspace within which the spectral data fall. Therefore, it could be expected that by representing data in the appropriate eigenvector subspace in which its first three principal axes have a large contribution in the construction of spectral variations, a sufficient accuracy is achievable.

It is clear that the number of dimensions that should be taken into account for most of variation of a given spectral data could be decreased if the similarity between the reflectance curves of target and samples in the dataset increases. Garcia *et al.*[10] used the clustering method based on samples' hues to increase the similarity and improve the performance of recovery process. Besides, Ansari *et al.*[15] and also Ayala *et al.*[16] presented the results of their research in this field and offered a selective database method, based on the similarity of samples, according to the color specifications of the proposed sample.

It should be noted that, in the recovery process by the standard PCA approach, all spectral data in the main dataset receive equal treatment. It means that all spectra have equal influence on the formation of the principal axes. In this work, to selectively control the influence of the reflectance data in the reconstruction process, samples attending in the database are weighted by different weights based on their color difference values and the proposed sample. This plan is achievable by the implementation of the weighted PCA (wPCA) method. The principals and some applications of this statistical algorithm have been described in detail by Skocaj *et al.*[19,20]

Let N measured reflectance spectrum with M sample point (wavelength) be aligned in the matrix $\mathbf{R} = [\mathbf{r}_1,\ldots,\mathbf{r}_N]_{M\times N}$. The main goal of the wPCA algorithm is to minimize the weighted squared reconstruction error:

$$e = \sum_{\lambda=400}^{700} w\left(r_\lambda - \left(\mathbf{v}_{0_\lambda} + \sum_{j=1}^{k} c_j \mathbf{v}_{j,\lambda}\right)\right)^2 \rightarrow \text{Min} \tag{9}$$

where w is determined according to the importance of each sample in the dataset in relation to others.

For a given dataset $\mathbf{R}_{M\times N}$ the weights are introduced in a diagonal matrix $\mathbf{W}$ as

$$\mathbf{W} = \begin{bmatrix} w_1 & 0 & \ldots & 0 \\ 0 & w_2 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & w_N \end{bmatrix}_{N\times N} \tag{10}$$

It is clear that if the weight of a given sample is larger than the weights of other samples, the reconstruction error of that sample should be smaller than reconstruction error of other ones.

To extract the principal axes that maximize the weighted variance of the projected data in the eigenvector subspace, the data matrix $\mathbf{R}$ is multiplied with the weighting diagonal matrix $\mathbf{W}$ and then eigenvector analysis is performed on the weighted covariance matrix Σ as

$$\hat{\mathbf{R}} = \mathbf{W}\mathbf{R}^{\mathrm{T}} \tag{11}$$

$$\boldsymbol{\Sigma} = \frac{1}{\sum_{i=1}^{N} W_i} \hat{\mathbf{R}}\hat{\mathbf{R}}^{\mathrm{T}} \tag{12}$$

By definition, the input data should be centered on the weighted mean which is calculated as

$$\mu_i = \left(1 \Big/ \sum_{j=1}^{N} w_j\right) \sum_{j=1}^{N} \hat{R}_{ij}, \quad i = 1, \ldots, M \tag{13}$$

Since the purpose of this work was to provide a three-dimensional eigenvector subspace in which its first three bases contain as much information as possible, the weighting matrix was chosen as

$$\mathbf{W} = \begin{bmatrix} \frac{1}{\Delta E_i + s} & 0 & \ldots & 0 \\ 0 & \frac{1}{\Delta E_i + s} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{\Delta E_i + s} \end{bmatrix}_{N\times N} \tag{14}$$

where ΔE_i refers to CIELAB ΔE^*_{ab} color difference values between the samples which form the original dataset ($i = 1,\ldots, N$) and the sample whose spectrum reconstruction is desirable. s is a very small constant, e.g., $s = 0.01$, which prevents from infinity in the case of $\Delta E_i = 0$. In fact, implementation of weighted version of PCA enables

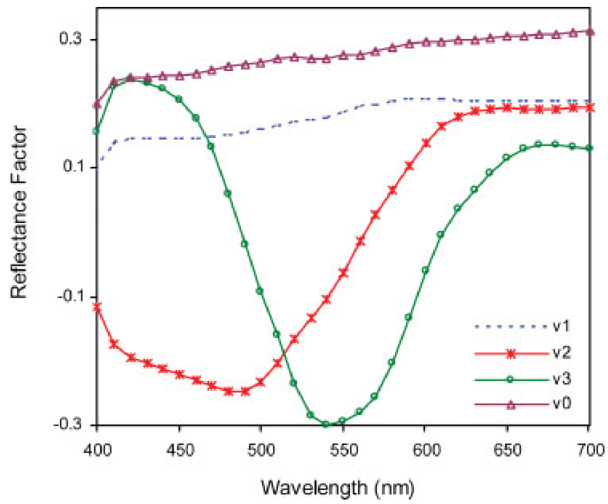

FIG. 1. The mean of reflectance values and the first three eigenvectors of the Munsell dataset. The eigenvectors are fixed and do not vary with sample. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

us to vary the influence of reflectance spectra of the original dataset in estimation of principal axes in a logical way. As Eq. (14) shows, by increasing the color difference value between a given colored sample and a sample in the original dataset, the corresponding weight in matrix **W**, and consequently the influence of that sample on the construction of eigenvector subspace, decreases. Therefore, applying the proposed weighting matrix to the original dataset before implementation of standard PCA, provides the capability to construct a suitable dynamic subspace which varies from one sample to the other one. It is expected that by representing each sample in the mentioned subspace, considerable improvements in recovery process can be obtained.

EXPERIMENTS

In this study we were loaned a dataset consisting of 1269 reflectance spectra of the chips in the Munsell Book of Color—Matte Finish Collection.[21] The spectra set were measured with Perkin Elmer Lambda 18 spectrophotometer

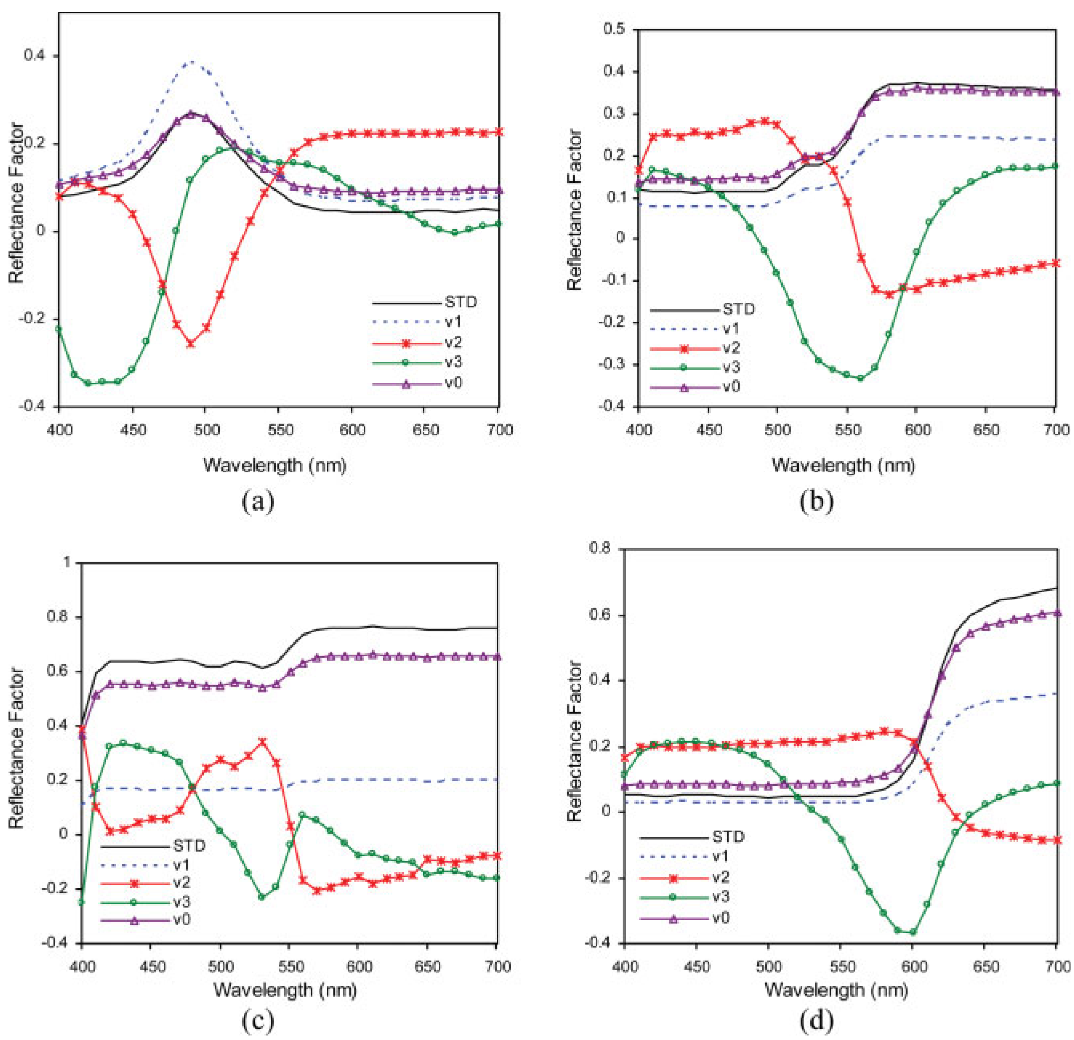

FIG. 2. The first three dynamic eigenvectors of the weighted Munsell dataset obtained by implementation of wPCA method. Targets belong to Munsell color atlas and are shown by solid line. The weighted mean of reflectance data is also shown. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

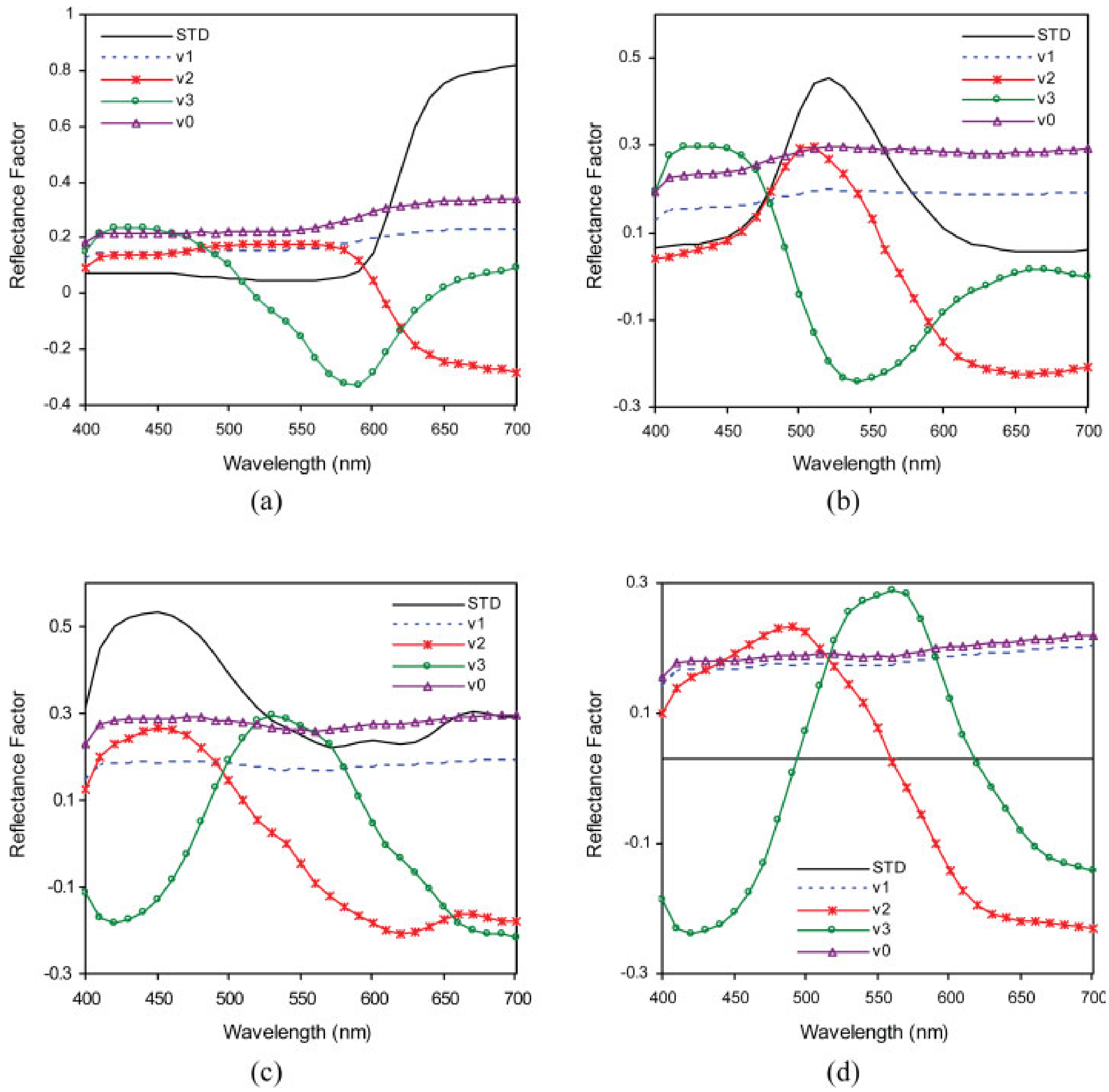

FIG. 3. The first three dynamic eigenvectors of the weighted Munsell dataset obtained by implementation of wPCA method. Targets have been selected from ColorChecker DC chart and are shown by solid line. The weighted mean of reflectance data is also shown. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

and the wavelength range was from 380 to 800 nm with 1-nm interval. In this research, the reflectance data were fixed between 400 and 700 nm at 10-nm intervals. The reflectance values of 237 colored samples of ColorChecker DC from GretagMacbeth as well as a collection of 200 woolen yarn samples dyed in different hues and depths were also measured, personally. They were only considered for the reconstruction sequence. The color measurements for later 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 to 700 nm with specular component excluded.

PROCEDURE

The Munsell data were considered as the original dataset and its bases were used to reconstruct the reflectance functions of Munsell as well as two other datasets. Before analyzing the data, the tristimulus values of all samples were calculated under D65 illuminant and CIE1964 standard observer. Then, we tried to reconstruct the spectral reflectance of all samples of each dataset from their tristimulus values by employing the wPCA method on Munsell data and compare the results by those obtained from standard PCA routine. For this purpose, CIELAB $\Delta E^*_{ab,D65}$ color difference values of all samples of three collections, including Munsell, ColorChecker, and textile materials, with Munsell color chips were calculated and stored for later use. The color difference values were arranged into 1269×1269, 237×1269, and 200×1269 matrices, respectively.

For the reconstruction of each sample by implementation of the wPCA method, the weighting diagonal matrix $\mathbf{W}$ was formed as Eq. (14) and multiplied by the original

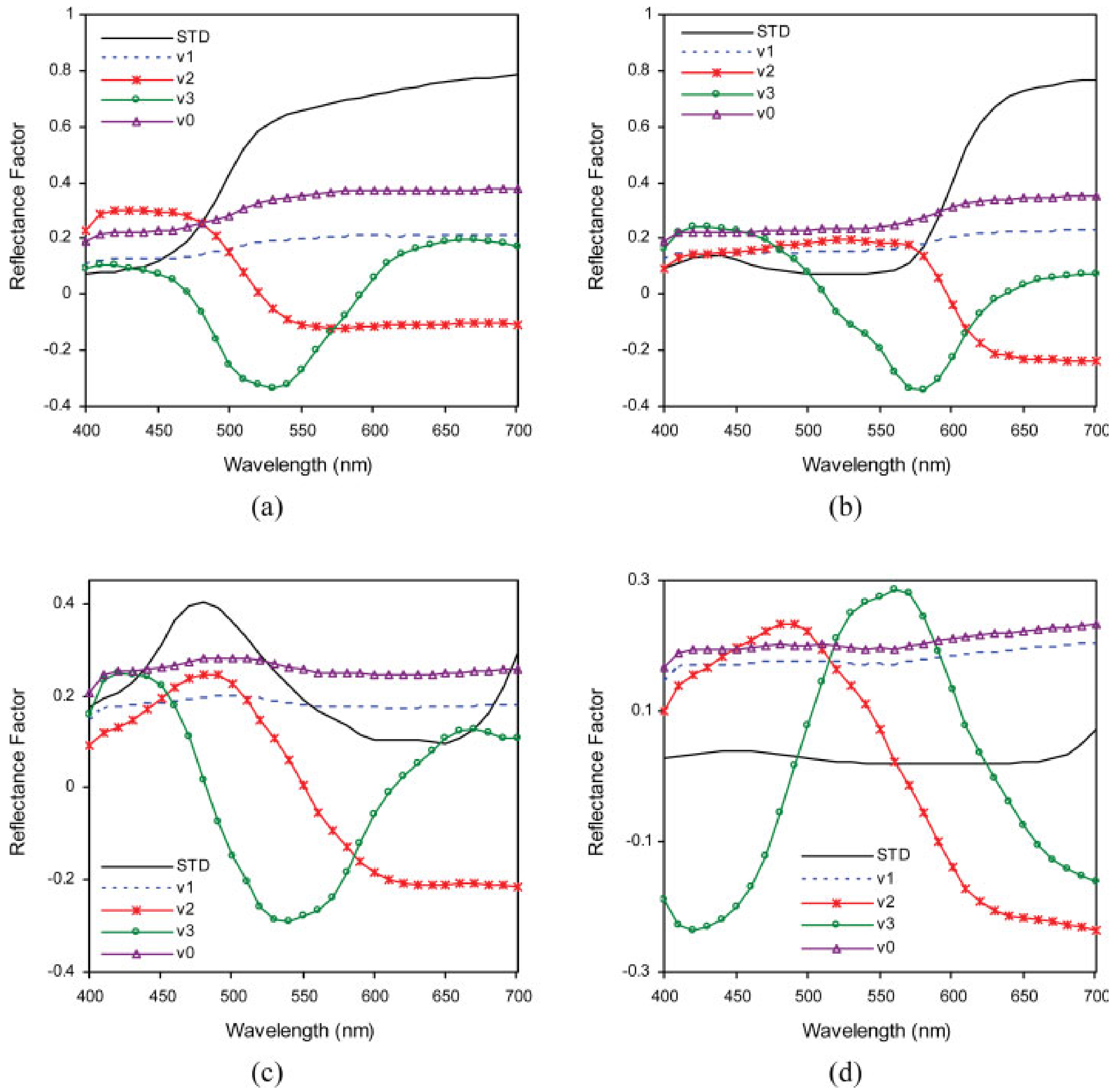

FIG. 4. The first three dynamic eigenvectors of the weighted Munsell dataset obtained by implementation of wPCA method. Targets have been selected from textile samples and are shown by solid line. The weighted mean of reflectance data is also shown. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Munsell dataset. The eigenvectors of weighted covariance matrix were calculated using the Matlab mathematical software package from Mathwork.[22] Figure 1 shows the mean of reflectance values and the first three eigenvectors of original Munsell dataset obtained from application of standard PCA. The achieved eigenvectors for four randomly selected samples of each dataset determined by employing wPCA technique are also illustrated in Figs. 2–4. The weighted mean of reflectance data are also shown in these figures.

Finally, the spectral reflectance of each sample was reconstructed by the use of three eigenvectors and com-

TABLE I. The spectral and colorimetric accuracy of spectral estimation by PCA and wPCA methods.

	PCA				wPCA			
	ΔE_A				ΔE_A			
	Mean	Max	Med	RMS	Mean	Max	Med	RMS
Munsell	1.68	11.17	1.24	0.024	0.01	0.19	0.00	0.000
ColorChecker	1.99	10.80	1.17	0.029	0.92	6.75	0.58	0.018
Textiles	3.20	9.74	2.54	0.073	1.82	5.04	1.66	0.059

TABLE II. Results of spectral recovery of four randomly selected samples of matte Munsell chips, ColorChecker, and textiles by using PCA and wPCA methods.

	PCA		wPCA	
Samples	ΔE_A	RMS	ΔE_A	RMS
Munsell				
a	11.17	0.042	0.004	0.00
b	1.49	0.029	0.004	0.00
c	0.97	0.030	0.011	0.00
d	8.70	0.149	0.001	0.00
ColorChecker				
a	9.64	0.200	3.03	0.098
b	5.94	0.058	1.49	0.017
c	1.44	0.019	0.48	0.016
d	2.36	0.011	0.40	0.001
Textiles				
a	2.91	0.076	2.37	0.048
b	5.77	0.102	0.30	0.023
c	4.46	0.061	2.67	0.052
d	3.13	0.011	0.63	0.009

The actual and the reconstructed reflectance spectra of samples are shown in Figs. 5–7.

pared with actual reflectance curves in addition to those obtained by the standard PCA method.

RESULTS AND DISCUSSIONS

The results of spectral reflectance recovery of three different sets of colored samples from their tristimulus values by using PCA and wPCA techniques are summarized in Table I. The spectral and colorimetric accuracies of methods were quantified by the calculation of RMS error between actual and estimated spectra of samples as well as mean, maximum, and median of CIELAB ΔE^*_{ab} color difference values under illuminant A for CIE1964 standard observer, respectively. Moreover, to demonstrate the performance of the suggested method, the results of the recovery of spectral reflectance for four randomly selected samples from Munsell, ColorChecker, and textile material datasets, whose bases of their eigenvector subspace were shown in Figs. 2–4, are illustrated in Table II and Figs. 5–7.

For better analysis of results, the values of maximum and minimum of color differences between four demon-

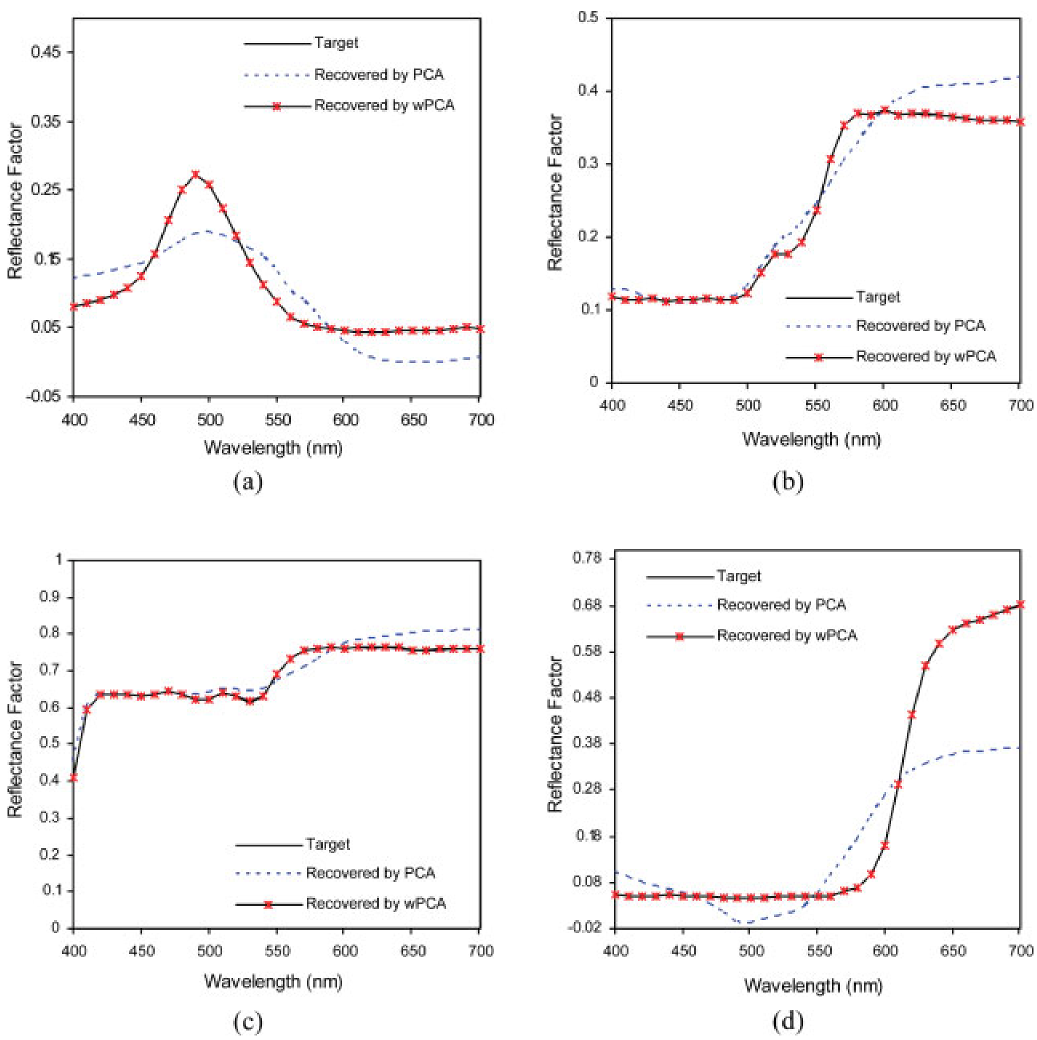

FIG. 5. Results of spectral recovery of four randomly selected samples of matte Munsell chips from their tristimulus values by using PCA and wPCA methods. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

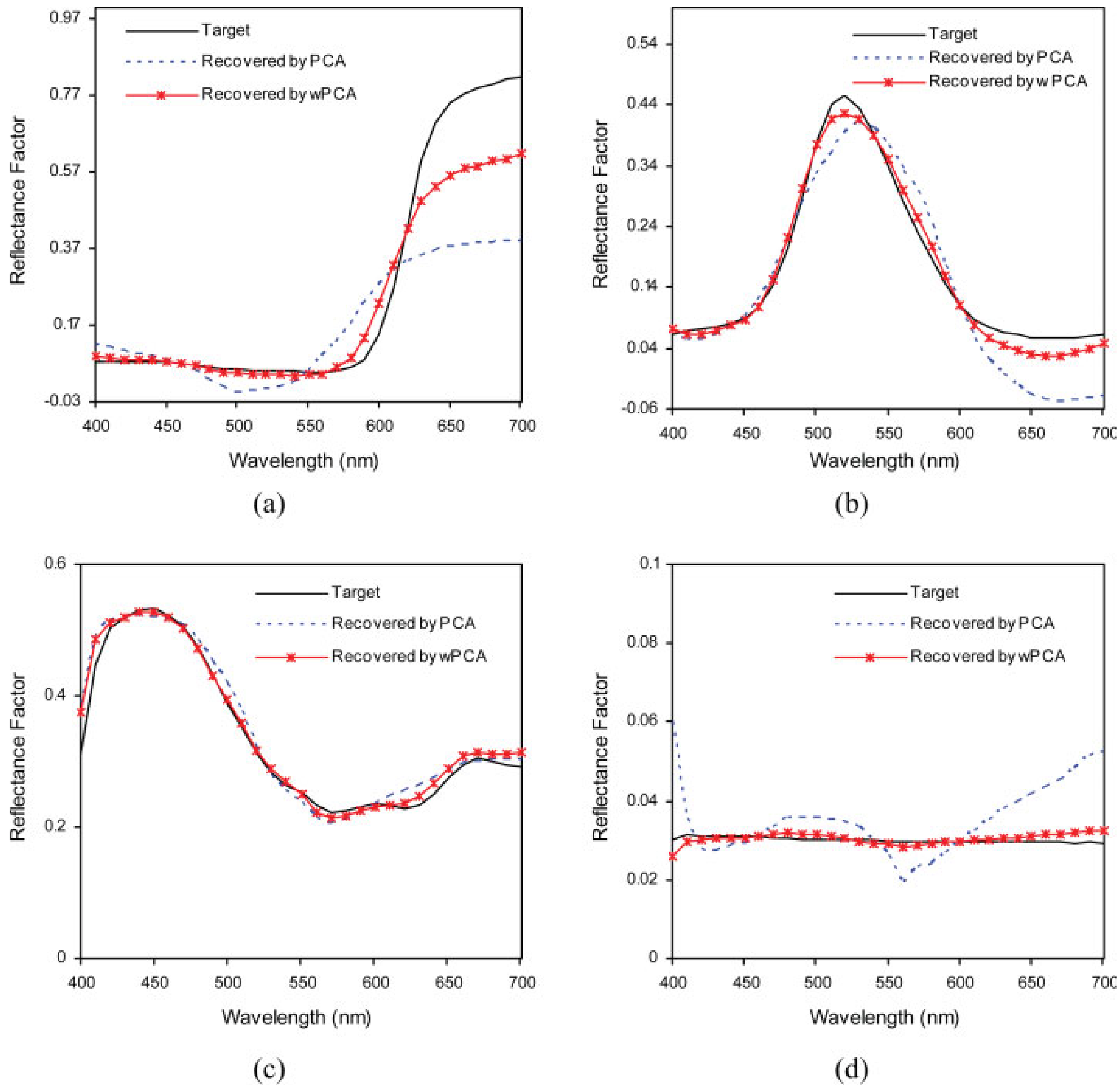

FIG. 6. Results of spectral recovery of four randomly selected samples of ColorChecker DC from their tristimulus values by using PCA and wPCA methods. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

strated samples and Munsell color chips have been reported in Table III. In addition, the weighting factors of these samples, which correspond to cited color differences, are given in this table. Note that the value of constant s has been considered as equal to 0.01.

As shown in Table I, implementation of weighted version of PCA leads to considerable decreasing of RMS errors as well as color difference values in comparison to the standard PCA technique. The explanation is that the wPCA method weights the samples of original dataset according to their distance from the given target and makes a special principal subspace so that the influence of samples which are similar or nearer to the proposed sample are significantly higher in construction of the mentioned subspace.

As shown in Figs. 1–4, the bases of the original Munsell dataset (Fig. 1) are completely different from the first three Munsell eigenvectors which have been extracted by implementation of the wPCA method (Figs. 2–4). As shown in these figures, unlike the standard PCA technique, the extracted eigenvectors by the wPCA method change with the color specifications of target samples. This feature could be considered as a key point in the application of the wPCA method.

According to literature, using as many as eight eigenvectors provide accurate representation of Munsell color chips.[6] It is obvious that the errors of spectral estimation of the Munsell dataset by application of wPCA approach and using just three eigenvectors have been significantly diminished, so that a fairly complete agreement between the target and the estimated spectra has been achieved (Fig. 5).

The performance of the method was also examined in nonideal conditions where the reconstruction of reflectance curves of ColorChecker and textile samples from

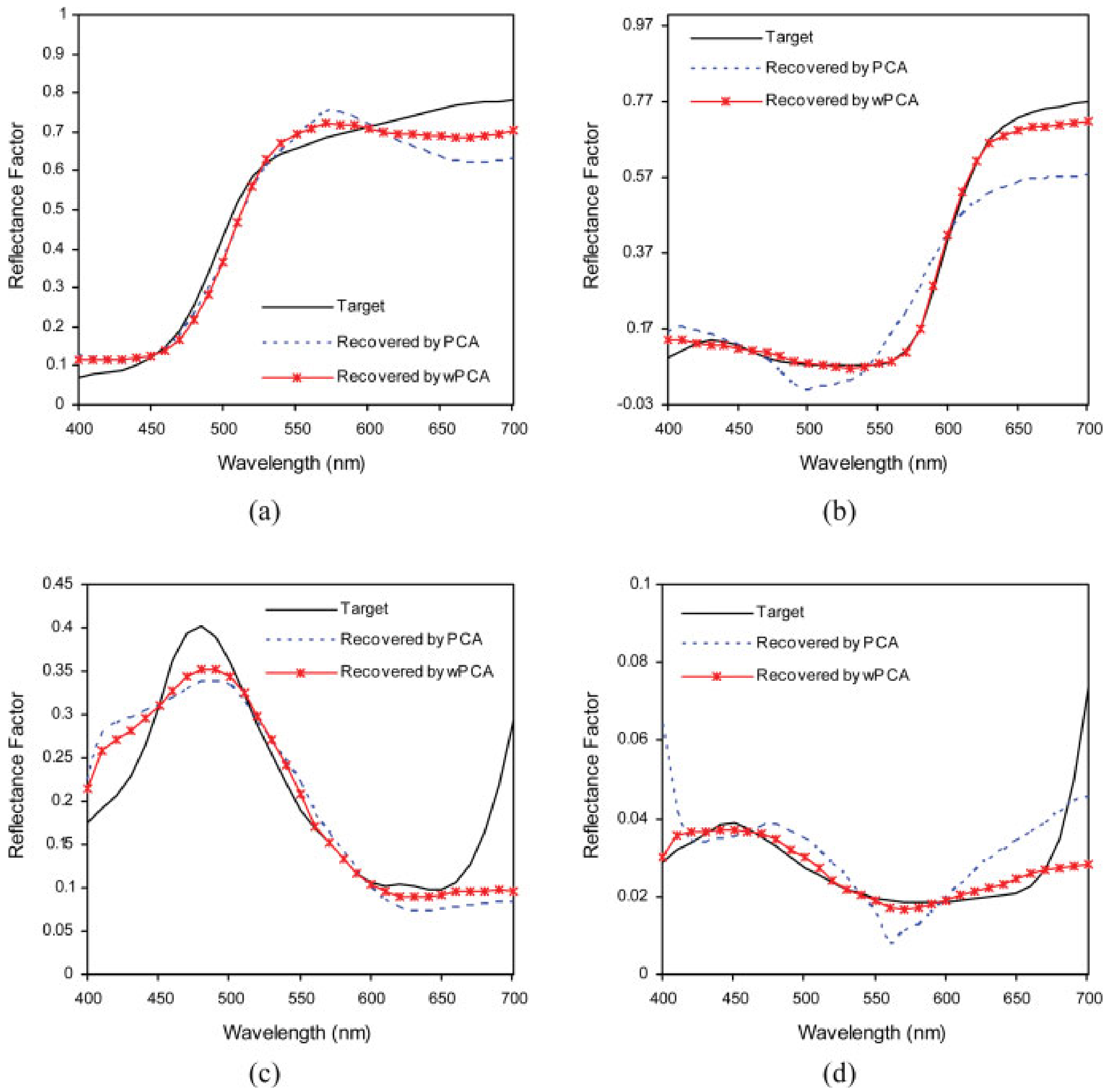

FIG. 7. Results of spectral recovery of four randomly selected textile materials from their tristimulus values by using PCA and wPCA methods. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

their tristimulus values were aimed. As mentioned earlier, in both cases, the reconstruction was carried out by the use of the Munsell bases.

According to Table I, the spectral and colorimetric accuracy of the reconstructed spectra by using three eigenvectors and application of the wPCA method are much better than those obtained by the standard approach with the same number of basis factors. The reflectance curves in Figs. 6 and 7 confirm this achievement.

As shown in Table I, although the application of the wPCA method leads to significant improvements in comparison to the PCA routine, the degree of improvement for the two mentioned data sets is not as good as the Munsell specimens. Obviously, the results for the reconstructed of Munsell spectra are optimal, because the same group of samples were employed to obtain the principal bases and to check the validity of the method. In other words, we used an idealized dataset for the extraction of

TABLE III. The maximum and minimum of color difference values between the four randomly selected target samples and Munsell color chips as well as corresponding weights.

	ΔE_{D65}		Weight	
Samples	Max	Min	Min	Max
Munsell				
a	101.6	0	0.0098	100
b	72.74	0	0.0137	100
c	73.29	0	0.0136	100
d	94.25	0	0.0106	100
ColorChecker				
a	97.27	1.63	0.0103	0.61
b	102.20	5.52	0.0098	0.18
c	106.66	4.01	0.0094	0.24
d	98.53	6.24	0.01	0.16
Textiles				
a	108.85	4.96	0.0092	0.20
b	98.26	2.33	0.0102	0.43
c	100.72	4.25	0.0099	0.23
d	107.89	9.18	0.0093	0.108

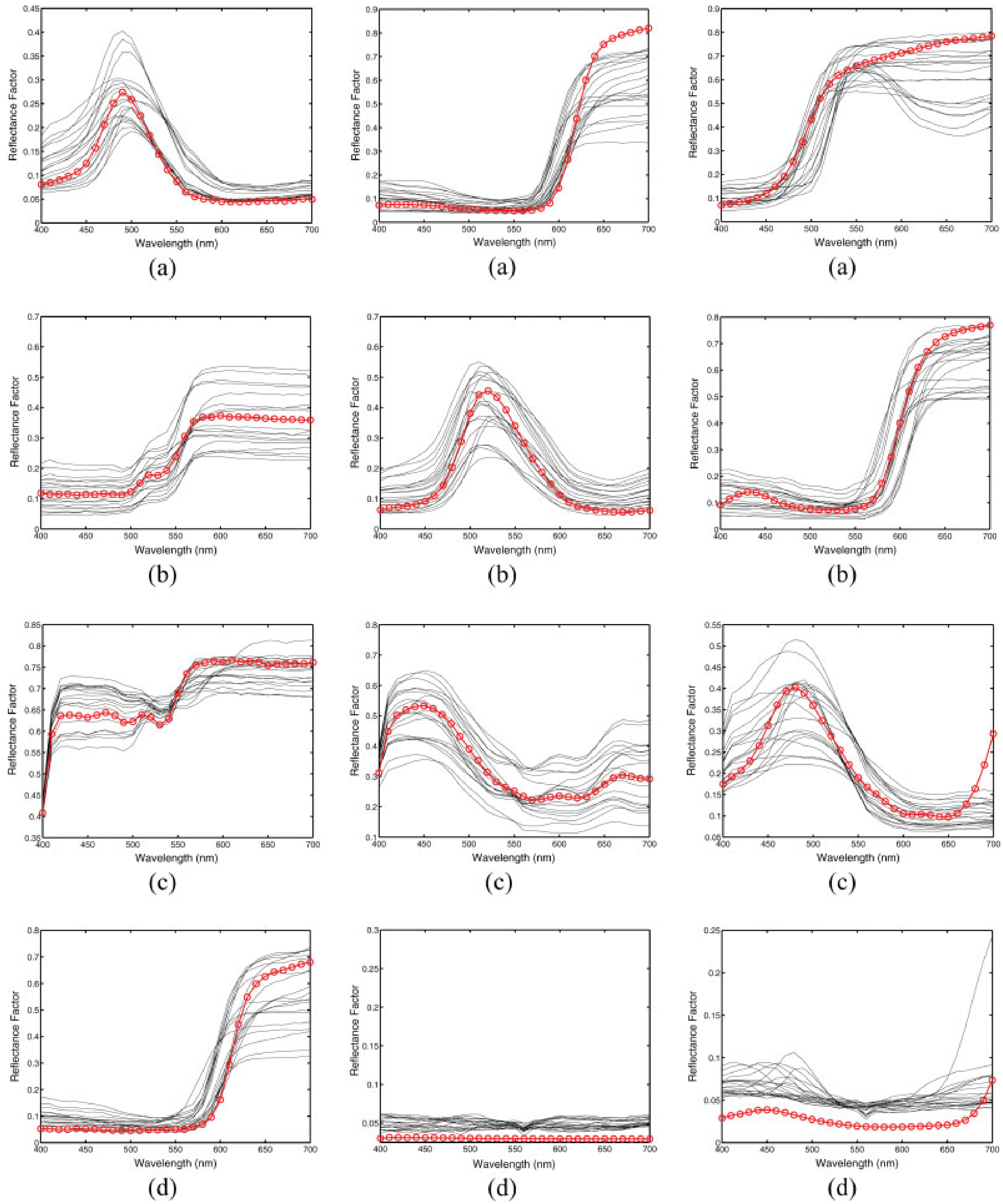

FIG. 8. Reflectance values of 20 Munsell chips which have the least color differences with randomly selected targets (target is shown by red "o" sign). Targets belong to Munsell, ColorChecker, and textiles from left column to right, respectively. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

eigenvectors in this situation; however, in practical conditions, the spectral reflectance that one desires to recover are not known.

Furthermore, in the original Munsell dataset, one of the samples is certainly identical with the proposed sample. This directs to zero color difference value, and consequently, very large value for the corresponding weight ($w_i = \frac{1}{0+s} \rightarrow \infty$) in weighting diagonal matrix $\mathbf{W}$.

On the other hand, as shown in Table III, the color difference values between the Munsell chips and Color-

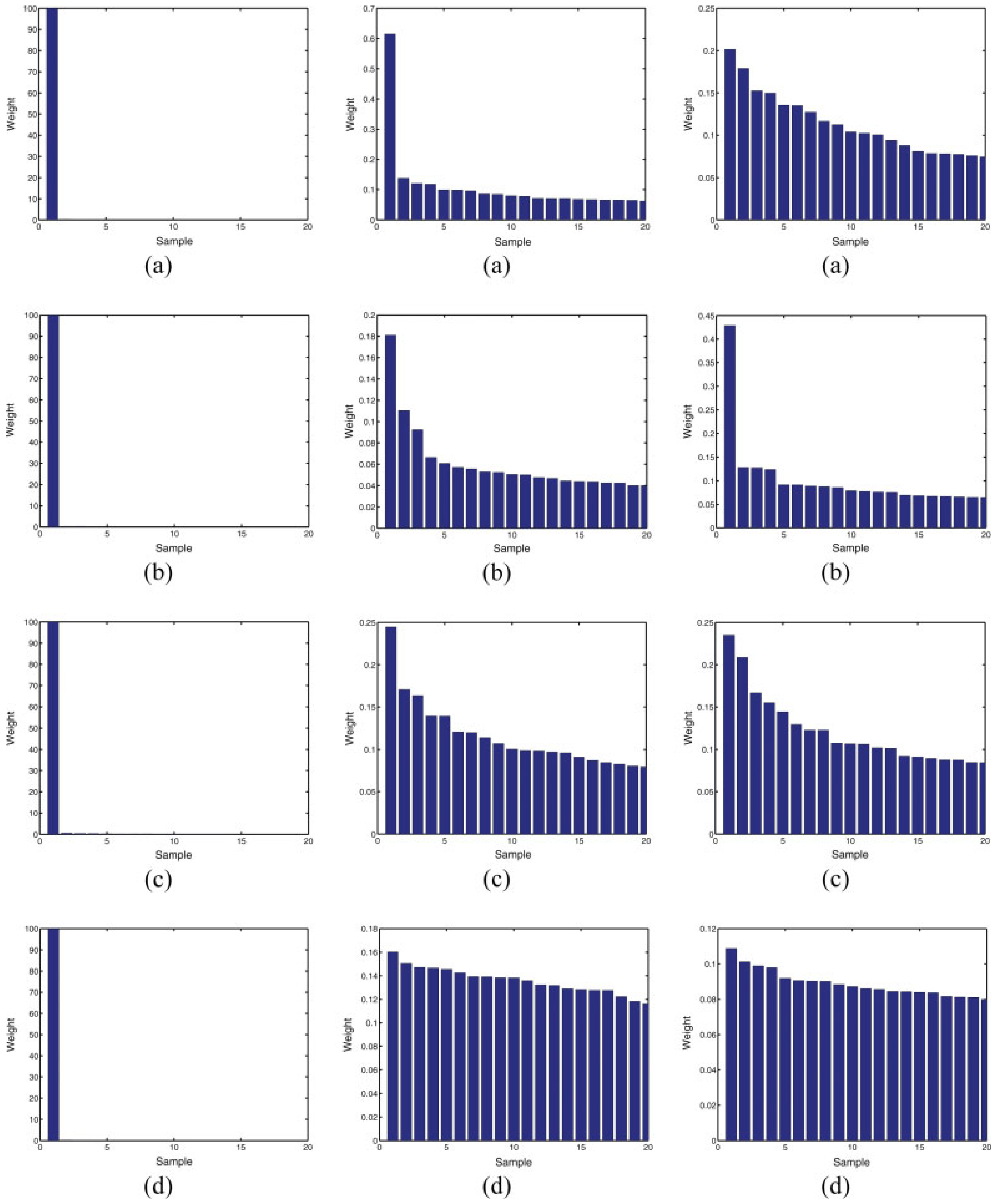

FIG. 9. The weights of 20 Munsell chips which have the least color difference with randomly selected targets shown in Fig. 8. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Checker or textile samples are not too small and do not lead to great values of weights for closer samples.

According to Table I, the results of recovery of ColorChecker are better than textile specimens due to the similarity between the natures of Munsell and ColorChecker samples. It should be noted that the textile samples show a type of incremental behavior at the end part of spectrum, which is different from the dominant spectral behavior of Munsell color chips which are manufactured from coated pigments. Figures 7(c) and 7(d) show this type of deviation between the actual reflectances and those reconstructed by using weighted Munsell bases due to spectral dissimilarity of datasets.

To show the mechanism of wPCA technique in weighting of samples, the spectra of 20 samples that benefited from the least values of color differences with the four target samples, shown in Figs. 5–7, were extracted from the Munsell dataset and are shown in Fig. 8. The weights of these samples for reconstruction of target specimens are also demonstrated in Fig. 9 in the bar form. As shown in Fig. 9, samples with larger values of color differences were gradually omitted during the extraction of eigenvectors with the wPCA technique and the final subspaces were formed by the great influence of closer samples.

CONCLUSION

A method based on implementation of wPCA was presented for recovery of spectral data from tristimulus values. Instead of using a fixed set of eigenvectors for the recovery process, a series of dynamic principal spectra, which form the bases of eigenvector subspace, were extracted for each target sample using eigenvector analysis of the weighted covariance matrix. The color differences between target and samples of original dataset under illuminant D65, i.e., Euclidean distance in CIELAB color space, were considered as a criterion for the determination of weighting factors. Therefore, the influence of samples in dataset on construction of eigenvector subspace has been varied according to their distances to target sample.

The performance of the suggested method was evaluated for recovery of three different sets of colored samples including 1269 matte Munsell color chips, 237 samples of ColorChecker DC, and a collection of 200 woolen yarn samples dyed in different shades. The Munsell bases were used to reconstruct Munsell and two other datasets in standard and weighted approaches.

To estimate the reconstruction accuracy of color spectra, the color difference values under illuminant A as well as RMS errors between the target and the estimated spectrum were calculated. The obtained results showed a significant improvement in colorimetric and spectrophotometric point of view, in comparison to the standard PCA approach.

ACKNOWLEDGMENTS

The authors express their appreciation to Dr. D. Skocaj from University of Ljubljana for sending the results of his study on the weighted principal component analysis, which significantly improved this work.

1. Marimont DH, Wandell BA. Linear models of surface and illuminant spectra. J Opt Soc Am A 1992;9:1905–1913.
2. Fairman HS, Brill MH. The principal components of reflectance. Color Res Appl 2004;29:104–110.
3. Connah D, Westland S, Thomson MGA. Recovering spectral information using digital camera systems. Color Technol 2001;117:309–312.
4. Westland S, Ripamonti C. Computational Colour Science. Chichester: Wiley; 2004.
5. Maloney LT. Evaluation of linear models of surface spectral reflectance with small number of parameters. J Opt Soc Am A 1986;3:29–33.
6. Parkkinen J, Halikainen J, Jaaskelainen T. Characteristics spectra of Munsell colors. J Opt Soc Am A 1989;6:318–322.
7. Jaaskelainen T, Parkkinen J, Toyooka S. Vector-subspace model for color representation. J Opt Soc Am A 1990;7:725–730.
8. Vrhel MJ, Gershon R, Iwan LS. Measurement and analysis of object reflectance spectral. Color Res Appl 1994;19:4–9.
9. Romero J, Garcia-Beltran A, Hernandez-Andres J. Linear bases for representation of natural and artificial illuminants. J Opt Soc Am A 1997;14:1007–1014.
10. Garcia-Beltran A, Nieves JL, Hernandez-Andres J, Romero J. Linear bases for spectral reflectance functions of acrylic paints. Color Res Appl 1998;23:39–45.
11. Jolliffe IT. Principal Component Analysis, 2nd edition (Springer Series in Statistics). New York: Springer-Verlag; 2002.
12. Tzeng DY, Berns RS. A review of principal component analysis and its applications to color technology. Color Res Appl 2005;30:84–98.
13. Hardeberg JY. Acquisition and Reproduction of Color Images: Colorimetric and Multispectral Approaches, PhD Thesis, Paris, France: Ecole Nationale Supérieure des Télécommunications, 1999.
14. Laamanen H, Jaaskelainen T, Parkkinen JPS. Comparison of PCA and ICA in color recognition. Proc SPIE 2001;4197:367–377.
15. Ansari K, Amirshahi SH, Moradian S. Recovery of reflectance spectra from CIE tristimulus values using a progressive database selection technique. Color Technol 2006;122:128–134.
16. Ayala F, Echavarri JF, Renet P. Use of three tristimulus values from surface reflectance spectra to calculate the principal components to reconstruct these spectra by using only three eigenvector. J Opt Soc Am A 2006;23:2020–2026.
17. Dupont D. Study of the reconstruction of tristimulus values: Comparison of methods of optimization. Color Res Appl 2002;27:88–99.
18. Zhao Y, Berns RS. Image based spectral reflectance reconstruction using matrix R method. Color Res Appl 2007;32:343–351.
19. Skocaj D. Robust Subspace Approaches to Visual Learning and Recognition, PhD Thesis, Slovenia: University of Ljubljana, 2003.
20. Skocaj D, Leonardis A, Bischof H. Weighted and robust learning of subspace representations. Pattern Recog 2007;40:1556–1569.
21. Spectral Database, University of Joensuu Color Group. Available at http://spectral.joensuu.fi/.
22. Matlab, Version 6.5. The MathWork Inc., 2002.