Optica Applicata, Vol. XLVI, No. 3, 2016
DOI: 10.5277/oa160304

Wavelength-sensitive-function-based spectral reconstruction using segmented principal component analysis

GUANGYUAN WU[1, 2], XIAOYING SHEN[3], ZHEN LIU[2*], JIANQING ZHANG[2]

[1]College of Printing and Packaging Engineering, Qilu University of Technology, Jinan, 250353, China

[2]College of Communication and Art Design, University of Shanghai for Science and Technology, Shanghai, 200093, China

[3]No. 18, Dahuisi Road, Beijing, 100081, China

[*]Corresponding author: lunaprint@163.com

Spectral images provide richer information than colorimetric images. A high-dimensional spectral data presents a challenge for efficient spectral reconstruction. In conventional reconstruction methods it is very difficult to obtain good spectral and colorimetric accuracy simultaneously. In this paper, a segmented principal component analysis (SPCA) method and a weighted segmented principal component analysis (wSPCA) method are proposed for efficient reconstruction of spectral color information. The methods require, firstly, partitioning the complete spectrum of wavelengths into two subgroups, considering the sensitivity of human visual system. Then the classical principal component analysis (PCA) carried out each subgroup of data separately. The results indicated that the spectral and colorimetric accuracy of the SPCA and wSPCA outperformed the PCA and weighted PCA, and wSPCA clearly retained more color visual information.

Keywords: spectral reconstruction, wavelength-sensitive function, segmented principal component analysis.

1. Introduction

The spectral reflectance can be called the object "fingerprinting" that accurately carries the fundamental color information, so spectral color information could match originals under arbitrary illuminants and observers. It is highly useful in various applications, such as print inspection, image color reproduction, art paintings and image classification [1–4]. However, high-dimensional spectral data need large storage space and computational complexity, so a significant effort is necessary for data compression or dimensionality reduction in such spectral color information. Consequently, more accurate reflectance reconstruction will become the key technology in multispectral images.

Since the reflectance spectra of natural spectral surfaces and most nonfluorescent dyes are mostly smooth spectral functions and they are strongly correlated across neighborhood spectral regions, the spectral reflectance can be adequately represented by a few numbers of the orthogonal basis vectors extracted from the dataset [2, 5]. Relying on this observation, multivariate statistical analysis methods such as the principal component analysis (PCA) can be the most efficient dimensional reduction methods for minimizing the error of spectral reconstruction. In color technology and science, PCA has become a standard method for reducing dimensionality of the data and minimizing the reconstruction error for over 50 years. In 1964, COHEN [6] applied PCA on a subset of the *Munsell Book of Color*, using only three principal components to represent 150 spectral reflectances. From then on, numerous papers state that the PCA has been used extensively to analyze different spectral datasets [7–10]. However, the PCA-based reconstruction process has treated equally the entire spectral reflectance along different wavelengths, which could not well reflect the human visual system. This is because the human eyes usually have different sensitives for different wavelengths. For this reason, the weighted version of PCA (wPCA), considering the wavelength-sensitivity function (WSF) of human visual system, was also proposed. LAAMANEN *et al.* [11] presented a wPCA-based method (wPCA$_1$) for the compression and reconstruction of spectral color information, which applied an appropriate weight function on spectral data before forming the correlation matrix and calculating the eigenvector basis. GUANGYUAN WU *et al.* [12] proposed a wPCA-based reconstruction method (wPCA$_2$), which used PCA to obtain the ordinary eigenvectors calculated from the unweighted spectral dataset and determined a proper weighting function to execute the weighted reconstruction of spectral color data. The wPCA is to attain much more reconstruction accuracy at wavelengths where the sensitivity of human vision is higher, which will improve the color reproduction accuracy in color technology and science. It is clear that the choice of weight function is arbitrary, AGAHIAN *et al.* [13] demonstrated that seven different weight functions involve the sensitivity of human visual system but each shows its own characteristic. And yet, in fact, the weighted function involving color-matching functions well reflects the brightness information and chromatic information of color. Recently, LAAMANEN *et al.* [11] presented two different weighted functions, one of which was formed as a combination of the CIE 1931 color -matching functions. JIANDONG TIAN and YANDONG TANG [14] showed the WSF, which generated by adding the three color matching functions. GUANGYUAN WU *et al.* [12] proposed the weighted function, which can be attained by the square root of arbitrary weight function that includes the CIE 1931 XYZ color-matching function. However, wPCA clearly improves the color reproduction accuracy, but fails to the spectral reproduction accuracy under different weight functions.

To obtain good spectral and colorimetric accuracy simultaneously, segmented principal component analysis (SPCA) method and weighted segmented principal component analysis (wSPCA) method are proposed in this study for reconstruction of spectral color information. First, the methods require partitioning of the complete spectrum of wavelengths into two subgroups, considering the sensitivity of human visual system. Then, the classical PCA is carried out in each subgroup of data separately.

2. Theoretical background

The spectral dataset can be represented adequately by a few numbers of the orthogonal basis vectors with a minimum mean square error of the residual $\min\|\mathbf{R} - \hat{\mathbf{R}}\|_2^2$, where $\mathbf{R} = [r_1, r_2, \ldots, r_m]^T$ and $\mathbf{R} = [\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_m]^T$ are two matrices that involve original and reconstructed spectral vectors, respectively. The solution of $\min\|\mathbf{R} - \hat{\mathbf{R}}\|_2^2$ can be usually generated by a PCA. A set of spectral vectors $r_i \in R^n$ $(i = 1, \ldots, m)$ can be represented by

$$r_i = \sum\nolimits_{j=1}^{n} u_j v_{ij} + \bar{r} \quad \text{for} \quad m \geq n$$

where u_j – the orthogonal basis vectors, v_{ij} – the coefficient of the j-th basis vector, $\bar{r}$ – the mean spectral reflectance value of dataset. Spectral reflectance can be approximated well to use only a few basis vectors

$$\hat{r}_i = \sum\nolimits_{j=1}^{d} u_j v_j + \bar{r} \quad \text{for} \quad d < n.$$

If we define the matrices $\mathbf{U} = [u_1, u_2, \ldots, u_n]^T$, $\mathbf{V} = [v_{n1}, v_{n2}, \ldots, v_{nm}]$, $\hat{\mathbf{U}} = [u_1, u_2, \ldots, u_d]^T$, $\hat{\mathbf{V}} = [v_{d1}, v_{d2}, \ldots, v_{dm}]$ and $\mathbf{h} = [1, 1, \ldots, 1]_m^T$, matrices $\mathbf{R}$ and $\hat{\mathbf{R}}$ can be expressed by:

$$\mathbf{R} = \mathbf{U}\mathbf{V} + \mathbf{h} \otimes \bar{\mathbf{r}}$$

$$\hat{\mathbf{R}} = \hat{\mathbf{U}}\hat{\mathbf{V}} + \mathbf{h} \otimes \bar{\mathbf{r}}$$

where sign $\otimes$ denotes the tensor product of vectors.

Since the classical PCA is a global transformation, it could not preserve local useful spectral color information to obtain a good spectral reconstruction, and therefore might not reflect the characteristics of all the spectral reflectance. So with the classical PCA and the wPCA it is very difficult to obtain good spectral and colorimetric accuracy simultaneously [11, 12, 15]. Spectral reconstruction using a SPCA could be useful. This is because the variances of the bands in each subgroup are much higher than the whole bands, and SPCA improves the performance of PCA [4, 16]. In addition, the PCA is the well-known linear model that equally treats spectral reflectance over the whole wavelength, but human visual system is a highly nonlinear system. For these reasons, we present two segmented PCA-based methods for the reconstruction of spectral color information, considering the human visual system.

The complete set of bands is segmented based on the following considerations. Since human visual system usually has different sensitivities over different wavelengths, and CIE XYZ color matching functions involve brightness information and chromatic information [11, 14], WSF can be generated by combination of color matching functions. If a whole spectrum of wavelengths is partitioned into several subgroups at wavelength where WSF has low sensitivity, the influence of color difference will be minimized because the junction of two subgroups could easily present atypical spikes (as shown in Fig. 1). This idea leads to the proposed SPCA method discussed below.

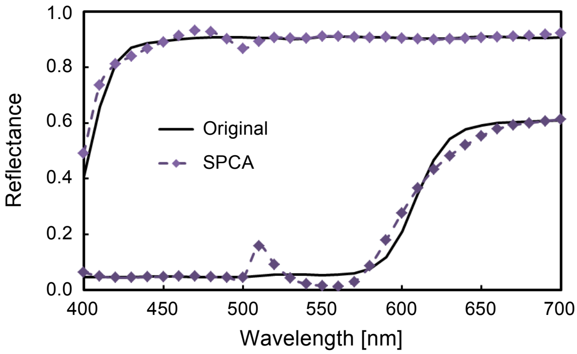

Fig. 1. Two examples of spectral reconstruction by the SPCA method.

The complete spectrum of wavelengths (400–700 nm) is first divided into two subgroups. Figure 2 shows the WSF, generated by adding three matching functions, and two subgroups of wavelengths. The PCA is then carried out in each subgroup of data separately.

It has been observed previously that when the wavelengths where WSF has high sensitivity are reconstructed accurately, more color information is retained and better

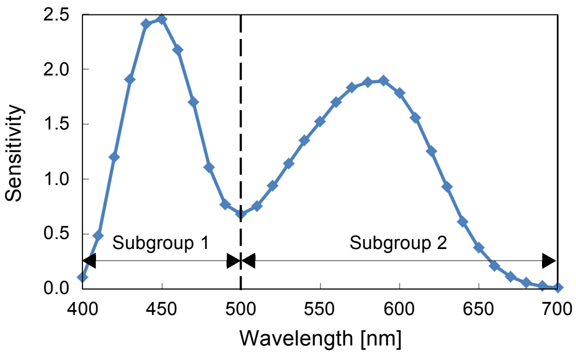

Fig. 2. The WSF generated by adding three matching functions and two subgroups of wavelengths.

color reproduction performance is achieved through the reconstruction process. The purpose of wPCA, considering the wavelength sensitivity of human visual system, is to improve the color reproduction accuracy in color technology and science. The wPCA is noted that after spectral reproduction, the same weight function **W** can be separated from the weighted spectral data to achieve representatives of the reconstructed spectral curves [11, 12],

$$\hat{\mathbf{R}} = \hat{\mathbf{U}}(\mathbf{W}\hat{\mathbf{U}})^{-1}\mathbf{W}(\mathbf{R} - \mathbf{h} \otimes \bar{\mathbf{r}}) + \mathbf{h} \otimes \bar{\mathbf{r}}$$

The weight function **W** is a diagonal matrix with the main diagonal of the values involved in WFS. Because WFS involves some very small values, it is necessary to add a constant function (*i.e.*, 1) to avoid computational instability when inverting values of the weight function [11]. Since human visual system usually has different sensitivities over different wavelengths in each subgroup, the wSPCA is similarly feasible.

3. Experiments and discussion

To evaluate the performance of the proposed SPCA and wSPCA methods for spectral reconstruction of spectral database, the PCA and wPCA methods (wPCA$_1$, wPCA$_2$), SPCA and wSPCA were implemented for comparison of the colorimetric accuracy and spectral accuracy. First, the spectra of *Munsell Atlas* were selected as training samples. The mixed spectrum sets (including *Munsell Atlas*, *ColorChecker 24*, *Acrylic Paints* and *NCS Atlas*) were employed as testing samples [8, 17, 18]. In addition, all the spectra and illuminants were sampled at 10 nm intervals between 400 and 700 nm. The goodness-of-fit coefficient (GFC) and CIELAB color differences under illuminants D65 and F2 between the original and reconstructed spectra of the testing samples were calculated to compare the five different methods. The GFC has values in the range [0, 1], GFC $\geq$ 0.999 and GFC $\geq$ 0.9999 represent good and excellent spectral matches, respectively.

Tables 1 and 2 show the mean CIELAB color differences and the maximum CIELAB color differences for the different numbers of the orthogonal basis vectors under different CIE illuminants. The tables also show the standard deviation of color difference statistics of the five methods. The standard deviations could represent the robustness of the five methods: the smaller the standard deviations, the more robust performance of the spectral reconstruction method under predefined viewing conditions. Figure 3 shows graphical representations of mean color differences to reconstruct the mixed spectrum sets under different CIE illuminants. As the results show, the colorimetric performance orders of the five methods are wSPCA, SPCA, wPCA$_2$, wPCA$_1$ and PCA. It is mainly due to the preserving of spectral color information that SPCA and wSPCA preserve more local information than the PCA and wPCAs, which minimizes the loss of color information in the reconstruction process. In addition, the colorimetric rep-

Table 1. The colorimetric reconstruction accuracy with five different methods, one calculated for non-weighted method (NW), two calculated for weighed PCA ($wPCA_1$ and $wPCA_2$), one calculated for SPCA and one calculated for weighted SPCA (wSPCA) under the condition of CIE D65 illuminant and CIE 1931 standard observer.

	Mean ΔE_{ab}					Max ΔE_{ab}					Standard deviation				
Components	NW	$wPCA_1$	$wPCA_2$	SPCA	wSPCA	NW	$wPCA_1$	$wPCA_2$	SPCA	wSPCA	NW	$wPCA_1$	$wPCA_2$	SPCA	wSPCA
3	7.4551	5.4273	5.4127	2.1383	0.7130	42.2028	31.2093	31.3082	18.1801	6.8867	7.5031	5.9398	5.9808	2.6546	0.9488
4	3.3029	1.5137	1.4567	0.5951	0.5473	31.5808	10.4223	9.5084	5.7131	5.4903	3.0285	1.5755	1.4629	0.6469	0.7451
5	1.5471	1.1947	1.1196	0.3425	0.2462	28.8372	8.7630	7.2604	4.2238	2.1037	1.1568	1.2704	1.1337	0.5423	0.3330
6	1.5786	0.9793	0.9045	0.1080	0.0496	28.8404	6.3658	6.0833	1.5236	0.6828	1.1941	0.9883	0.9269	0.1771	0.0550
7	0.7734	0.4860	0.4985	0.0251	0.0462	6.6279	4.2114	3.6736	0.2099	0.6767	0.7749	0.5387	0.5165	0.0254	0.0534
8	0.5928	0.2681	0.3343	0.0186	0.0312	4.0334	2.2671	2.7201	0.1487	0.2777	0.4905	0.3183	0.3850	0.0163	0.0324
9	0.5609	0.2367	0.2441	0.0140	0.0128	4.0642	1.5629	1.6903	0.1667	0.2431	0.5006	0.2000	0.2439	0.0133	0.0125
10	0.4119	0.0735	0.0564	0.0052	0.0056	3.8236	0.6417	0.4587	0.1194	0.0970	0.4293	0.0820	0.0597	0.0093	0.0078

Table 2. The colorimetric reconstruction accuracy with five different methods, one calculated for non-weighted method (NW), two calculated for weighed PCA ($wPCA_1$ and $wPCA_2$), one calculated for SPCA and one calculated for weighted SPCA (wSPCA) under the condition of CIE F2 illuminant and CIE 1931 standard observer.

	Mean ΔE_{ab}					Max ΔE_{ab}					Standard deviation				
Components	NW	$wPCA_1$	$wPCA_2$	SPCA	wSPCA	NW	$wPCA_1$	$wPCA_2$	SPCA	wSPCA	NW	$wPCA_1$	$wPCA_2$	SPCA	wSPCA
3	5.5841	3.1018	3.1445	2.8315	1.3039	29.8920	17.6460	17.9716	28.5624	12.4292	5.2742	3.2163	3.2620	3.7577	1.6969
4	3.7869	1.8356	1.7893	1.2352	0.1950	29.2006	13.7615	13.2923	13.0198	1.2981	3.8649	2.0865	1.9419	1.7078	0.1578
5	1.9067	0.6929	0.5226	0.9259	0.2104	26.9666	5.2330	4.4110	10.5980	2.0500	1.5263	0.7214	0.4575	1.3792	0.2361
6	1.7886	0.3709	0.4263	0.1163	0.0792	26.9778	4.3183	4.3788	0.8733	0.5454	1.4437	0.3472	0.3747	0.0938	0.0676
7	1.3127	0.1877	0.2410	0.1111	0.0785	12.4995	2.0234	2.1300	0.8866	0.5419	1.5015	0.1790	0.2556	0.0917	0.0632
8	0.6098	0.1361	0.1286	0.0556	0.0387	4.2464	1.6567	1.2025	0.5425	0.3282	0.5019	0.1150	0.1053	0.0518	0.0341
9	0.6210	0.1429	0.1283	0.0355	0.0238	4.2430	1.6594	1.3794	0.9837	0.4326	0.5055	0.1336	0.1121	0.0292	0.0181
10	0.4037	0.1047	0.1011	0.0235	0.0202	3.9401	0.6630	0.7149	0.1945	0.1373	0.3925	0.0844	0.0803	0.0211	0.0159

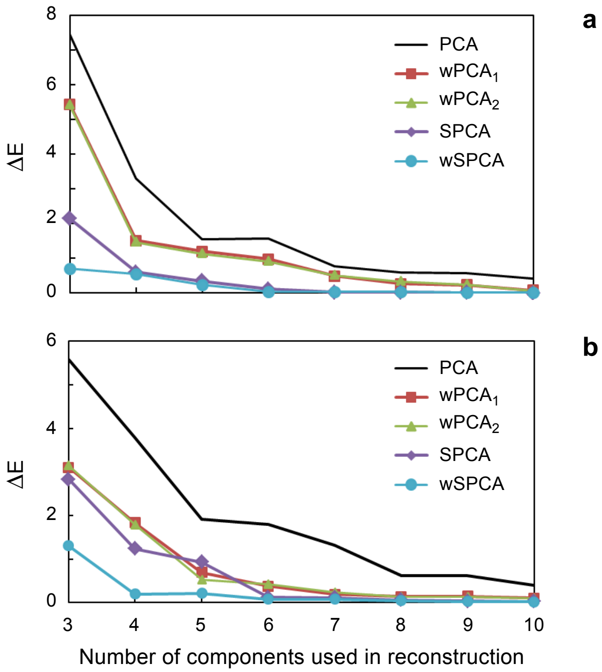

Fig. 3. Graphical representation of mean color differences for the mixed spectrum sets in Table 1 (**a**) and in Table 2 (**b**).

resentation accuracy of wSPCA performed better than SPCA. The main reason is that the wSPCA achieves more accurate reconstruction at high sensitivity wavelength of human visual system.

Spectral reconstruction accuracy was estimated by using the GFC between the original and reconstruction spectra. Table 3 shows the minimum of the GFC values, the mean of the GFC values for different numbers of the orthogonal basis vectors used in the reconstruction of the mixed spectrum sets. Also, percentage of testing samples with the GFC values greater than 0.999 was recorded in each case, where GFC $\geq$ 0.999 represents the condition for good spectral matches. Figure 4 shows graphical representations of the reconstructed results for the mixed spectrum sets used in five different methods, and average spectral residuals between the reconstructed and original spectra in the mixed spectrum sets with three orthogonal basis vectors. It is easy to find that the spectral reconstruction accuracy of the weighted reconstruction method is less than that of the non-weighted reconstruction method. This result presents a strong agreement with the conclusion made by numerous previous studies [11–13], and is due to cause spectral representation errors to increase in low sensitivity wavelength.

Figure 5 shows the example of spectral reconstructions of one sample from the mixed spectrum sets. It can be seen from Fig. 5 that the middle part of spectrum obtained with the weighted reconstruction method is more accurate than that obtained with the non-weighted reconstruction method, but the both ends of the spectrum are

T a b l e 3. The spectral reconstruction accuracy with five different methods, one calculated for non-weighted method (NW), two calculated for weighed PCA ($wPCA_1$ and $wPCA_2$), one calculated for SPCA and one calculated for weighted SPCA (wSPCA).

	Min GFC					Mean GFC					Samples where GFC > 0.999 [%]				
Components	NW	$wPCA_1$	$wPCA_2$	SPCA	wSPCA	NW	$wPCA_1$	$wPCA_2$	SPCA	wSPCA	NW	$wPCA_1$	$wPCA_2$	SPCA	wSPCA
3	0.7714	0.7471	0.7516	0.8194	0.7958	0.9800	0.9781	0.9788	0.9927	0.9917	8.0913	7.0258	7.3770	39.0047	34.6019
4	0.8072	0.7639	0.7785	0.8965	0.8841	0.9905	0.9886	0.9895	0.9960	0.9953	17.5293	16.2061	15.5621	59.6136	55.9368
5	0.8482	0.8207	0.8218	0.9356	0.9268	0.9934	0.9915	0.9925	0.9982	0.9978	27.6230	24.6487	24.3326	75.0820	72.4707
6	0.8482	0.8204	0.8214	0.9684	0.9638	0.9964	0.9948	0.9959	0.9993	0.9992	37.7963	31.2295	32.0960	88.2904	87.1077
7	0.9354	0.8928	0.9237	0.9877	0.9865	0.9973	0.9956	0.9967	0.9998	0.9997	55.1171	40.9251	47.4473	96.0304	95.2927
8	0.9635	0.9164	0.9573	0.9884	0.9875	0.9983	0.9977	0.9979	0.9998	0.9998	65.2927	52.2600	58.5480	97.5644	97.1897
9	0.9692	0.9571	0.9650	0.9915	0.9907	0.9991	0.9986	0.9989	0.9999	0.9999	81.3349	71.5457	78.7939	98.1382	98.0445
10	0.9701	0.9569	0.9650	0.9920	0.9911	0.9993	0.9990	0.9992	0.9999	0.9999	86.7330	83.2319	84.9883	98.3372	98.1616

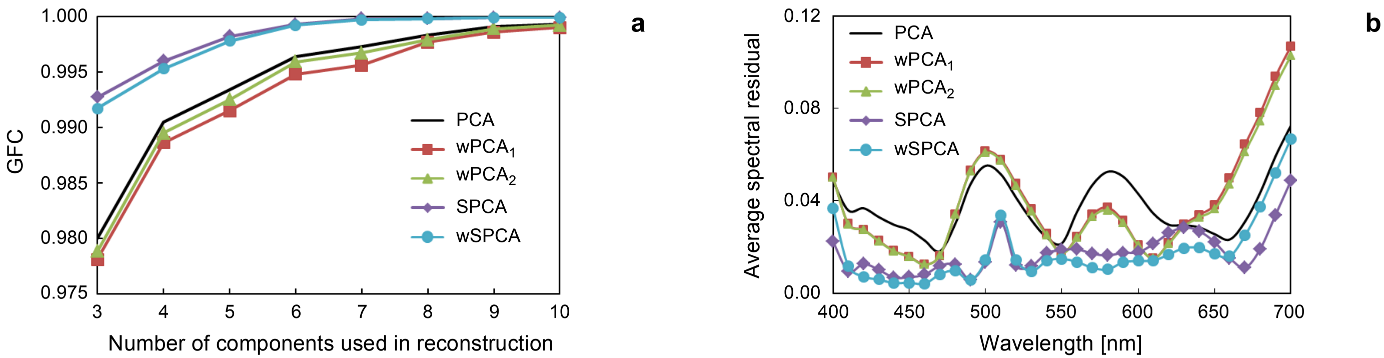

Fig. 4. Graphical representation of average spectral differences for the mixed spectrum sets in Table 3 (**a**). Average spectral residuals between reconstructed and original spectra for the mixed spectrum sets with three orthogonal basis vectors (**b**).

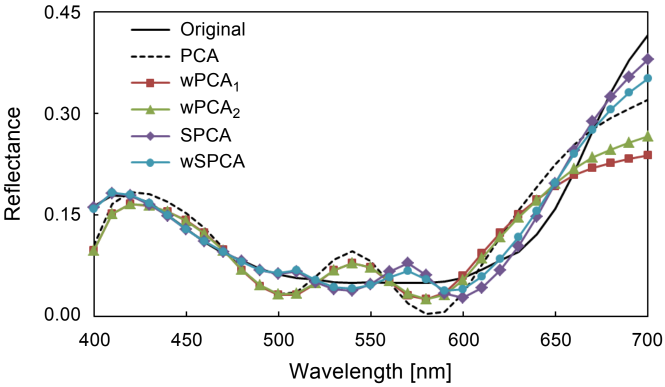

Fig. 5. Results of spectral reconstruction by using PCA, SPCA and wSPCA methods.

just the reverse. The same phenomenon can be seen more clearly from the average spectral residuals shown in Fig. 4**b**.

4. Conclusions

In this paper, we presented the segmented principal component analysis (SPCA) method and weighted version (wSPCA) method for reconstruction of spectral color information. The bands partition and the weighted function are connected with the CIE color -matching function, which is done to retain more color visual information in the reconstruction process. The feasibility of the SPCA and wSPCA were tested by reconstructing the mixed spectrum sets (including *Munsell Atlas*, *ColorChecker 24*, *Acrylic Paints* and *NCS Atlas*). The results indicated that the SPCA and wSPCA achieved higher spectral and colorimetric accuracy for all the testing samples than the classical PCA and wPCAs. In addition, the wSPCA retained clearly more color visual information.

Acknowledgements – This work was supported by the National Natural Science Foundation of China (No. 61301231) and the Innovation Fund Project for Graduate Student of Shanghai (No. JWCXSL1401).

References

[1] Valero E.M., Yu Hu, Hernández-Andrés J., Eckhard T., Nieves J.L., Romero J., Schnitzlein M., Nowack D., *Comparative performance analysis of spectral estimation algorithms and computational optimization of a multispectral imaging system for print inspection*, Color Research and Application **39**(1), 2014, pp. 16–27.

[2] Di-Yuan Tzeng, Berns R.S., *A review of principal component analysis and its applications to color technology*, Color Research and Application **30**(2), 2005, pp. 84–98.

[3] Haneishi H., Hasegawa T., Hosoi A., Yokoyama Y., Tsumura N., Miyake Y., *System design for accurately estimating the spectral reflectance of art paintings*, Applied Optics **39**(35), 2000, pp. 6621–6632.

[4] Xiuping Jia, Richards J.A., *Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification*, IEEE Transactions on Geoscience and Remote Sensing **37**(1), 1999, pp. 538–542.

[5] Barakzehi M., Amirshahi S.H., Peyvandi S., Afjeh M.G., *Reconstruction of total radiance spectra of fluorescent samples by means of nonlinear principal component analysis*, Journal of the Optical Society of America A **30**(9), 2013, pp. 1862–1870.
[6] Cohen J., *Dependency of the spectral reflectance curves of the Munsell color chips*, Psychonomic Science **1**(1–12), 1964, pp. 369–370.
[7] Vrhel M.J., Gershon R., Iwan L.S., *Measurement and analysis of object reflectance spectra*, Color Research and Application **19**(1), 1994, pp. 4–9.
[8] García-Beltrán A., Nieves J.L., Hernández-Andrés J., Romero J., *Linear bases for spectral reflectance functions of acrylic paints*, Color Research and Application **23**(1), 1998, pp. 39–45.
[9] Kohonen O., Parkkinen J., Jääskeläinen T., *Databases for spectral color science*, Color Research and Application **31**(5), 2006, pp. 381–390.
[10] Shams-Nateri A., *Wavelength intervals effect on reflectance spectra reconstruction*, Optica Applicata **42**(4), 2012, pp. 737–742.
[11] Laamanen H., Jetsu T., Jaaskelainen T., Parkkinen J., *Weighted compression of spectral color information*, Journal of the Optical Society of America A **25**(6), 2008, pp. 1383–1388.
[12] Guangyuan Wu, Zhen Liu, Enyin Fang, Haiqi Yu, *Reconstruction of spectral color information using weighted principal component analysis*, Optik – International Journal for Light and Electron Optics **126**(11–12), 2015, pp. 1249–1253.
[13] Agahian F., Funt B., Amirshahi S.H., *Spectral compression: weighted principal component analysis versus weighted least squares*, Proceedings of SPIE **9014**, 2014, article 90140Z.
[14] Jiandong Tian, Yandong Tang, *Wavelength-sensitive-function controlled reflectance reconstruction*, Optics Letters **38**(15), 2013, pp. 2818–2820.
[15] Flinkman M., Laamanen H., Tuomela J., Vahimaa P., Hauta-Kasari M., *Eigenvectors of optimal color spectra*, Journal of the Optical Society of America A **30**(9), 2013, pp. 1806–1813.
[16] Qian Du, Wei Zhu, He Yang, Fowler J.E., *Segmented principal component analysis for parallel compression of hyperspectral imagery*, IEEE Geoscience and Remote Sensing Letters **6**(4), 2009, pp. 713–717.
[17] *Spectral Database*, University of Eastern Finland, http://www2.uef.fi/fi/spectral/spectral-database
[18] Ayala F., Echávarri J.F., Renet P., Negueruela A.I., *Use of three tristimulus values from surface reflectance spectra to calculate the principal components for reconstructing these spectra by using only three eigenvectors*, Journal of the Optical Society of America A **23**(8), 2006, pp. 2020–2026.

Received August 6, 2015
in revised form November 19, 2015