Evaluation of Linear Models for Spectral Reflectance Dimensionality Reduction

Zhaojian Li, Roy S. Berns, Munsell Color Science Laboratory, Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, New York, USA

Abstract

It has been shown that dimensionality reduction techniques can be applied to a large dataset of spectral reflectance to reconstruct the spectra by the linear model with a small number of basis functions. Principal Component Analysis (PCA) and Independent Component Analysis (ICA) are two popular techniques to perform the dimensionality reduction. For each technique, there are two approaches to perform the spectral reconstruction. One approach is to use the mean-centered data while another approach excludes the mean offset. This paper presents these four linear models mathematically. The colorimetric and spectral accuracy of the spectral reconstructions using the four models were compared. It was found that ICA had slightly better performance than PCA using not only the mean-centered data but also the data excluding the mean offset. More importantly, ICA without the mean had very close performance to PCA with the mean.

Introduction

It has been shown that spectral reflectance can be represented using a linear model with a limited number of dimensions [1-7]. That is, hundreds or thousands of reflectance spectra are described by a considerably smaller set of basis functions, termed *statistical colorants* by Tzeng and Berns [7]. Different linear models have been used for the spectral reflectance dimensionality reduction with varying results. Cohen was the first to analyze the characteristic spectra of the Munsell colors using Principal Component Analysis (PCA) [1]. Parkkinen, *et al.* [2] and Fairman, *et al.* [3] used similar methods for a larger dataset with 1257 Munsell color chips. Laamanen, *et al.* [4] compared Independent Component Analysis (ICA) and PCA in color recognition using 1269 reflectance spectra of the Munsell Color chips. Their results showed that ICA had better reconstruction performance than PCA with the same number of dimensions. Ramanath, *et al.* [5] used the same 1269 Munsell samples to perform the spectral dimensionality reduction using ICA, PCA, and Neural networks. They found that PCA performed better than ICA in reproducing the spectra of the samples. Given the inconsistent results, these analyses were repeated in this research. In addition, considering the fact that the traditional PCA and ICA use the mean-centered data while in some applications, the mean is discarded [7] the linear models of PCA and ICA without a mean offset are evaluated.

Dimensionality Reduction Techniques

Principal Component Analysis (PCA)

Principal component analysis (PCA) is concerned with explaining the variance-covariance structure through a few linear combinations of the original variables [6]. Its goal is to find an uncorrelated representation of a set of correlated n-dimensional vectors. Given the spectral reflectance of one sample set, $\mathbf{X}$, a $n \times q$ matrix with n spectral bands (wavelength) and q number of samples. First, the variance-covariance matrix $\mathbf{C}$ is computed as:

$$\mathbf{C} = \frac{1}{q} \sum_{i=1}^{q} (\mathbf{X_i} - \overline{\mathbf{X}})(\mathbf{X_i} - \overline{\mathbf{X}})^T \tag{1}$$

where $\overline{\mathbf{X}}$ is the mean spectrum of the sample set and T is the matrix-vector transpose operation. After n eigenvectors $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ of $\mathbf{C}$ are obtained, the first m $(m < n)$ principal components coordinates using and discarding the mean spectrum are given as the following

$$\mathbf{Y} = \mathbf{V}^T (\mathbf{X} - \overline{\mathbf{X}}) \tag{2}$$

$$\mathbf{Y}^* = \mathbf{V}^T \mathbf{X} \tag{3}$$

where the $(n \times m)$ matrix $\mathbf{V}$ an orthogonal matrix formed by the first m $(m < n)$ eigenvectors. The spectral reflectance can be reconstructed using and discarding the mean spectrum respectively by

$$\hat{\mathbf{X}} = \overline{\mathbf{X}} + \mathbf{V}\mathbf{Y} = \overline{\mathbf{X}} + \mathbf{V}\mathbf{V}^T (\mathbf{X} - \overline{\mathbf{X}}) \tag{4}$$

$$\hat{\mathbf{X}}^* = \mathbf{V}\mathbf{Y}^* = \mathbf{V}\mathbf{V}^T \mathbf{X} \tag{5}$$

Independent Component Analysis (ICA)

Independent component analysis (ICA) is another technique for dimensionality reduction. Its goal is to produce basis functions that give rise to maximum statistical independence of the data [8]. When applying dimensionality reduction of spectral data, the model of ICA is given by

$$\mathbf{X} = \mathbf{A}\mathbf{S} \tag{6}$$

where $\mathbf{X}$ is a $q \times n$ matrix with n spectral bands (wavelength) and q number of samples representing observed mixtures (linear combination of the original source signals and mixing matrix). $\mathbf{S}$ is a $m \times n$ matrix with m independent components (ICs) representing the original source signals. For the intent of the dimensionality reduction, m is less than q. $\mathbf{A}$ is a $q \times m$ scalar matrix of mixing coefficients to construct $\mathbf{X}$ from the various independent components.

ICA with the mean

For the ICA approach using the mean-centered data, before attempting to estimate $\mathbf{A}$ and $\mathbf{S}$, both "centering" and "whitening" preprocessing are performed. First, the obtained data $\mathbf{X}$ are centered by subtracting their mean value $\overline{\mathbf{X}}$ to make $\mathbf{X}$ zero-mean.

$$\overline{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n}(x_1, x_2, x_3, ... x_q)_i \qquad (7)$$

$$\vec{\mathbf{X}} = \mathbf{X} - \overline{\mathbf{X}} \qquad (8)$$

where $\overline{\mathbf{X}}$ is a $1 \times q$ vector. It should be noted that in this case the mean vector is not the mean spectrum as in PCA since q represents the number of the samples in the dataset. That is, it has not a physical interpretation in terms of spectral reflectance. Then the centered data are whitened, which means they are linearly transformed so that the components are uncorrelated and have unit variance. "Whitening" can be performed via eigenvalue decomposition of the covariance matrix $E\{\vec{\mathbf{X}}\vec{\mathbf{X}}^T\} = \mathbf{U}\mathbf{D}\mathbf{U}^T$. $\mathbf{U}$ is here the orthogonal matrix of eigenvectors of $E\{\vec{\mathbf{X}}\vec{\mathbf{X}}^T\}$ and $\mathbf{D}$ is the diagonal matrix of its eigenvalues. Whitening can now be done by

$$\vec{\vec{\mathbf{X}}} = \mathbf{D}^{-1/2}\mathbf{U}^T\vec{\mathbf{X}} \qquad (9)$$

In addition to simplifying the ICA algorithm, whitening can also perform the dimensionality reduction at the same time. When calculating the whitening matrix $\mathbf{D}^{-1/2}\mathbf{U}^T$, first m largest eigenvalues and corresponding eigenvector of $E\{\vec{\mathbf{X}}\vec{\mathbf{X}}^T\}$ are selected and the rest are discarded, as in often done in PCA. That is, $\mathbf{D} = diag(d_1, d_2, ... d_m)$ and $\mathbf{U}$ now is the orthogonal matrix with m eigenvectors. So far, we can see that PCA is the preprocessing of ICA. In this case, $\vec{\vec{\mathbf{X}}}$ is the centered and whitened data, which is a $m \times n$ matrix. The number of dimensions is reduced from q to m rather than from n to m in PCA.

After the preprocessing of the data, the ICA algorithm is performed. There are many different algorithms such as CoBliss, FastICA, and JADE. The JADE algorithm (Joint Approximate Diagonalization of Eigenvalues) [9, 10] was used in Laamanen's approach as well as the new approach in this research. The direct output of the JADE algorithm is the separation matrix $\mathbf{B}$, which is the inverse of mixing matrix $\mathbf{A}$. The independent components $m \times n$ matrix $\mathbf{S}$ is estimated by

$$\mathbf{S} = \mathbf{B}\vec{\vec{\mathbf{X}}} \qquad (10)$$

In this approach, the independent components are used as the m basis functions to form the new coordinate system. The original data could be reconstructed by

$$\hat{\mathbf{X}} = \mathbf{S}^T\mathbf{S}\vec{\mathbf{X}} + \overline{\mathbf{X}} \qquad (11)$$

ICA without the mean

Considering the fact that the mean vector has no physical interpretation in terms of spectral reflectance, the original data without a mean vector offset can be used to perform ICA in this new approach. The "whitening" preprocessing is also required to perform the dimensionality reduction, as shown in the following:

$$\vec{\vec{\mathbf{X}}} = \mathbf{D}^{-1/2}\mathbf{U}^T\mathbf{X} \qquad (12)$$

where $\mathbf{X}$ is the original data, $\mathbf{D} = diag(d_1, d_2, ... d_m)$ and $\mathbf{U}$ is the orthogonal matrix with m eigenvectors coming from $E\{\mathbf{X}\mathbf{X}^T\}$.

Applying the same JADE algorithm as the approach of ICA with the mean, the independent components $m \times n$ matrix $\mathbf{S}$ is estimated by

$$\mathbf{S} = \mathbf{B}\mathbf{X} \qquad (13)$$

The spectral reconstruction can be performed by

$$\hat{\mathbf{X}} = \mathbf{S}^T\mathbf{S}\mathbf{X} \qquad (14)$$

Results and Discussions

In this study, 1269 reflectance spectra of the chips in the Munsell Book of Color-Matte Finish Collection defined the data set [11]. The wavelength range was from 400 nm to 700 nm with 10 nm intervals. That is, the number of spectral bands n equaled 31 and the number of samples q was 1269. The three basis functions of the Munsell data set from PCA and the two ICA approaches are shown in Figures 1, 2, and 3, respectively.
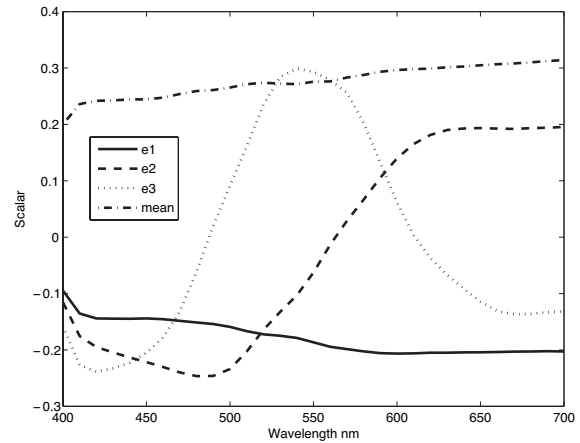

Figure 1. The three basis functions of the Munsell spectral dataset from PCA and the mean spetrum of the Munsell spectral dataset.
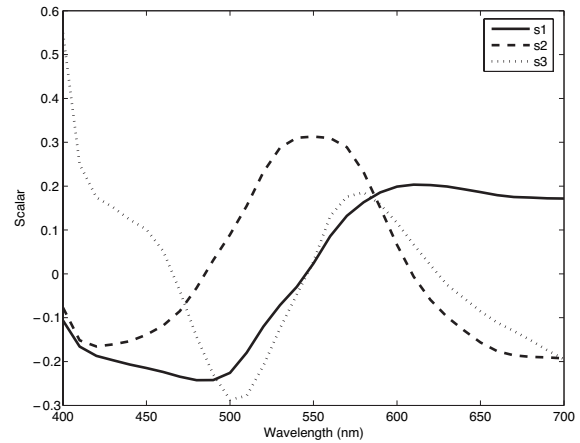

Figure 2. The three basis functions of the Munsell spectral dataset from ICA with the mean.
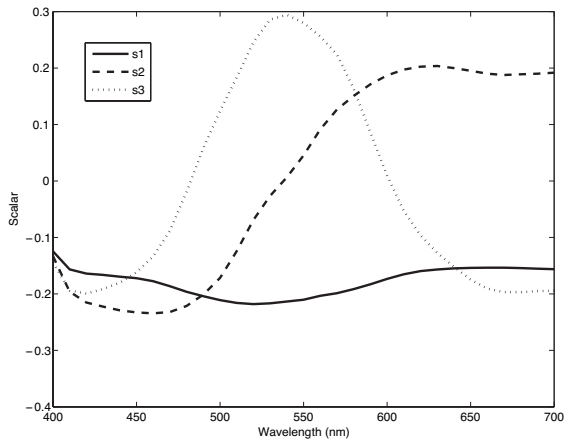
Figure 3. *The three basis functions of the Munsell spectral dataset from ICA without the mean.*

The three basis functions of the Munsell spectral dataset using PCA and ICA with the mean are consistent with that in the literature but the opposite signs in some basis functions and different normalization. For PCA, since the covariance matrix of the mean-centered data and un-mean-centered data is identical, two approaches have the same eigenvectors. For ICA, it's goal is to seeks directions in feature such that the resulting signal show independence, so the mean-centered data would result in different independence components from that of the data without a mean offset. Comparing Figure 2 and 3, two sets of basis functions are different. Interestingly, the basis functions from ICA without the mean are very similar to that of PCA.

The average colorimetric and spectral accuracy of spectral reconstruction using the various dimensionality reduction techniques is listed in Table 1. The colorimetric accuracy was calculated using CIEDE2000 for 1931 standard observer under illuminant D65. The spectral accuracy is expressed by the RMS of spectral reflectance [12].

Table 1: The average colorimetric and spectral accuracy of the spectral reconstruction for the Munsell dataset with different number of basis functions from PCA and ICA.

| Function number | PCA with mean | | PCA without mean | | ICA with mean | |
	DE2000 (D65, 2^0)	RMS Spectral Error	DE2000 (D65, 2^0)	RMS Spectral Error	DE2000 (D65, 2^0)	RMS Spectral Error
1	16.20	0.077	17.43	0.080	12.02	0.044
2	12.01	0.042	12.47	0.043	1.97	0.023
3	2.18	0.019	1.91	0.021	1.70	0.016
4	1.22	0.013	1.37	0.015	0.64	0.012
5	0.56	0.009	0.76	0.012	0.24	0.008
6	0.56	0.008	0.76	0.011	0.23	0.006
7	0.12	0.006	0.16	0.007	0.13	0.005
8	0.10	0.004	0.15	0.007	0.10	0.003
9	0.09	0.003	0.12	0.004	0.08	0.003
10	0.06	0.002	0.08	0.003	0.08	0.002
11	0.05	0.002	0.07	0.002	0.02	0.002
12	0.05	0.002	0.08	0.002	0.01	0.002
13	0.01	0.001	0.01	0.002	0.01	0.001
14	0.01	0.001	0.01	0.001	0.01	0.001
15	0.01	0.001	0.01	0.001	0.01	0.001
16	0.01	0.001	0.01	0.001	0.01	0.001
17	0.00	0.001	0.00	0.001	0.00	0.001
18	0.00	0.001	0.00	0.001	0.00	0.001
19	0.00	0.001	0.00	0.001	0.00	0.001
20	0.00	0.001	0.00	0.001	0.00	0.001
21	0.00	0.001	0.00	0.001	0.00	0.001
22	0.00	0.001	0.00	0.001	0.00	0.001
23	0.00	0.001	0.00	0.001	0.00	0.001
24	0.00	0.001	0.00	0.001	0.00	0.000
25	0.00	0.000	0.00	0.001	0.00	0.000

The spectral reconstruction for an arbitrary sample in the Munsell data set using the derived three basis functions from PCA and ICA, are plotted in Figure 4.
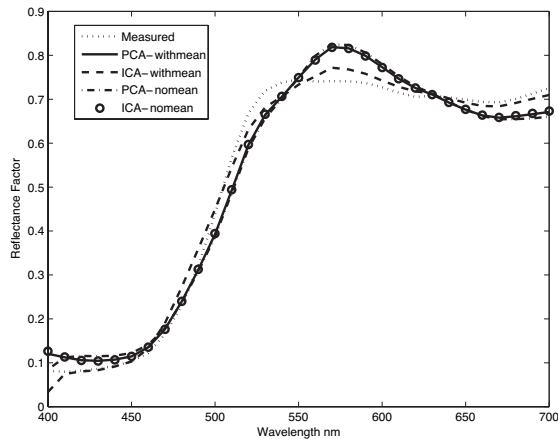

Figure 4. *An example of the spectral reconstruction using three basis functions from PCA and ICA.*

It is obvious that both PCA and ICA built effective linear models for spectral reflectance dimensionality reduction. That is, using a small number of basis functions, the spectral reflectance of the dataset can be reconstructed with tolerable accuracy.

For both PCA and ICA, the spectral reconstruction performance with the mean was better than that without the mean using the same number of basis functions. However, if the mean spectrum was treated as a basis function, the results of PCA without the mean would be better than that of PCA with the mean using the same number of basis functions. However, when the number of functions is larger than two, the difference becomes very small.

For the mean-centered data, ICA has the better spectral reconstruction performance than PCA. This conclusion is consistent with the result of Laamanen, *et al*, however, opposite to that of Ramanath, *et al*. For the data without a mean offset, ICA also has smaller spectral reconstruction accuracy than PCA.

A significant result is that ICA without the mean has very close performance to PCA with the mean. The latter needs to know the mean spectrum of the dataset before reconstructing it. This indicates that in order to obtain the same spectral reconstruction, there is an extra parameter to know for PCA with the mean comparing with ICA without the mean.

Conclusions

Four linear models for spectral reflectance dimensionality reduction using PCA and ICA were evaluated in terms of spectral reconstruction accuracy. Three basis functions are enough to reconstruct the measured data with tolerable accuracy. The comparison between PCA and ICA shows ICA has slightly better performance than PCA using not only the mean-centered data but also the un-mean-centered data. Although ICA with the mean performs better than other models, the mean vector has no physical interpretation in terms of spectral reflectance. In addition, ICA without the mean has very close performance to PCA with the

mean, which require an extra parameter. So for the purpose of dimensionality reduction and data compression, ICA without the mean is a better approach than the other three models.

Acknowledgments

References

[1] J. B. Cohen, Dependency of the spectral reflectance curves of the Munsell Color chips, Psychon. Sci. 1, 369 (1964).
[2] J. Parkkinen, J. Hallikainen, T. Jaaskelainen, Characteristic spectra of Munsell colors, J. Opt. Soc. Am. A, 6, 318 (1989).
[3] H. S. Fairman, M. H. Brill, The principal components of reflectances, Col. Res. Appl., 29, 104 (2004).
[4] H. Laamanen, T. Jaaskelainen, J. Parkkinen, Comparison of PCA and ICA in color recognition, Proc. SPIE, pg. 367. (2001).
[5] R. Ramanath, R. G. Kuehni, W. E. Snyder, D. Hinks, Spectral spaces and color spaces, Col. Res. Appl., 29, 29 (2004).
[6] R. A. Johnson, D. W. Wichern. Applied multivariate statistical analysis (Prentice Hall, New Jersey, 2002) pg. 100.
[7] D. Tzeng, R. S. Berns, A review of principal component analysis and its application to color technology, Color Res. Appl., 30, 84 (2005).
[8] A. Hyvarinen, E. Oja, Independent component analysis: Algorithms and applications, Neural Networks, 13, 411 (2000).
[9] J. F. Cardoso, A. Souloumiac, Blind beamforming for non Gaussian signals, Proc. IEEE, pg. 1878. (1998).
[10] The JADE algorithm in Matlab is available at: http://www.tsi.enst.fr/icacentral/Algos/cardoso/
[11] The Munsell spectra database is available at: http://cs.joensuu.fi/~spectral/databases/download/munsell_spec_matt.htm
[12] R. S. Berns, Billmeyer and Saltzman's principles of color technology, (Wiley, New York, NY, 2000) pg. 126.

Author Biography

Zhaojian Li received his B. S. degree in Physics from Central University for Nationalities, China in 2001 and his M. S. degree in Optical Engineering from Beijing Institute of Technology, China in 2004. Now he is the second year M. S. student in color science in the Munsell Color Science Laboratory at Rochester Institute of Technology.