PROGRESS IN COLOR, COLORANTS AND COATINGS

available online @ www.pccc.icrc.ac.ir

Prog. Color Colorants Coat. 6(2013), 129-139

Ministry of Science, Research and Technology
Institute for Color Science and Technology

Evaluating Dye Concentration in Bi-Component Solution by PCA-MPR and PCA-ANN Techniques

A. Shams-Nateri *

Associate Professor, (a) Center of Excellence for Color Science and Technology and (b) Textile Engineering Department, University of Guilan, P.O. Box: 3756, Rasht, Iran.

ARTICLE INFO

Article history:
Received: 29-10-2012
Final Revised: 06-07-2013
Accepted: 16-07-2013
Available online: 17-07-2013

Keywords:
Dye concentration prediction
Principal component analysis
Neural network
Polynomial regression
Spectrophotometry

ABSTRACT

This paper studies the application of principal component analysis, multiple polynomial regression, and artificial neural network techniques to the quantitative analysis of binary mixture of dye solution. The binary mixtures of three textile dyes including blue, red and yellow hues were analyzed by PCA-Multiple polynomial regression and PCA-artificial neural network methods. The obtained results indicate that the accuracy of PCA-artificial neural network technique is higher than PCA-Multiple polynomial regression and normal spectroscopy methods. The PCA-artificial neural network technique is applicable for dye concentration bicomponent solution with both overlapping and non-overlapping spectra. The developed method can be a practical solution to quantitative analysis of binary mixture of dye solutions with overlapping. Prog. Color Colorants Coat. 6(2013), 129-139. © Institute for Color Science and Technology.

1. Introduction

The Beer-Lambert theory explained the relationship between the amount of light absorbance and the concentration of the substance absorbing the light and the light path. The Beer- Lambert's law is written as equation 1.

$$A = \varepsilon \times L \times C \qquad (1)$$

where A: absorbance, ε: absorptivity (liter mole^{-1} cm^{-1}), L:path length (cm), and C: concentration (mole liter^{-1}).

According to Beer-Lambert theory, the absorption spectrum is used to quantitative analysis of the compound. The absorbance value at wavelength of maximum absorption (λ_{max}) is used in most application. UV-Visible spectrophotometery can be used to determine several characteristics of the compound. The wavelength of maximum absorbance is not fixed and depends partly on the solvent in which the substance is dissolved, temperature and pH, which may change the intensity and the wavelength of maximum absorbance. The performance of UV-Visible spectrophotometery for multiple components is less than single-component. However, modern curve-fitting techniques improve the performance of UV-Visible spectrophotometery in multiple component analysis. According to beer-lambert's law, the absorbance at any wavelength is equal to the sum of the absorbance of each component in the mixture. In spectroscopy method, the effect of random

*Corresponding author: a_shams@guilan.ac.ir

error can be reduced by using additional absorbance spectral information and least squares curve fitting technique [1-8].

Principal component analysis is a classical statistical technique for dataset analyzing, simplifying, and compression by reducing multidimensional datasets to lower dimensions. Principal component analysis is an orthogonal linear transformation that transforms the data to a new coordinate system such that the maximum variance of the data comes to lie on the first coordinate, the second greatest variance on the second coordinate, and so on. In datasets dimensional reduction and reducing the number of variables, the most important component like first and second principal components are used and higher-order principal components was ignored. Several principles and methods such as principal component analysis (PCA), factor analysis (FA), and independent component analysis (ICA) have been used to linearize transformation of dataset. Principal component analysis is a simple, non-parametric method of extracting applicable information from confusing datasets. It is used abundantly in all forms of analysis from neuroscience to computer graphics. Principal component analysis is a classical second-order method of linear transform, which uses the information contained in the covariance matrix of the data vector. Principal component analysis is widely used in several application areas such as signal processing, statistics, neural computing and color science [9-18].

In color science, principal component analysis has been used for data reduction, spectral reconstruction and color matching [11-18]. Agahian and Amirshahi [16] suggested a new matching approach based on the matching of the first three principal components. They obtained some improvements in comparison with normal colorimetric algorithms in terms of spectral accuracy as well as colorimetric accuracy. Shams-Nateri [17] suggested principal component analysis-based spectrophotometric match prediction method for color matching colored fabrics. In addition, principal component analysis and derivative spectrophotometry techniques were used to improve the accurateness of Beer's law to predict the dyes concentration [18].

The goal of present research is to suggest new accurate spectrophotometric based method for the dye concentration determination in binary mixtures by using principal component analysis and artificial neural network techniques.

2. Experimental

2.1. Materials and Methods

Cibacron blue GNE, Cibacron Red GE, Cibacron yellow 2GE from CIBA Chemical Corporation were used to construct bi-component mixture. Blue-Yellow, Red-Blue and Red-Yellow binary mixtures were prepared by using Blue, Red and Yellow dyes with concentrations of 0.01, 0.025, 0.04, 0.055, 0.07 and 0.085 g/l. The absorbance spectrum from 400 to 700 nm was measured by CINTRA10-UV-Visible spectrophotometer. The dataset was randomly divided into two groups as training and testing dataset. The 24 samples were selected for training, and 12 samples were used to assess the performance of the methods.

2.2. Determination of concentration:

2.2.1. Normal spectrophotometry method

In analyzing the binary mixture of dye solution, two wavelengths are selected as the wavelengths of maximum absorbance (λ_{max}) of the components. Then, the concentration of each dye was calculated from the following simultaneous equations:

$$\begin{cases} at \quad \lambda_{\max 1} \quad A_{\lambda_1} = \varepsilon_{1,\lambda_1} \times C_1 + \varepsilon_{2,\lambda_1} \times C_2 \\ at \quad \lambda_{\max 2} \quad A_{\lambda_2} = \varepsilon_{1,\lambda_2} \times C_1 + \varepsilon_{2,\lambda_2} \times C_2 \end{cases} \tag{2}$$

where $\lambda_{\max 1}$ and $\lambda_{\max 2}$ are the wavelengths of maximum absorbance of components, $\varepsilon_{j,\lambda_i}$ is absorptivity of j[th] component at wavelength λ_i. C_1 and C_2 are the concentration of first and second components of the binary mixture.

The accuracy of prediction is evaluated by using binary relative error (E_{Bi}), which is calculated by equation 3:

$$E_{Bi} = \sqrt{\frac{(C_{a,1} - C_{p,1})^2 + (C_{a,2} - C_{p,2})^2}{(C_{a,1})^2 + (C_{a,2})^2}} \tag{3}$$

where E_{Bi} is binary relative error, $C_{a,i}$ is actual

concentration value of i[th] component and $C_{p,i}$ is predicted concentration value of i[th] component.

2.2.2. PCA-Multiple polynomial regression method

In this method, the relationship between the dye concentration and principal components of absorbance spectra were obtained by means of multiple polynomial regression technique with two and three principal components. Initially, the transformation matrix was calculated for transferring principal components into dye concentration. Several regression schemes were employed as detailed in Table 1 and Table 2 for two and three principal components, respectively. In next stage, this transformation matrix was used to predict dye concentration of binary mixture as equations 4 and 5 respectively for two and three principal components.

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & . & . & . & a_{1,m} \\ a_{2,1} & a_{2,2} & . & . & . & a_{2,m} \end{bmatrix}_{2\times m} \times \begin{bmatrix} PC_1 \\ PC_2 \\ (PC_1)^2 \\ (PC_2)^2 \\ . \\ . \\ . \\ . \\ . \\ . \end{bmatrix}_{m\times 1} \quad (4)$$

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & . & . & . & a_{1,m} \\ a_{2,1} & a_{2,2} & . & . & . & a_{2,m} \end{bmatrix}_{2\times m} \times \begin{bmatrix} PC_1 \\ PC_2 \\ PC_3 \\ (PC_1)^2 \\ (PC_2)^2 \\ (PC_3)^2 \\ . \\ . \\ . \\ . \end{bmatrix}_{m\times 1} \quad (5)$$

Where C_1 and C_2 are dye concentration of mixtures components, PC_1, PC_2 and PC_3 are first three principal components with highest eigenvalue.

Table 1: PCA- Multiple Polynomial regression models with two principal components.

No.	Number of terms	Polynomial matrix (m,R,G,B)
1	2	PC1 PC2
2	3	1 PC1 PC2
3	5	1 PC1 PC2 $PC1^2$ $PC2^2$
4	6	1 PC1 PC2 $PC1^2$ $PC2^2$ PC1*PC2
5	8	1 PC1 PC2 $PC1^2$ $PC2^2$ PC1*PC2 $PC1^3$ $PC2^3$
6	10	1 PC1 PC2 $PC1^2$ $PC2^2$ PC1*PC2 $PC1^3$ $PC2^3$ $PC1^4$ $PC2^4$

Table 2: PCA- Multiple Polynomial regression models with three principal components.

No.	Number of terms	Polynomial matrix (m,R,G,B)
1	3	PC1 PC2 PC3
2	4	1 PC1 PC2 PC3
3	7	1 PC1 PC2 PC3 $PC1^2$ $PC2^2$ $PC3^2$
4	10	1 PC1 PC2 PC3 $PC1^2$ $PC2^2$ $PC3^2$ PC1*PC2 PC1*PC3 PC2*PC3
5	11	1 PC1 PC2 PC3 $PC1^2$ $PC2^2$ $PC3^2$ PC1*PC2 PC1*PC3 PC2*PC3 $PC1^3$ $PC2^3$ $PC3^3$ PC1*PC2*PC3
6	12	1 PC1 PC2 PC3 $PC1^2$ $PC2^2$ $PC3^2$ PC1*PC2 PC1*PC3 PC2*PC3 $PC1^3$ $PC2^3$ $PC3^3$ PC1*PC2*PC3 $PC1^4$ $PC2^4$ $PC3^4$

2.2.3. PCA-Artificial neural network (PCA-ANN) method

In this method, neural network technique was used to obtain a relationship between the dye concentration of mixtures and principal components of absorbance spectra.The used multilayer perceptrons neural networks contained two outputs as the concentration of the binary mixture, and two and three inputs for two and three principal components, respectively. Figure 1 shows the structure of neural network with three principal components and one hidden layer with 4 nodes. The topologies of neural networks are shown in Table 3 and Table 4 respectively for two and three principal components.

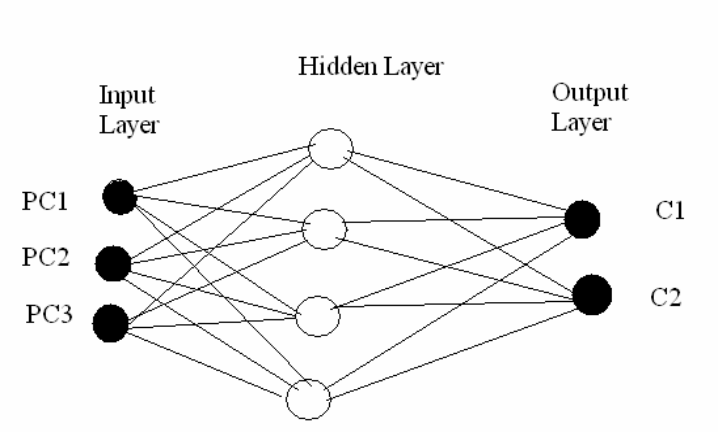

Figure 1: Structure of neural network with three principal components.

Table 3: PCA- Artificial neural network models with two principal components.

No.	Number of layers	Number of nodes per layer
1	3	2 3 2
2	3	2 4 2
3	3	2 5 2
4	3	2 6 2
5	4	2 3 2 2
6	4	2 4 2 2
7	4	2 5 2 2
8	4	2 5 3 2

Table 4: PCA- Artificial neural network models with three principal components.

No.	Numberof layers	Number of nodes per layer
1	3	3 3 2
2	3	3 4 2
3	3	3 5 2
4	3	3 6 2
5	4	3 3 2 2
6	4	3 4 2 2
7	4	3 5 2 2
8	4	3 5 3 2

The neural network was trained with backpropagation algorithm. The hyperbolic tangent sigmoid transfer function (TANSIG) was used to calculate a layer's output from its input. The training was occurred according to TRAINLM training function. During the training, the weights of network are modified to reduce the error between the predicted and actual output. The training error goal was 0.0001 and neural network training was continued over 500 epochs by back propagation algorithm. The generalization performance of trained neural network is evaluated by using testing dataset. The trained neural networks were used to evaluate the dye concentration of bi-component mixture from their principal components of absorbance spectra. The performance of neural networks was evaluated by binary relative error (E_{Bi}), which is evaluated by equation 3.

3. Results and discussion

The obtained results in evaluating dye concentrations of Blue-Yellow, Red-Blue and Red-Yellow binary mixtures by normal spectroscopy, PCA-Multiple polynomial regression and PCA-Artificial neural network are shown in Tables 5 to 9.

Table 5: Binary relative error in spectrophotometry method.

Combination	$\Delta\lambda=\lvert\lambda_{max,1}-\lambda_{max,2}\rvert$	Mean	STD	Max	Min
Yellow- Blue	170	0.764	1.245	6.995	0.012
Red -Blue	80	1.067	1.908	11.521	0.094
Yellow -Red	90	1.342	3.717	22.217	0.067

Table 6: Binary relative error in PCA- multiple polynomial regression method with two principal components.

Combination	Method	Mean	STD	Max	Min
Yellow-Blue	1	0.767	0.230	1.000	0.000
	2	0.133	0.128	0.673	0.022
	3	0.136	0.136	0.747	0.016
	4	0.128	0.128	0.667	0.016
	5	0.123	0.136	0.697	0.013
	6	0.147	0.138	0.664	0.003
Red-Blue	1	0.840	0.161	1.000	0.395
	2	0.311	0.335	2.071	0.058
	3	0.282	0.275	1.562	0.023
	4	0.262	0.265	1.492	0.037
	5	0.231	0.297	1.454	0.021
	6	0.212	0.280	1.343	0.021
Yellow-Red	1	0.904	0.125	1.000	0.530
	2	0.370	0.240	1.317	0.067
	3	0.288	0.203	1.083	0.032
	4	0.187	0.178	0.821	0.006
	5	0.174	0.148	0.567	0.005
	6	0.150	0.130	0.504	0.008

Table 7: Binary relative error in PCA- multiple polynomial regression method with three principal components.

Combination	Method	Mean	STD	Max.	Min.
Yellow-Blue	1	0.795	0.191	1.000	0.271
	2	0.074	0.046	0.206	0.012
	3	0.072	0.054	0.259	0.009
	4	0.067	0.047	0.199	0.009
	5	0.080	0.085	0.428	0.008
	6	0.098	0.149	0.836	0.002
Red-Blue	1	0.840	0.161	1.000	0.395
	2	0.258	0.183	0.770	0.035
	3	0.203	0.173	0.675	0.013
	4	0.194	0.155	0.632	0.029
	5	0.166	0.260	1.045	0.001
	6	0.157	0.275	1.019	0.001
Yellow-Red	1	0.904	0.125	1.000	0.530
	2	0.340	0.185	0.714	0.051
	3	0.248	0.171	0.799	0.031
	4	0.137	0.126	0.540	0.021
	5	0.117	0.135	0.534	0.002
	6	0.111	0.136	0.511	0.000

3.1. Normal spectrophotometry method

The obtained results in spectrophotometry method are summarized in Table 5 as mean, standard deviation, maximum and minimum of binary relative error. From this table, the performance of normal spectrophotometry method for Yellow- Blue mixture with 0.764 binary relative errors is better than Red–Blue and Yellow-Red mixtures respectively with 1.067 and 1.342 binary relative errors.

3.2. Principal component analysis

The principal component analysis (PCA) was performed over training spectral databases for understanding the statistical nature absorbance spectra. The eigenvalue of principal components and three first principal components of Red-Blue, Red-Yellow and Blue-Yellow binary mixture are shown in Figures 2, 3 and 4, respectively. As shown in these figures, the first three eigenvector have significant eigenvalue and others are ignored. While, the first two and three principal components are used to evaluate dye concentration of mixtures by using PCA-Multiple polynomial regression and PCA-neural network techniques.

3.3. PCA-Multiple Polynomial Regression (PCA-MPR) Method

The obtained results in PCA-multiple polynomial regression with two and three principal components are summarized in Table 6 and Table 7, respectively. In these tables, the binary relative error (E_{Bi}) is evaluated by equation 3. The tables indicate that the binary relative error decreases by increasing polynomial terms. In multiple polynomial regression with two principal components, the best accuracy is obtained by 8 polynomial terms with 0.123 binary relative error for Yellow-Blue mixture and 10 polynomial terms with 0.212 and 0.150 binary relative errors for Red-Blue and Yellow-Red mixtures, respectively.

Table 8: Binary relative error in PCA- Artificial neural network method with two principal components.

Combination	Method	Mean	STD	Max.	Min.
Yellow-Blue	1	0.223	0.341	1.000	0.000
	2	0.210	0.339	1.018	0.000
	3	0.144	0.208	0.769	0.002
	4	0.117	0.219	0.785	0.000
	5	0.305	0.811	4.579	0.000
	6	0.116	0.214	0.731	0.000
	7	0.175	0.321	1.000	0.000
	8	0.168	0.298	1.064	0.000
Red-Blue	1	0.137	0.181	0.757	0.001
	2	0.242	0.376	1.427	0.000
	3	0.235	0.399	1.700	0.000
	4	0.589	1.640	9.248	0.000
	5	0.738	2.577	14.873	0.000
	6	0.755	1.512	5.835	0.000
	7	0.640	1.558	6.704	0.000
	8	0.205	0.423	1.894	0.000
Yellow-Red	1	0.118	0.284	1.699	0.003
	2	0.267	0.887	5.011	0.000
	3	0.201	0.421	1.754	0.000
	4	0.131	0.267	0.946	0.000
	5	0.178	0.318	0.973	0.000
	6	0.222	0.371	1.093	0.000
	7	0.171	0.343	1.000	0.000
	8	0.176	0.317	1.040	0.000

In multiple polynomial regression with three principal components, the best accuracy is obtained by 6 polynomial terms with 0.067 binary relative error for Yellow-Blue mixture and 12 polynomial terms with 0.157 and 0.111 binary relative errors for Red-Blue and Yellow-Red mixtures, respectively. As a result, the performance of PCA-multiple polynomial regression with three principal components is better than PCA-multiple polynomial regressions with two principal components. As shown in Figures 1 to 3, the eigenvalue of third principal components is more than zero and significant.

3.4. PCA-Artificial Neural Network (PCA-ANN) method

Obtained results by principal components analysis-artificial neural network (PCA-ANN) are summarized in Table 8 and Table 9 as binary relative error (E_{Bi}). The optimal topology of neural network is determined by the number of layers, the number of nodes per layer. In neural network with two principal components, the best accuracy for Yellow-Blue mixture is obtained by four layers containing 2, 6 and 2 nodes per layers with 0.117 binary relative error. The best accuracy for Red-Blue, and Yellow-Red mixtures are obtained by three layers containing 2, 3 and 2 nodes per layers with 0.137 and 0.118 binary relative errors, respectively. In three principal components neural network, the best accuracy for Yellow-Blue mixture is obtained by four layers with 3, 5, 3, and 2 nodes per layers with 0.045 binary relative error.

Table 9: Binary relative error in PCA- Artificial neural network method with three principal components.

Combination	Method	Mean	STD	Max.	Min.
Yellow-Blue	1	0.152	0.361	1.904	0.000
	2	0.080	0.172	0.885	0.001
	3	0.174	0.324	1.276	0.000
	4	0.070	0.134	0.524	0.000
	5	0.159	0.294	0.928	0.000
	6	0.082	0.146	0.528	0.000
	7	0.051	0.109	0.552	0.000
	8	0.045	0.085	0.296	0.000
Red-Blue	1	0.083	0.179	0.750	0.000
	2	1.158	2.670	11.438	0.000
	3	0.196	0.349	1.064	0.000
	4	0.080	0.179	0.753	0.000
	5	0.163	0.336	1.228	0.000
	6	0.104	0.227	0.781	0.000
	7	0.113	0.224	0.753	0.000
	8	0.147	0.260	0.872	0.000
Yellow-Red	1	0.073	0.126	0.429	0.000
	2	0.147	0.350	1.366	0.000
	3	0.077	0.162	0.848	0.000
	4	0.086	0.176	0.646	0.000
	5	0.035	0.067	0.282	0.000
	6	0.094	0.155	0.485	0.000
	7	0.159	0.320	1.000	0.000
	8	0.126	0.231	0.888	0.000

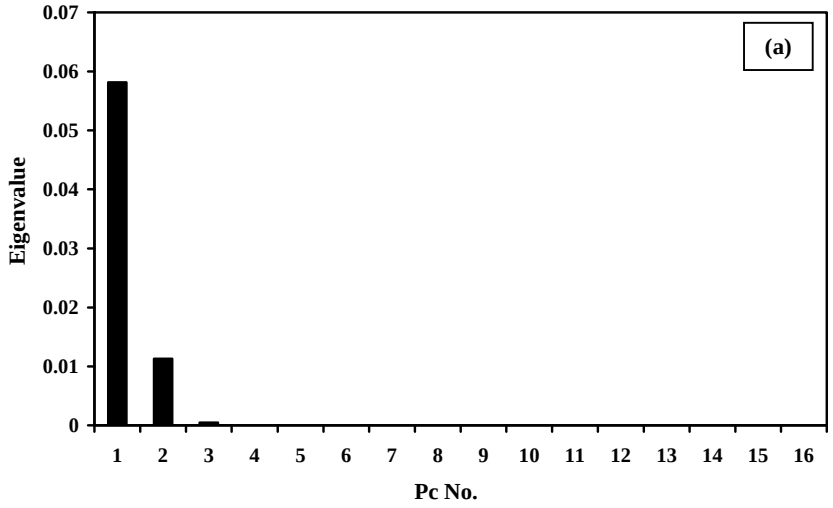

Figure 2: Eigenvalues (a) and first three principal components (b) of Red-Blue mixture.

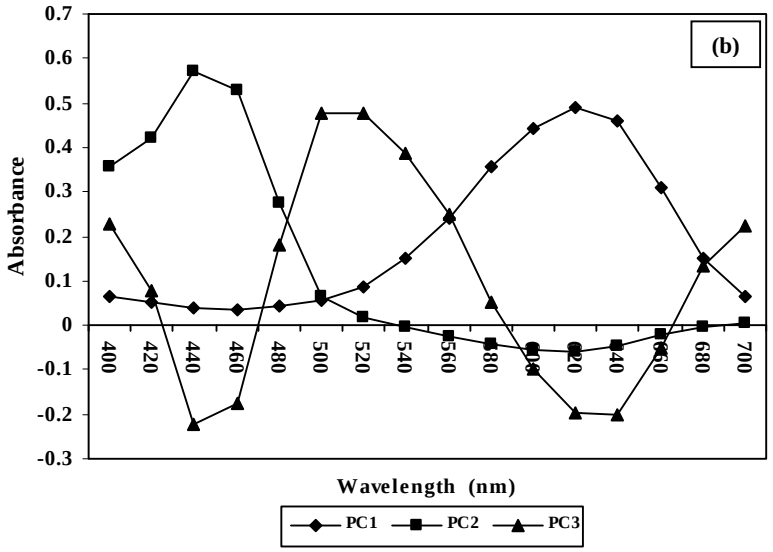

Figure 2: Continued.

The best accuracy for Red-Blue mixture is obtained by three layers with 3, 6 and 2 nodes per layers with 0.08 binary relative error, respectively. The best accuracy for Yellow-Red mixture is obtained by four layers with 3, 3, 2, and 2 nodes per layers with 0.035 binary relative error. These results suggest that the accuracy of neural network with three principal components is better than neural network with two principal components. However, the neural network method needs more samples than normal spectrophotometric based method.

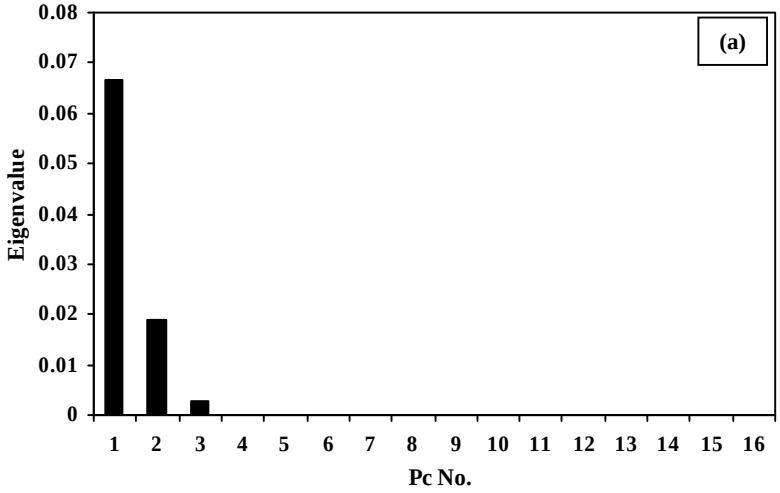

Figure 3: Eigenvalues (a) and first three principal components (b) of Red-Yellow mixture.

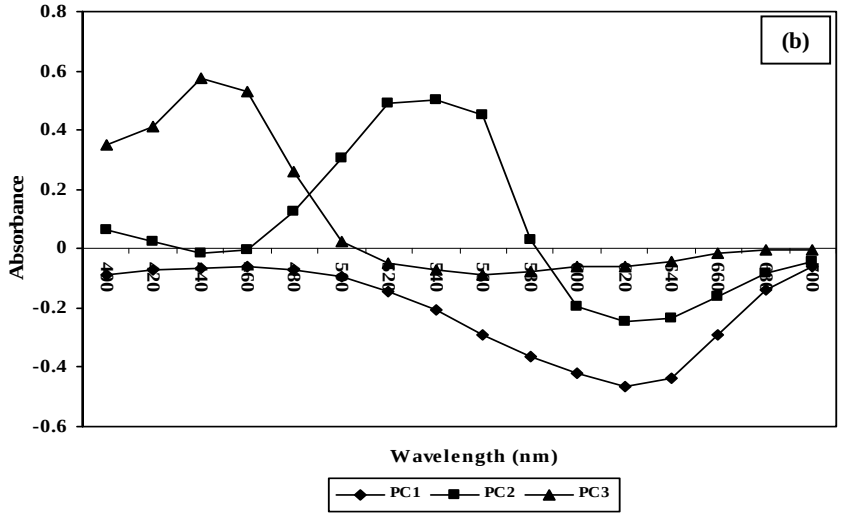

Figure 3: Continued.

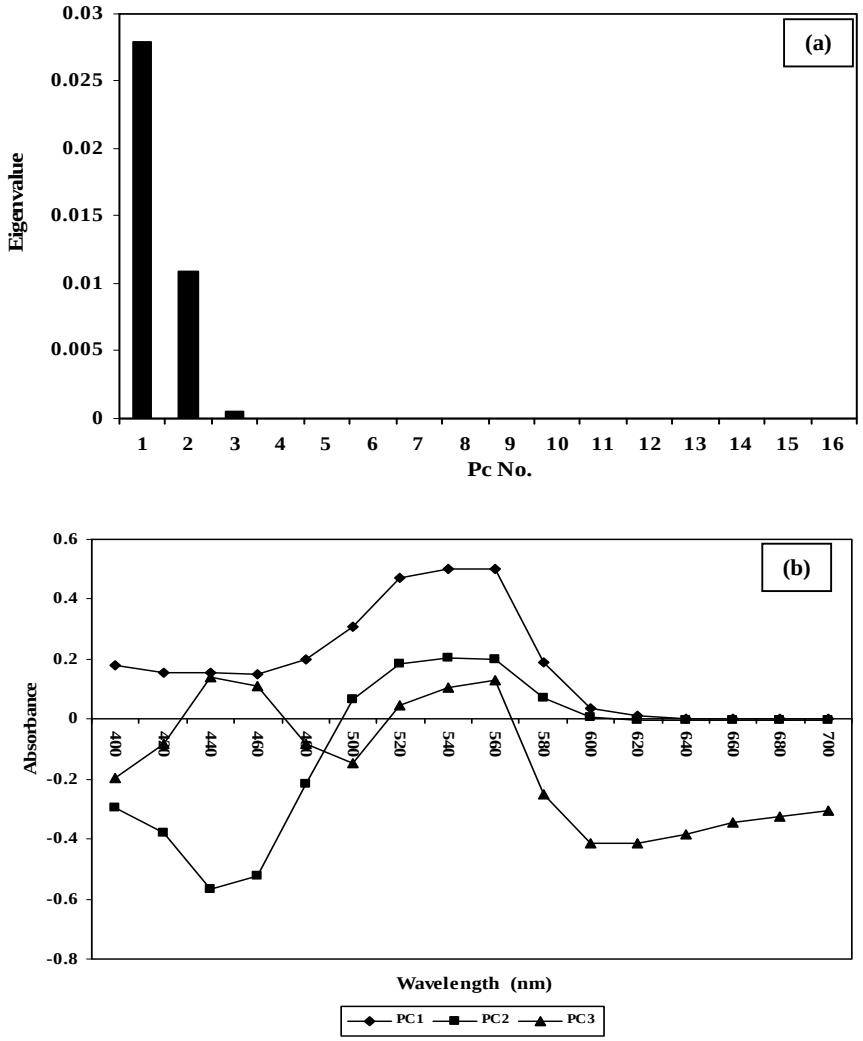

Figure 4: Eigenvalues (a) and first three principal components (b) of Blue-Yellow mixture.

4. Conclusions

The purpose of the present work is a novel technique for evaluating dye concentration in bi-component mixture. Initially, the absorbance spectra of training dataset are analyzed by using principal component analysis technique. Results indicate that the first three principal components are significant and others are unimportant. Then, the PCA-multiple polynomial technique with two and three principal components is used to calculate the dye concentration of bi-component mixture. Afterward, the dye concentration of bi-component mixture is evaluated by PCA-neural network with two and three principal components. The obtained results indicate that the accuracy of the estimation with three principal components is better than two principal components. Also, the performance of PCA-neural network is higher than PCA-multiple polynomial method.

Acknowledgement

The authors sincerely thank the Center of Excellence for Color Science and Technology for making this investigation possible.

5. References

1. N. M. Gary, Modern concepts of color and appearance, Science Publishers Inc., U.S.A, 2000.
2. T. Owen, Fundamentals of modern UV-Visible spectroscopy, agilent technologies, Germany, 2000.
3. R. Mc Donald, Color physics for industry, Society of Dyers and Colorists, 1997.
4. W.J. Jasper, E. V. Kovacs, G. I. Berkstresser, Using neural networks to predict dye concentrations in multiple-dye mixtures, *Text. Res. J.*, 63(1993), 545-551.
5. M. Marjoniemi, E. Mantysalo, Neuro-fuzzy modeling of spectroscopic data, part A: modeling of dye solutions, *J. Soc. Dyers Color.,* 113(1997), 13-17.
6. H. Logan, UV and visible spectrophotometry in organic chemistry, http://members.aol.com/ logan20/uv.html, (open accessed 5.10.09).
7. E. I. Stearns, The practice of absorption spectrophotometry, Wiley-Interscience, 1969.
8. E. Ekrami, A. Shams Nateri**,** Application of the ratio derivative method in quantitative analysis of samples dyed with ternary dye mixtures, *Prog. Color Colorants Coat.,* 1(2008), 1-9.
9. A. Hyvärinen, Survey on independent component analysis, *Neural Comput.,* 2(1999), 94-128.
10. I. S. Lindsay, A tutorial on principal components analysis, http://www.cs.otago.ac.nz /cosc453/student _tutorials/principal_components.pdf, (open accessed 5.10. 09).
11. E. Bingham, Advances in independent component analysis with applications to data mining, Helsinki university of technology, http://lib.hut.fi/Diss/ 2003/isbn9512268205, (open accessed 5.10.09).
12. H. S. Fairman, M. H. Brill, The principal components of reflectance, *Color Res. Appl.,* 29(2004), 104-110.
13. D. Y. Tzeng, R. S. Berns, A review of principal component analysis and Its applications to color technology, *Color Res. Appl.,* 30(2005), 84-98.
14. A. Shams-Nateri, Effect of standard colorimetric observer on the reconstruction of reflectance Spectra of colored fabrics, *Color. Technol.,* 124(2008), 14-18.
15. K. Ansari, S. H. Amirshahi, S. Moradian, Recovery of reflectance spectra from CIE tristimulus values using a progressive database selection technique, *Color. Technol.,* 122(2006), 128-134.
16. F. Agahian, S. H. Amirshahi. A new matching strategy: trial of the principal component coordinates, *Color Res. Appl.,* 33(2008), 10-18.
17. A. Shams-Nateri, Color matching by principal component analysis-based spectrophotometric technique, *Color. Technol.,* 125(2009), 36-42.
18. A. Shams-Nateri, Prediction of dye concentrations in a three-component dye mixture solution by a PCA-derivative spectrophotometry technique, *Color Res. Appl.,* 35(2010), 29-35.