Principal Component Analysis for Skin Reflectance Reconstruction

Kaida Xiao[1,2], Zhenghui Qin[1], Tushar Chauhan[2] , Changjun Li[1], and Sophie Wuerger[2]*
[1]School of Electronics and Information Engineering, University of Science and Technology Liaoning, China
[2]Department of Psychological Science, University of Liverpool, UK

Abstract

Principal component analyses (PCA) were conducted for skin spectral reflectance of a new skin colour database. Results for skin colours from different ethnic groups were analysed and reflectance re-construction based on first three components obtained by PCA were investigated. Significant differences in the derived principal components for the three ethnic groups were found. Furthermore, a new nonlinear optimization model based on all reflectance data was developed for generating the three basis functions for reflectance reconstruction. Comparison results for reflectance predictions show the basis functions obtained from the new optimisation model are better than those obtained from PCA.

Introduction

Interest in human skin colour has been greatly stimulated by the increased need for skin colour reproduction for graphic arts [1], skin analysis for cosmetic industry [2], skin-based diagnosis of cutaneous disease [3] and matching of skin colour for body and maxillofacial soft tissue prostheses [4, 5]. CIE Colorimetry [6] has been widely used for more than 80 years to provide objective measures for those applications. However, more accurate predictions of skin appearance can be made by taking into account the entire spectrum as opposed to the three colorimetric values; spectral-based calculations are particularly important for predicting skin appearance under varying illumination conditions. The purpose of this paper is to test a method for reconstructing skin spectra and assessing whether systematic differences exist between different ethnicities.

Previous research has focused on predicting spectral reflectance functions for each pixel of a facial image: spectral reflectance reconstruction algorithms were developed to transform either camera RGB or CIE XYZ to skin spectral reflectance [1,7]. The main result was that the three basis functions obtained with Principal Component Analysis (PCA) [8] are sufficiently accurate to describe the spectral reflectance of human skin [7] and therefore the spectral reflectance of each pixel of the skin image can be estimated from the values of three colour channels and the spectral radiance of the illuminant used.

The performance of skin reflectance reconstruction depends not only on the algorithm used, but also on the quality of skin colour database. While it is known that skin spectra differ between ethnic groups, it is poorly understood how these differences affect skin reflectance reconstruction algorithms and which components of the skin spectra best explain the ethnic differences. Moreover, for practical purposes, when the ethnicity is unknown, is important to know whether a set of fixed basis functions can provide a satisfactory reconstruction performance.

There are various skin colour databases available for different ethnic groups. However, none of the presently available skin colour databases is comprehensive. Since measurement instruments and protocols varied, direct comparisons between these databases are difficult.

In the present study, three recently established skin colour databases, collected in three countries using over 500 subjects and 9 body areas from each subject, were used. Crucially, the same protocol and measurement device (spectrophotometer [9]) was employed to obtain skin reflectance spectra for three ethnic groups. Principal Component Analysis was used to obtain the basis functions for each ethnic group. Performance of reflectance reconstruction was evaluated in terms of root-mean-square reflectance difference and the CIELAB colour differences under different illuminants. Based on these new skin colour databases, a new set of three basis functions was derived by minimising the predictive error for reflectance reconstruction. The model performance was further evaluated using a new set of skin data and satisfactory results were achieved.

Methodology

Acquisition of the Skin Colour Database

Skin spectra were obtained at six universities: University of Sheffield (UK), University of Liverpool (UK), Manchester Metropolitan University (UK), Beijing Institute of Technology (China), Beijing Institute of Graphic Communication (China) and University of Sulaimani (Iraq). Using a Minolta CM-2600d spectrophotometer and the SpectraMagic NX Colour Data Software, skin reflectance spectra (range: 360nm to 740 nm; 10nm spectral interval) were obtained. During the measurement, a viewing geometry of d/8 (diffuse illumination, 8-degree viewing) was used with the specular component included and the aperture size was set to 3mm. For each subject, nine body areas were measured: forehead, tip of nose, cheek, ear lobe, chin, back of hand, palm, outer forearm and inner forearm. For each subject, age, gender and ethnicity was recorded. To date, skin colours of 486 subjects covering three different ethnics groups have been obtained (Table 1) and are used in the current study.

Table 1. Skin colour data for different ethnics groups

	Caucasian	Chinese	Arabs
No. of Subject	138	202	146
No. of data	1242	1818	1314

Colour gamuts and colorimetric properties were reported in reference [10]; it is suffice to say that the skin colours of all three ethnicities show a significant overlap and discrimination based on the colorimetric skin properties is not easy to achieve.

 © 2014 Society for Imaging Science and Technology

Furthermore, a skin colour database provided by Spectromatch Limited was used to evaluate performance of skin reflectance reconstruction. These spectra were measured with a Konica Minolta CM2300d spectrophotometer with an aperture size of 11mm diameter and measurement geometry was set to the specular component included. In this set of data, spectral reflectances of 250 subjects were collected in a spectral range of 400 to 700nm at 10nm spectral intervals. The ethnicity of this data base is not known to us (due to confidential policy) and we will use this database only to evaluate the performance of the PCA-based reflectance reconstruction and the nonlinear optimisation model.

Principal Component Analysis

Let R be the skin colour database and it has m samples. Thus, R can be represented by an n by m matrix:

$$R = \begin{pmatrix} r^{(1)} & r^{(2)} & \cdots & r^{(m)} \end{pmatrix} \tag{1}$$

Each column of R is a reflectance having n components. The PCA method is to represent a reflectance r by principal component basis vectors $b^{(1)}, \ b^{(2)}, \ \cdots, \ b^{(n)}$, i.e.,

$$r = \alpha_1 b^{(1)} + \alpha_2 b^{(2)} + \cdots + \alpha_n b^{(n)} = B\alpha \qquad (2).$$

Here, the first principal component basis function $b^{(1)}$ has the most important characteristic (i.e. explains most of the variance) of the training set of reflectances, the second principal component basis vector $b^{(2)}$ has the second most important characteristic, and so on. For any given reflectance vector r, it can be uniquely represented by equation (2). The matrix B contains all basis vectors and the column vector α all coefficients α_j which are the coordinates of the reflectance r under the set of basis vectors; α_j can be represented uniquely as the inner product of jth basis vector $b^{(j)}$ and the reflectance r since the set of basis vectors is orthogonal.

In this study, basis functions of first three principal components were derived for skin colours of the Caucasian, Chinese and Arabs samples and the overall samples. These sets of basis functions will be further studied and used for reflectance reconstruction.

Reconstruction of the reflectance spectra

A reflectance reconstruction algorithm was developed to transform CIE XYZ tristimulus values to spectral reflectance spectra for human skin. For a given skin spectral reflectance, their CIE XYZ tristimulus values under standard D65 illuminant were calculated using the ASTM D65 weighing table.

Since only three basis functions are used, for any skin reflectance r we can express r as follows:

$$r \approx \beta_1 b^{(1)} + \beta_2 b^{(2)} + \beta_3 b^{(3)} = B_3\beta \tag{3}$$

Note that, if we know the reflectance r, it is better to choose the coefficient β_j to be its coordinates under the set of basis functions, i.e., $\beta_j = \alpha_j$ for j=1,2,3. However, in real applications, we have a set of tristimulus values X, Y and Z, while the corresponding reflectance (r) is unknown. We can express the relationship between the tristimulus values and the basis functions as follows:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = Mr = MB_3\beta \tag{4}$$

Where M, a 3xn matrix, is the ASTM 10nm weighting table obtained from the spectral power distribution of the CIE D65 and CIE 1931 colour matching functions. Thus, β can be calculated using equation (5).

$$\beta = (MB_3)^{-1} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{5}$$

Finally, using Equation (3), skin reflectance can be predicted based on pre-defined basis functions B_3 and the predicted coefficient vector β .

Performance evaluation of reflectance re-construction

To evaluate performance of skin reflectance reconstruction, spectral root mean square (RMS) errors were calculated between original skin reflectance (from the skin colour database) and the predicted skin reflectance spectra. Of particular importance for colour reproduction applications are the perceptual errors; we therefore calculated colour differences for each pair of skin spectra under four typical Illuminants (D50, A, CWF and F11). Note that, since CIE XYZ in Equation (4) were calculated using D65 weighting table, there is no colour difference under D65 illuminant.

Derivation of new basis functions

The basis vectors derived from the PCA analysis have useful properties (e.g. being orthogonal to each other) and provide an efficient way of summarising the main characteristics of the skin spectra. However, this does not guarantee that the reconstructed reflectance r based on Equations (3-5) is optimal since the estimated coefficients β_j given by Equation (5) depend on the given tristimulus values XYZ directly rather than the corresponding reflectance r and are generally different from the true coordinates α_j .

Now we want to find a new set of three basis functions $V_3 = \begin{pmatrix} v^{(1)} & v^{(2)} & v^{(3)} \end{pmatrix}$ and let r be any reflectance in the training dataset, i.e., be any column vector of the matrix R defined by equation (1). Thus as discussed in the above, the estimated reflectance $r^{(e)}$ based tristimulus values and this set of basis vectors can be given by:

$$r^{(e)} = V_3(MV_3)^{-1}(Mr) \tag{6}$$

Here, *Mr* factor is the tristimulus values for the given reflectance r under the D65/31. $(MV_3)^{-1}(Mr)$ is the estimated combination coefficients under the set of basis vectors V_3 . Hence we can determine the set of basis vectors V_3 so that is as small as possible. Note $\| p \|_2$ represents the Euclidian length of the vector p. Note also that there are 3 x n unknowns for the wanted basis vector set V_3 . Using only one reflectance r cannot generate a better set of basis vector. However, all the available skin reflectance spectra can be used for deriving the new set of basis vector V_3 , which leads to the nonlinear optimization problem:

$$\text{minimise}\{\frac{1}{m}\sum_{k=1}^{m}\sqrt{\frac{1}{m}\| V_3(MV_3)^{-1}(Mr^{(k)}) - r^{(k)} \|_2^2}\} \tag{7}$$

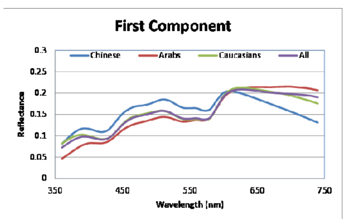
where, all $r^{(k)}$ are the known reflectance functions listed in Table 1, m is the total number of reflectance functions, M is the ASTM 10-nm weighting table under D65/31 and V_3 is the set of basis vector having 3n unknowns to be determined by the optimization. MATLAB fmincon function was used for solving the above nonlinear unconstrained optimization problem.

Compared with the PCA approach, the new basis vectors can not only optimize the match between the reconstructed and the original reflectance spectra under D65/31, but also make them more similar in shape, which are confirmed by the comparison results below.

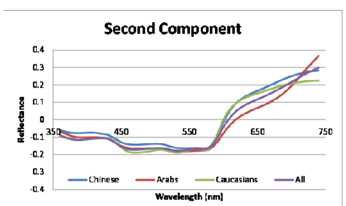
Results and Analysis

Basis functions from PCA

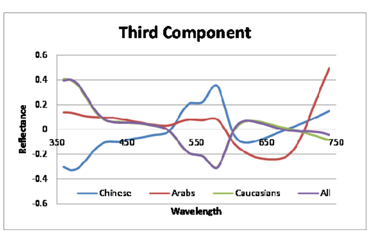
Basis functions were obtained using PCA for skin colours for each individual ethnic group (Chinese, Caucasians and Arabs) and for the combined set ('ALL'). The first, second and third basis vectors are shown in Figure 1a, 1b and 1c. The first principal component for Caucasians is very similar to that for 'All', whereas basis functions for Chinese and Arabs are quite different. For the second principal component, all basis functions show smaller variations in the short wavelength, but large variations in the long-wavelength range. Largest differences are observed in the third principal component, although the component for the Caucasians is very similar to that for 'ALL'.



(a)



(b)



(c)

Fig.1 *First (a), second (b) and third (c) basis functions for each skin database derived using PCA*

Table 2. The cumulative coverage percentages of first three principle component

CCP (%)	First Component	Second Component	Third Component	Sum
Chinese	84.1	12.3	2.2	98.5
Arabs	86.6	7.9	3.1	97.6
Caucasians	87.4	8.4	2.6	98.5
Combined	85.4	9.4	2.3	97.1

The cumulative coverage percentages (CCP) for the first three basis functions were calculated and are shown in Table 2. The results in Table 2 indicate that the first three basis functions account for more than 97% variance for all spectral data for different skin database.

New Basis functions

Using all the three skin colour spectral reflectance data, a new set of basis functions (denoted by BNEW) was derived using the nonlinear optimization method defined by Equation (7). Comparisons of the new basis functions (BNEW) with the ones based on PCA (BPCA) are shown in Figure 2a, 2b and 2c: the best agreement is found in the first component, whereas slight differences are apparent in the second and third components.

Evaluation of Reflectance Reconstruction

Based on those basis functions, skin reflectance reconstructions were performed to predict skin spectral reflectance from CIE XYZ tristimulus values using equations (4) and (5). CIE XYZ tristimulus values were calculated using ASTM weighting table to simulate stimuli of reference reflectance under a standard D65 illuminant. Performance of reflectance reconstruction was evaluated using both RMS difference of spectral reflectance and colour difference under four typical illuminants. By applying different testing skin colour database, performances with different basis functions were evaluated as follows:

(1). Basis functions from PCA for each training set of skin colours

Based on four set of basis functions derived by PCA, performance of reflectance re-construction was evaluated using each corresponding training skin colour database. The results in terms of spectral and colour differences are listed in Table 3. Last column shows the mean of the four median colour differences.

Table 3. Performance of reflectance re-construction using different sets of basis functions obtained using PCA

Basis functions	RMS Error (Reflectance)	Colour Difference (CIE ΔE*ab)				
		A	D50	CWF	F11	Mean
Chinese	0.017	0.27	0.09	0.29	0.39	0.26
Caucasians	0.028	0.50	0.16	0.44	0.62	0.43
Arabs	0.223	1.99	0.77	1.54	3.50	1.95
Combined	0.016	0.28	0.10	0.33	0.53	0.31

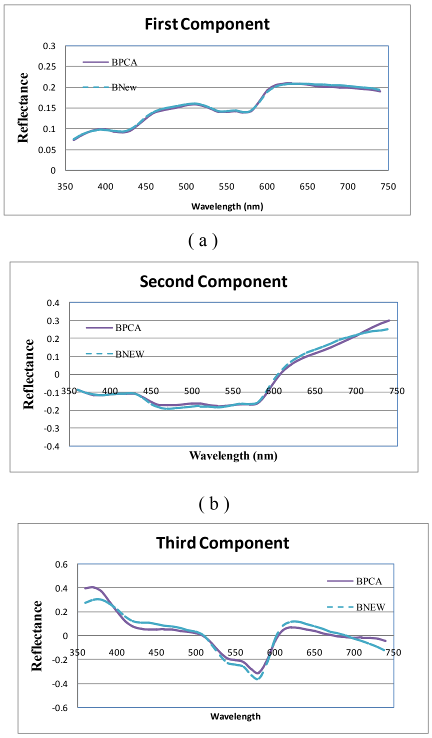
From Table 3, it can be seen that spectral difference (RMS) is between 0.016 and 0.223 and the mean colour difference under the four illuminants is between 0.26 and 1.95 ΔE*ab units, indicating a good reconstruction performance. Although all sets of basis functions were derived using the PCA approach, they perform differently when they are used for reconstructing reflectance. The basis functions derived from Chinese skin provided the best performance overall, whereas basis

© 2014 Society for Imaging Science and Technology

functions based on the Arabs skin data gave the worst prediction. Good predictions were achieved for all skin colours, based on the combined ethnic skin data set. Furthermore, when D65 was used as a reference illuminant, the reflectance reconstruction gave the smallest colour difference under D50 and the largest colour difference under illuminant F11.

(2). New Basis function obtained using nonlinear optimisation

Based on basis functions Bnew, reflectance re-construction was evaluated for all three individual and the combined skin database (Table 4).



(c)

Fig. 2. First (a), second (b) and third (c) basis functions for the two sets of basis functions B_{PCA} and B_{NEW}

Table 4. Performance of reflectance re-construction using basis functions (BNEW) derived using the nonlinear optimization model trained using the combined set.

New Basis functions	RMS Error (Reflectance)	Colour Difference (CIE ΔE*ab)				
		A	D50	CWF	F11	Mean
Chinese	0.008	0.21	0.08	0.22	0.32	0.21
Caucasians	0.012	0.38	0.14	0.32	0.57	0.35
Arabs	0.017	0.31	0.10	0.26	0.48	0.29
Combined	0.012	0.28	0.10	0.26	0.42	0.26

Comparing results from Table 4 with that from Table 3, it is clear that both the RMS Difference and the colour difference is reduced, indicating performance of reflectance re-construction could be better for all skin spectra when the new set of basis functions was used.

(3). Test of a new skin colour database

Skin spectra from the Spectromatch skin colour database were used to evaluate proposed skin reflectance re-construction algorithm. Since the ethnicities are unknown to us, the combined training data sets were used. The previously derived sets of basis functions (BPCA and BNEW) were used to predict this new set of reflectance data. All results are listed in Table 5. Table 5 shows that the average RMS difference is approximately 0.02 and the average colour difference is approximately 0.5 ΔE*ab units. Taken into consideration that different measurement instruments were used for two different skin colour database, it is clear that proposed method has a good performance for reflectance re-construction of human skin. Results also show that, using the newly developed basis functions, a better performance is achieved in terms of for both RMS difference and colour difference for all four illuminants.

Table 5. Performance of reflectance re-construction for the Spectromatch skin colour database using the two sets of basis functions BPCA and BNEW

Basis functions	RMS Error (Reflectance)	Colour Difference (CIE ΔE*ab)				
		A	D50	CWF	F11	Mean
B_{PCA}	0.023	0.49	0.18	0.49	0.86	0.51
B_{NEW}	0.016	0.49	0.18	0.42	0.83	0.48

Conclusion

In this paper, Principal Component Analysis was used to identify basis functions for a newly acquired skin colour database. Reflectance re-construction was performed to transform CIE XYZ tristimulus values to spectral reflectance. It was found that the selection of basis functions can significantly affect the overall performance of skin reflectance re-construction. Based on the skin colour database, a set of basis functions was derived to provide the smallest RMS difference between the original and the predicted skin spectra. The performance of the proposed method for reflectance re-construction was evaluated using a different skin colour database and achieved a satisfactory result.

Acknowledgment

This project was partially supported by the Royal Academic of Engineering for Research Exchanges with China and India, EPSRC grant EP/K040057/1 and Spectromatch Limited, UK. The work of Changjun Li was also partially supported by National Natural Science Foundation of China (Grant number: 61178053).

References

[1]. Q. Sun and M.D. Fairchild, "Statistical characterization of face spectral reflectances and its application to human portraiture spectral estimation," Journal of Imaging Science and Technology, 46 498-506 (2003).

[2]. K. Kikuchi, Y. Masuda, T. Hirao, Imaging of hemoglobin oxygen saturation ratio in the face by spectral camera and its application to evaluate dark circles, Skin research and technology, 19, 499-507, 2013.

[3]. Leonardi, A, Buonaccorsi, S, Pellacchia, V, Moricca, LM, Indrizzi, E and Fini G, "Maxillofacial prosthetic rehabilitation using extraoral implants". J Craniofac Surg, 19(2), pp. 398-405 (2008).

[4]. K. Xiao, F. Zardawi, R. van Noort R. J.M Yates, Colour reproduction for advanced manufacture of soft tissue facial prostheses, Journal of Dentistry (2013), 41:e15-e23.

[5]. K. Xiao K, F. Zardawi, R. van Noort R, J.M Yates, Developing a 3D colour image reproduction system for additive manufacturing of facial prostheses, International Journal of Advanced Manufacturing Technology, 70, 2043-2049,2014.

[6]. CIE Publication Nos. 15.3, "Colorimetry", (2004).

[7]. F. H. Imai, N. Tsumura, H. Haneishi and Y. Miyake, "Principal Component Analysis of Skin Color and Its Application to Colorimetric Color Reproduction on CRT Display and Hardcopy," J. of Imag. Sci. and Tech., 40, pp. 422-430, (1996).

[8]. H. Abdi, L.J. Williams, L.J, "Principal component analysis.". Wiley Interdisciplinary Reviews: Computational Statistics, 2: 433–459, (2010).

[9]. K. Xiao, N. Liao, H. Liu, F. Zardawi, R V Noort, Z. Yang, M. Huang, J.M Yates, An investigation of Chinese skin colour and appearance for skin colour reproduction, Chinese Optics Letter (2012), 10: 083301 1-5.

[10]. K. Xiao, F. Zardawi, R.V. Noort and J.M. Yates, Investigation of colour and appearance for human skin, Proceeding of the midterm meeting of the International Colour Association (AIC), Taipei, Taiwan, (2012).

Author Biography

Kaida Xiao received his PhD in colour science from the Colour and Imaging Institute, University of Derby (2007). Since then he worked in the Samsung Advanced Institute of Technology, Tru-Colour Limited, The University of Sheffield and now he is working in University of Liverpool as a Senior Research Fellow. His research interest covers colour appearance, image quality enhancement, 3D colour image reproduction and colour and appearance in medicine and dentistry.

© 2014 Society for Imaging Science and Technology