Improved method for skin reflectance reconstruction from camera images

Kaida Xiao,[1,2,3] Yuteng Zhu,[2,4] Changjun Li,[5,*] David Connah,[6] Julian M Yates,[7] and Sophie Wuerger[2]

[1]*School of Electronics and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, China*
[2]*Department of Psychological Sciences, University of Liverpool, Liverpool, L69 7ZA, United Kingdom*
[3]*School of Design, University of Leeds, Leeds, LS2 9JT, United Kingdom*
[4]*College of Optical Engineering, Zhejiang University, Hangzhou, 310027, China*
[5]*College of Physics and Electronic Information Engineering, Wenzhou University, Wenzhou, 325035, China*
[6]*Centre for Visual Computing, University of Bradford, Bradford, BD7 1DP, United Kingdom*
[7]*School of Dentistry, University of Manchester, M13 9PL, United Kingdom*
[*]*cjliustl@sina.com*

Abstract: A improved spectral reflectance reconstruction method is developed to transform camera RGB to spectral reflectance for skin images. Rather than using conventional direct or two-step processes, we transform camera RGB to skin reflectance directly using a principal component analysis (PCA) approach. The novelty in our direct method (RGB to spectra) is the use of a skin-specific colour characterisation chart with spectra closer to human skin spectra, and a new database of skin reflectances to derive the PCA bases. The experimental results using the facial images of 17 subjects demonstrate that our new direct method gives a significantly better performance than conventional, two-step methods and direct methods with traditional characterization charts. This new spectral reconstruction algorithm is sufficiently precise to reconstruct spectral properites relating to chromophores and its performance is within the acceptable range for maxillofacial soft tissue prostheses (error < 3 ΔE^*_{ab} units).

OCIS codes: (330.1710) Color, measurement; (110.4234) Multispectral and hyperspectral imaging; (300.6550) Spectroscopy, visible.

References and links

1. M. Uzair, A. Mahmood, F. Shafait, C. Nansen, and A. Mian, "Is spectral reflectance of the face a reliable biometric?" Opt. Express **23**(12), 15160–15173 (2015).
2. Q. Sun and M. D. Fairchild, "Statistical characterization of face spectral reflectances and its application to human portraiture spectral estimation," J. Imaging Sci. Technol. **46**, 498–506 (2002).
3. K. Kikuchi, Y. Masuda, and T. Hirao, "Imaging of hemoglobin oxygen saturation ratio in the face by spectral camera and its application to evaluate dark circles," Skin Res. Technol. **19**(4), 499–507 (2013).
4. A. Leonardi, S. Buonaccorsi, V. Pellacchia, L. M. Moricca, E. Indrizzi, and G. Fini, "Maxillofacial prosthetic rehabilitation using extraoral implants," J. Craniofac. Surg. **19**(2), 398–405 (2008).
5. K. Xiao, F. Zardawi, R. van Noort, and J. M. Yates, "Developing a 3D colour image reproduction system for additive manufacturing of facial prostheses," Int. J. Adv. Manuf. Technol. **70**(9-12), 2043–2049 (2014).
6. I. Nishidate, T. Maeda, K. Niizeki, and Y. Aizu, "Estimation of Melanin and Hemoglobin Using Spectral Reflectance Images Reconstructed from a Digital RGB Image by the Wiener Estimation Method," Sensors (Basel) **13**(6), 7902–7915 (2013).
7. F. H. Imai and R. S. Berns, "Spectral estimation using trichromatic digital cameras," in *International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives* (Chiba University, 1999), pp. 42–49.
8. F. H. Imai, "Multi-spectral image acquisition and spectral reconstruction using a trichromatic digital camera system associated with absorption filters," in *MCSL Technical Report* (1998).
9. X. Zhang, Q. Wang, J. Li, X. Zhou, Y. Yang, and H. Xu, "Estimating spectral reflectance from camera responses based on CIEXYZ tristimulus values under multi-illuminants," Color Res. Appl. in press (2016).
10. H.-L. Shen, P.-Q. Cai, S.-J. Shao, and J. H. Xin, "Reflectance reconstruction for multispectral imaging by adaptive Wiener estimation," Opt. Express **15**(23), 15545–15554 (2007).
11. H.-L. Shen and J. H. Xin, "Spectral characterization of a color scanner based on optimized adaptive estimation," J. Opt. Soc. Am. A **23**(7), 1566–1569 (2006).

12. H. Haneishi, T. Hasegawa, A. Hosoi, Y. Yokoyama, N. Tsumura, and Y. Miyake, "System design for accurately estimating the spectral reflectance of art paintings," Appl. Opt. **39**(35), 6621–6632 (2000).
13. K. Barnard and B. Funt, "Camera characterization for color research," Color Res. Appl. **27**(3), 152–163 (2002).
14. C. Li, G. Cui, and M. R. Luo, "The accuracy of polynomial models for characterizing digital cameras," in *Proceedings of AIC2003 Bangkok: Color Communication and Management* (2003), pp. 166–170.
15. V. Cheung, S. Westland, C. Li, J. Hardeberg, and D. Connah, "Characterization of trichromatic color cameras by using a new multispectral imaging technique," J. Opt. Soc. Am. A **22**(7), 1231–1240 (2005).
16. M. Shi and G. Healey, "Using reflectance models for color scanner calibration," J. Opt. Soc. Am. A **19**(4), 645–656 (2002).
17. 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. Imaging Sci. Technol. **40**, 422–430 (1996).
18. S. Chen and Q. Liu, "Modified Wiener estimation of diffuse reflectance spectra from RGB values by the synthesis of new colors for tissue measurements," J. Biomed. Opt. **17**(3), 030501 (2012).
19. H.-L. Shen and J. H. Xin, "Estimation of spectral reflectance of object surfaces with the consideration of perceptual color space," Opt. Lett. **32**(1), 96–98 (2007).
20. J. Tian and Y. Tang, "Wavelength-sensitive-function controlled reflectance reconstruction," Opt. Lett. **38**(15), 2818–2820 (2013).
21. K. Xiao, J. M. Yates, F. Zardawi, S. Sueeprasan, N. Liao, L. Gill, C. Li, and S. Wuerger, "Characterising the variations in ethnic skin colours: a new calibrated data base for human skin," Skin Res. Technol. in press (2016).
22. ASTM E308–13. Standard practice for computing the colors of objects by using the CIE system. American Society for Testing and Materials (ASTM; 2013).
23. R. Shrestha, R. Pillay, S. George, and J. Y. Hardeberg, "Quality evaluation in spectral imaging - Quality factors and metrics," J. Int. Colour Assoc. **10**, 22–35 (2014).
24. M. A. Changizi, Q. Zhang, and S. Shimojo, "Bare skin, blood and the evolution of primate colour vision," Biol. Lett. **2**(2), 217–221 (2006).
25. R. D. Paravina, G. Majkic, M. Del Mar Perez, and S. Kiat-Amnuay, "Color difference thresholds of maxillofacial skin replications," J. Prosthodont. **18**(7), 618–625 (2009).

1. Introduction

Measuring the spectral reflectance of human skin has gained importance due to the increasing number of applications that need an accurate spectral representation of skin colour, including biometric identification based on spectral information [1], skin colour reproduction for graphic arts [2], skin pigmentation predictions for the cosmetics industry [3], skin colour analysis for the diagnosis of cutaneous diseases [4], and skin colour matching for maxillofacial soft tissue prostheses [5]. In contrast to colorimetric skin values, skin spectral reflectance data contain the necessary information to model skin appearance under different illumination conditions and also allow the extraction of information about skin chromophores, which is important for detecting various skin diseases including cancers, monitoring health status and tissue metabolism, and evaluating convalescence [6]. Since high-resolution RGB camera images are readily available compared to spectral measurements, a methodology that recovers the skin reflectance spectra from a camera image is important for the above applications. An algorithm that accurately reconstructs the spectral reflectance for each camera pixel, results in a high-resolution multi-spectral representation of a skin patch containing the full information about the subtle spatial texture as well as the spectral properties of skin.

To transform the device-dependent digital RGB camera signals to spectral reflectance, two main classes of methods exist (cf. Table 1), direct methods mapping RGB directly into spectral reflectance [7, 8], and 2-step methods [9] which map the camera RGB signals into spectra, but requiring an intermediate step: camera sensitivities are either known [10–12], or need to be estimated [13]. The two steps, camera characterization and spectral reconstruction are optimized separately, usually using different sets of training colours.

The choice of training colours is crucial for both, reflectance reconstruction and camera characterization, and performance is generally better if training and test samples are as similar as possible [14]. However, large differences in colorimetric and spectral properties between human skin and conventional training colour charts currently exist. Human skin is a multi-layer material and skin colour is determined by melanin, blood concentration and oxygenation. These material differences between training and test spectra are likely to affect the spectral reconstruction performance. Similarly, for camera characterization, colour charts

that cover the entire gamut of visible colours are likely to introduce errors in the reconstruction performance for skin colours, with the latter only occupy a small area in colour space.

Reflectance reconstruction techniques such as Wiener estimation [11, 12] and finite-dimensional modeling [15, 16] are widely used in multispectral imaging systems. For skin spectral reconstruction, Imai and colleagues [17] have applied finite-dimensional modelling to skin images and concluded that three basis functions obtained with Principal Component Analysis (PCA) are sufficiently accurate to describe the spectral reflectance of human skin. Chen and Liu [18] developed a modified Wiener estimation to predict skin spectra for tissue measurements.

The purpose of this paper is to develop a more accurate spectral reconstruction method of camera RGB images, optimized for skin images. The novelty in our approach is the simultaneous optimization of both steps (cf. Table 1) and the use of a skin-specific calibration chart and a skin spectra database. Performance for the conventional and the new method are evaluated by predicting the skin spectra (2 facial areas) of 17 subjects. Two performance measures will be used, the root-mean-square error applied to the reflectance spectra and a perceptual error measurement taking into account the properties of the visual system [19, 20].

2. Methods

2.1. Conventional 2-step method

2.1.1. For the characterisation step, polynomial regression [14] is often used to derive the transformation from RGB to XYZ, as defined in Eq. (1)

$$\begin{aligned} X &= \sum_{0\le j_1+j_2+j_3\le m} a_{X,j_1,j_2,j_3} R^{j_1} G^{j_2} B^{j_3} \\ Y &= \sum_{0\le j_1+j_2+j_3\le m} a_{Y,j_1,j_2,j_3} R^{j_1} G^{j_2} B^{j_3} \\ Z &= \sum_{0\le j_1+j_2+j_3\le m} a_{Z,j_1,j_2,j_3} R^{j_1} G^{j_2} B^{j_3} \end{aligned} \tag{1}$$

where R, G and B are camera signals and *X, Y* and *Z* are the CIE tristimulus values; *m* is the order of the polynomial and j_1, j_2, j_3 are nonnegative integer indices; *a* are the model coefficients to be determined via minimizing the *X, Y* and *Z* approximation error for a characterisation chart with known colorimetric values (e.g. the X-Rite ColorChecker). Performance of the above transform is affected by the order *m* of the polynomial [14]. The MATLAB solver: lsqlin is used for training the polynomial model.

2.1.2. For the reflectance reconstruction step, CIE XYZ tristimulus values are transformed to spectral reflectance estimates based on the following Eq. (2)

$$u = W^T r \tag{2}$$

where *u* is a vector of the tristimulus values *(XYZ)*, *W* is the *n* by 3 matrix formed by a weighting table under a particular illuminant and CIE standard observer, *T* is the transpose, and *r* is a column vector with *n* components, representing reflectance. *n* is 31 since the wavelength range from 400nm to 700nm is sampled uniformly at 10nm intervals.

Equation (2) describes three linear equations, with *n* (= 31) unknowns. The problem is therefore underdetermined, and there are an infinite number of solutions for *r*. However, the true dimensionality of skin reflectance is much lower than 31 which can be used to constrain the solution space. Principal component analysis (PCA) applied to skin reflectances has shown that skin spectra can be represented by three basis functions $U_3 = (u_1, u_2, u_3)$ [17], hence Eq. (2) can be rewritten as follows:

$$u = W^T U_3 \alpha \tag{3}$$

where α is a column vector with three components. We can then solve for α and the reflectance r using Eq. (4):

$$\alpha = (W^T U_3)^{-1} u \text{ and } r = U_3 \alpha \quad (4)$$

In principle, more than three basis vectors can be used in a generalised inverse approach defined by Eq. (3), but when applied to skin images the method performs best with three basis vectors [21].

These two processing steps, i.e. the camera colour characterisation and the subsequent reflectance reconstruction, are completely independent and the overall error of the two-step method is therefore the sum of the individual errors. Moreover, reflectance samples used to characterise the camera are often drawn from different sources to those used to derive the linear model, and hence have different statistical distributions. As a result, reducing characterization error for a given set of training patches, may not necessarily improve spectral reconstruction error for a distinct set of test surfaces.

In summary, for the implementation of the conventional 2-step method, a chart such as the X-Rite ColorChecker 'SG' is needed to obtain the camera RGBs. The chart can also be measured for obtaining the reflectance set and its associated tristimulus values under a particular viewing condition. Thus, the polynomial model P can be built based on the camera RGBs and XYZs of the chart (see Eq. (1)). Next, we can use the measured reflectance set from the chart or a large database from any application area under consideration, to obtain the three basis functions u_1,u_2,u_3, hence the matrix U_3 in Eq. (3). Let W (n by 3 array) be the weigthing table (ASTM method [22]) for the selected viewing conditions, and u be the vector formed from a given XYZ. Now for any camera RGB, the u vector or XYZ can be predicted using the trained model P. Using Eq. (4), the reflectance r can be estimated.

2.2. Modified method: Silicon skin chart and calibrated skin database

We propose and evaluate three significant modifications of existing methods to enhance the accuracy of reflectance reconstruction for skin images. The modifications are summarized here and explained in more detail in the next section.

(1) Instead of using a standard colour chart (X-Rite ColorChecker, 'SG'; cf. Table 1) which includes a range of 140 skin colours, we use a new silicone skin colour chart ('SS'; cf. Table 1) consisting of 90 colour samples. This silicon skin colour chart was developed by Spectromatch Ltd. to provide an accurate reference chart for soft tissue prostheses applications. The colour specifications (in CIELAB space) for both charts are plotted in Fig. 1, redness-greeness versus yellowness-blueness (a*-b* diagram) on top and chroma/lightness (C*-L* diagram) on the bottom. On the left the colours of the X-Rite ColourChecker are shown ('SG'), in the middle the silicon skin samples ('SS', Spectromatch Ltd), and on the right a sample of 100 typical skin colours (sampled from three ethnicities, Caucasian, Kurdish, and Chinese). The reflectance spectra of the 90 silicon skin samples are shown in Fig. 2.

(2) We use a new skin reflectance database [21] to improve the quality of the linear model derived by PCA ('DB', cf. Table 1). The skin reflectance database consists of 4392 colorimetric and corresponding spectrophotometric measurements; nine body areas of 482 subjects from three ethnic groups (Caucasians, Chinese and Kurdish) were measured using a Minolta CM-2600d spectrophotometer (direct contact measurement), in a spectral range from 400 to 700 nm with a 10 nm spectral interval.

(3) Rather than mapping RGB to XYZ and then to reflectance as a two-step process, we map RGB direct to reflectance ('Direct', cf. Table 1). For the implementation of these direct methods (SS/SS and SS/DB), we need a chart as explained in detail in the next section. Camera RGBs and the reflectance functions are obtained from the chart and the PCA is used to calculate the basis functions $U_K = (u_1,u_2,...,u_k)$ from the measured chart reflectances (SS/SS method in the Table 1) or the reflectances in our

database DB (SS/DB method in the Table 1). For each chart reflectance r, the best coordinate vector β is obtained from Eq. (5). Thus, for each colour from the chart, we have the camera RGB and the β hence the polynomial model P can be trained to map the RGB to the best coordinate vector β (see Eq. (6) and upper part of Fig. 3). Using the trained polynomical model P we can now estimate, for any given camera RGB, the associated best coordinate vector β. The reflectance function r can be obtained using the basis function U_k and the estimated best coordinate β and Eq. (7) as shown in the lower part of Fig. 3.

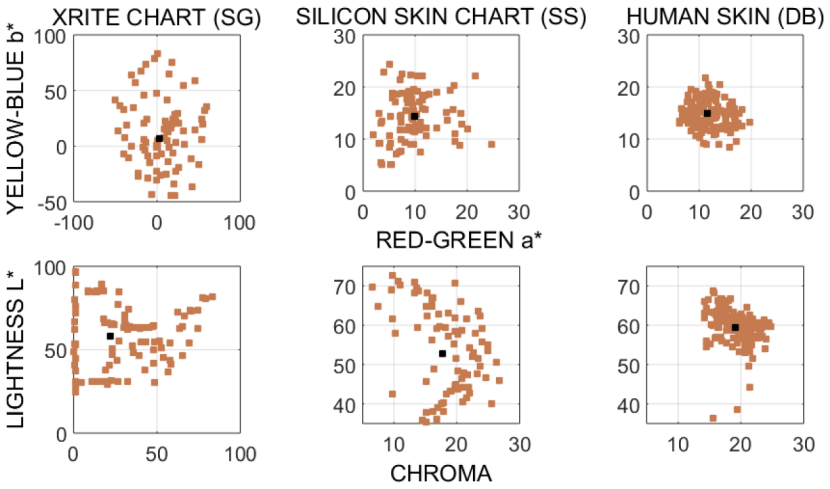

Fig. 1. Colour specifications in CIELAB of the two different colour charts used for calibration. (a) on the left: the 140 X-Rite SG colour samples, (b) in the middle, the 90 Silicon skin chart (Spectromatch Ltd), and (c) on the right, a subset (100) of typical human skin colours. Please note that the scales are different between (a) and (b)/(c) to visualize the subtle but important differences between the Silicon and the human skin colours.

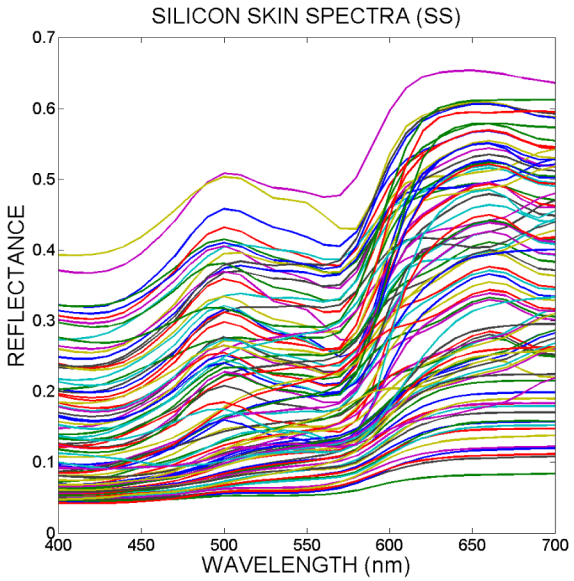

Fig. 2. Spectra of Silicon skin charts (SS), provided by Spectromatch Ltd.

Table 1. A summary of the properties of the four methods

METHOD	Recovery Method	Calibration Chart	PCA Source data
Two-step SG/DB	RGB->XYZ-> Reflectance	X-Rite SG	Skin Database (DB)
Two-step SS/DB	RGB->XYZ-> Reflectance	Spectromatch Silicon (SS)	Skin Database (DB)
Direct SS/SS	RGB-> Reflectance	Spectromatch Silicon (SS)	Spectromatch Silicon (SS)
Direct SS/DB	RGB-> Reflectance	Spectromatch Silicon (SS)	Skin Database (DB)

Table 1 shows the four different methods to be evaluated, with the first two rows being the convential methods. The first conventional two-step approach ('Two-step SG/DB') uses the X-Rite ColorChecker SG chart (with 140 coloured patches) as a calibration chart. The reference skin colour chart is imaged with a camera and measured with a spectrophotometer, so that for ech colour in the chart the camera RGB data and the reflectance, hence, the *XYZ* are available. Mapping from RGB to XYZ can be computed using Eq. (1). Then the n basis functions $u_1, u_2, ..., u_n$ are derived from the skin reflectance database using PCA; the first K basis functions form a matrix denoted by U_K. This matrix is used in Eqs. (3) and (4) with K = 3 to map from XYZ (u vector transformed from RGB) to reflectance (r). Note that, for each reflectance r from the chart, the best first K ($K \geq 3$) coordinates (denoted by a column vector β) under this set of basis functions are computed as:

$$\beta = (U_K)^T r \tag{5}$$

However, the two-step SG/DB method computes the coordinate vector α using Eq. (4), which is an approximation to the best coordinate vector β.

The 2nd modified two-step method ('Two-step SS/DB') makes use of the silicon skin colour chart instead of using the X-rite SG chart (Table 1). Otherwise it is identical to the first method.

The 3rd method also uses the silicon skin colour chart ('Direct SS/SS'), but maps the camera RGB directly into the vector $\hat{\beta}$ (Fig. 3), where $\hat{\beta}$ is again an approximation to the best coordinate vector β. We use a simple modification of the polynomial of Eq. (1), which maps RGB to the first K ($K \geq 3$) reflectance basis weights of vector $\hat{\beta}$ using polynomial regression (Eq. (6):

$$\hat{\beta}_i = \sum_{0 \leq j_1 + j_2 + j_3 \leq m} a_{\beta_i, j_1, j_2, j_3} R^{j_1} G^{j_2} B^{j_3} \tag{6}$$

The polynomial regression (Eq. (6) uses as input the RGB values of the Silicon Skin chart ('SS'). To derive the basis functions, the silicon skin spectra are used ('SS'). Reflectance r is then predicted using Eq. (7):

$$r = U_K \beta \tag{7}$$

Finally, to evaluate the effect of the skin database, we use the human skin database in method #4 ('Direct SS/DB') instead of the silicon skin spectra chart to derive the basis functions (Eqs. (5) and (7)).

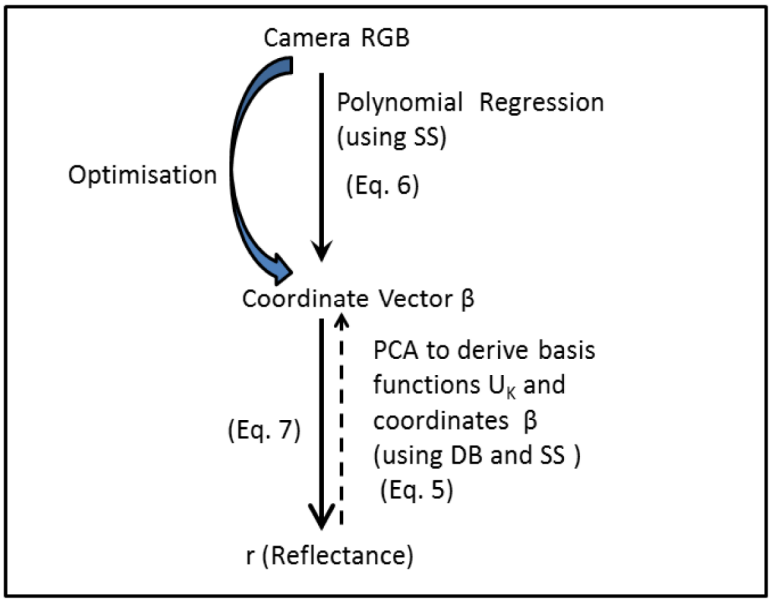

Fig. 3. Direct recovery method using the skin database (DB) and the silicon skin chart (SS)

2.3. Evaluation

2.3.1. Test images and skin spectra

To evaluate the performance of each method, test data were obtained using a skin imaging system consisting of a large viewing cabinet and a digital camera. The viewing cabinet (Verivide Ltd) is sufficiently spacious for a subject to sit inside, and provides a consistent daylight environment. A Nikon D7000 digital SLR camera controlled by the DigiEye system software, was used to capture images with an aperture = 7.1, ISO = 100, and the white balance fixed at D65. A Nikon AF-S DX 18-105 mm lens was used, with a fixed focus of 50mm.The image capture distance between camera lens and training colour charts (or the subject's face) was fixed to 57.5cm and the capture angle was 0 degrees. After image capture, a uniformity correction was performed to compensate for the non-uniformity of illumination. A Konica Minolta (CM700d) spectrophotometer was used to measure spectral reflectance (400 nm to 700 nm; 10nm intervals) for both the training chart and the facial skin areas of the subjects. Facial images for 17 subjects were captured and the reflectance of their foreheads and cheeks were measured, resulting in 34 skin reflectance measurements (Fig. 4(b)). The subjects consisted of 8 Caucasians, 8 Chinese and 1 Indian and were either lab members or academic staff associated with the lab. The faces of the subjects were captured as shown in Fig. 4(a).

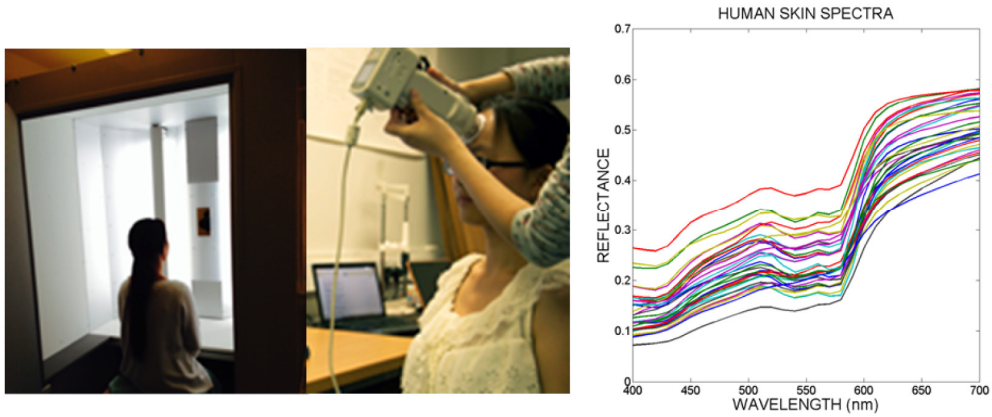

Fig. 4. (a) On the left the capture of the facial camera image is shown while subjects are seated in the viewing cabinet. Spectral reflectances are obtained using a CM700d spectrophotometer. (b) 34 Test skin spectra (forehead and cheek) for 17 individuals

2.3.2. Evaluation procedure

The four methods were evaluated in the following way. For each image of the 34 images, the RGB values were averaged across all pixels and then transformed to skin spectral reflectance using each of the four methods (Table 1). Assuming a standard observer (*CIE* 1931 standard *2 deg* observer), the CIELAB colour differences between the predicted skin spectra and the measured skin spectra were calculated. This colour difference reflects a perceptual difference and is often used as measure for the goodness of fit [19]. The mean of the median colour differences under each of the five illuminants (A, D50, F02, F11, and D65) is used as the overall performance of each method. Using multiple illuminants provides a robust performance assessment since it guards against obtaining small errors for metameric, rather than spectral matches.

Two spectral shape differences were also used for the evaluation of the spectral differences between the original ($r^{(o)}$) and the reconstructed ($r^{(c)}$) spectra: the root mean square error (RMSE) and the spectral similarity value (SSV) [23], using Eqs. (8) and (9):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(r_k^{(o)} - r_k^{(o)})^2} \tag{8}$$

$$\mathrm{SSV} = \sqrt{\mathrm{RMSE}^2 + \mathrm{S}^2} \tag{9}$$

Here n is the number of the components of the reflectance. The square of S is defined by Eq. (10)

$$S^2 = 1 - \{[\frac{1}{n}\sum_{k=1}^{n}(r_k^{(o)} - \mu^{(o)})(r_k^{(c)} - \mu^{(c)})] / (\sigma^{(o)}\sigma^{(c)})\}^2 \tag{10}$$

Here, $\mu^{(c)}$ and $\mu^{(o)}$ are the means of the reconstructed and original reflectances, and $\sigma^{(c)}$ and $\sigma^{(o)}$ are the standard deviations of the reconstructed and original reflectances.

The accuracy of skin reflectance reconstruction depends not only on the type of the characterisation chart and the spectral skin database but also on the order of the polynomial (*m*), and the number of basis functions *(K)* in the linear model. To investigate how those parameters affect the model performance, the polynomial order *m* is varied from 1 to 5 and the number of basis functions *K* from 3 to 6 respectively and the median colour difference

under five different illuminants is calculated. The parameters *(m, K)* minimizing the error are reported in the results.

3. Results

The performance for all four methods is shown in Table 2. Below the median performance, the maximum errors are shown for each of the tested illumination conditions. The best performance, i.e. the smallest median colour difference between the predicted and the measured spectra, is achieved by the new direct method (DIRECT SS/DB) which uses the skin database to derive the PCA basis functions and the Silicon Skin Chart for calibration. Median colour differences range from 2.3 (Illuminant F02) to 3.1 ΔE^*_{ab} units (Illuminant A), with an overall mean of 2.7 ΔE^*_{ab} units. A statistical analysis using ANOVA and a post-hoc multiple comparison tests shows the performance between each pair of methods to be statistically significant ($p<0.001$).

Table 2. Performance of the four reconstruction methods across different illuminations

METHOD	ΔE^*_{ab}	A	D50	F02	F11	D65	Mean
Two-step SG/DB	Median (Max)	4.6 (11.3	4.3 10.4	4.2 10.1	4.7 10.7	4.2 10.0	4.4 10.0)
Two-step SS/DB	Median (Max)	3.7 (8.6	3.6 8.7	3.7 8.6	4.2 8.7	3.5 8.7	3.7 8.7)
Direct SS/SS	Median (Max)	3.4 (8.3	3.2 8.4	3.1 8.3	3.7 8.3	3.1 8.4	3.3 8.3)
Direct SS/DB	Median (Max)	3.1 (8.4	2.6 8.6	2.3 8.4	3.0 8.6	2.6 8.6	2.7 8.5)

3.1. Reconstruction performance in terms of perceptual error

All three modifications lead to a commensurable reduction in the spectral reconstruction error. (1) Mapping directly from camera RGB to the skin spectral reflectances yields smaller errors than mapping from camera RGB to XXZ, and then from XYZ to spectral reflectances (2.7 and 3.3 ΔE^*_{ab} compared to 4.4 and 3.7 ΔE^*_{ab}), across all five illumination conditions. (2) Using the newly developed skin-specific Silicon chart (Spectromatch Ltd) reduces the mean error from 4.4 to 3.7 ΔE^*_{ab}); this improvement is systematic across all illumination conditions. (3) Using the spectral skin database (DB) instead of the Silicon skin chart to derive the PCA basis functions leads to an error reduction from 3.3 to 2.7ΔE^*_{ab}.

For each method, performance for the optimal polynomial order and number of basis functions is reported in Table 2. Best-fitting parameters (polynomial order / number of basis functions) for each of the four methods are as follows: Two-step SG/DB – 4th order, 3 bases; Two-Step SS/DB – 2nd order, 3 bases; Direct SS/DB - 4th order, 5 bases; Direct SS/SS – 5th order; 3 bases.

3.2. Spectral measures of reconstruction performance

To evaluate the spectral reconstruction performance, in section 3.1. we have used perceptual error (ΔE^*_{ab}) which incorporpates properties of the visual system. For many applications, conversion of the spectra into a uniform colour space such as CIELAB might not be practical and sufficient and a measure based on the spectral difference is required.

Two spectral shape differences can be used for the evaluation of the spectral differences between the original and the reconstructed spectra: the root mean square error (RMSE) and the spectral similarity value (SSV) [23]. Spectral shape differences are evaluated RMSE and SSV in the full spectral range (between 400nm and 700nm at 10nm interval) in terms of average (Ave), maximum (Max) and median (Med) of the differences. The Appendix shows each of the 34 spectra generated using different method together with the original. Each

individual RMSE error and perceptual error (ΔE^*_{ab}) are also shown in the Appendix. Here, statistical results are given in Tables 3 and 4 respectively. Using the RMSE, the methods that perform best to worst are: DIRECT SS/DB, DIRECT SS/SS, TWO-STEP SS/DB, TWO-STEP SG/DB (Table 3), consistent with the perceptual error ranking. When using the SSV (Table 4), however, the best-to-worst performing methods are: TWO-STEP SG/DB, DIRECT SS/DB, DIRECT SS/SS, and TWO-STEP SS/DB, which is not consistent with the performance ranking based on the perceptual error. The results show the RMSE and SSV measures are not consistent with each other, but RMSE is consistent with the perceptual error ranking.

Table 3. Spectral shape differences in terms of RMSE for each method in the full spectral range between 400 nm and 700 nm at 10 nm intervals)

RMSE (400-700 *nm*)	Ave	Max	Med
TWO-STEP SG/DB	0.0449	0.1178	0.0420
TWO-STEP SS/DB	0.0341	0.0934	0.0262
DIRECT SS/SS	0.0341	0.0902	0.0269
DIRECT SS/DB	0.0329	0.0909	0.0254

Table 4. Spectral shape differences in terms of SSV for each method in the full spectral range (between 400 nm and 700 nm at 10 nm intervals)

SSV (400-700 *nm*)	Ave	Max	Med
TWO-STEP SG/DB	0.0773	0.1482	0.0718
TWO-STEP SS/DB	0.0885	0.1834	0.0897
DIRECT SS/SS	0.1042	0.1628	0.1013
DIRECT SS/DB	0.0831	0.1402	0.0819

For some biomedical applications reconstruction of the spectra is most important in a limited medium wavelength range (500-600nm) where the visual system exhibits its highest sensitivity and the spectral skin properities are diagnostic for the health of the individual (see discussion for further details). Skin spectra often exhibit a 'W' feature (see Fig. 5) between 520 and 580nm reflecting the oxygen saturation of haemoglobin in the blood flow [24].We therefore further evaluated the spectral differences in the W-shape range (between 500nm and 600nm at 10nm interval) and the results are given in Tables 5 and 6. In this medium wavelength range the rankings based on the two measures are in better agreement and the lowest spectral error is consistent with the perceptual error.

Table 5. Spectral shape differences in terms of RMSE for each method in the W shape range (between 500 nm and 600 nm at 10 nm intervals)

RMSE (500-600 nm)	Ave	Max	Med
TWO-STEP SG/DB	0.0322	0.0840	0.0275
TWO-STEP SS/DB	0.0275	0.0904	0.0208
DIRECT SS/SS	0.0262	0.0865	0.0194
DIRECT SS/DB	0.0256	0.0891	0.0190

Table 6. Spectral shape differences in terms of SSV for each method in the W shape range (between 500 nm and 600 nm at 10 nm intervals)

SSV (500-600 nm)	Ave	Max	Med
TWO-STEP SG/DB	0.1373	0.3406	0.1177
TWO-STEP SS/DB	0.1922	0.4392	0.1893
DIRECT SS/SS	0.1995	0.3149	0.1961
DIRECT SS/DB	0.1218	0.3006	0.1131

3.3. Correlation between spectral and perceptual error measures

The previous analysis showed that, for skin images, the RMSE is a more robust and useful measure to evaluate spectral reconstruction performance than the SSV is. In the following we therefore compare the perceptual errors with the RMSE and evaluate their strength of association and whether the perceptual error can be predicted from the RMSE.

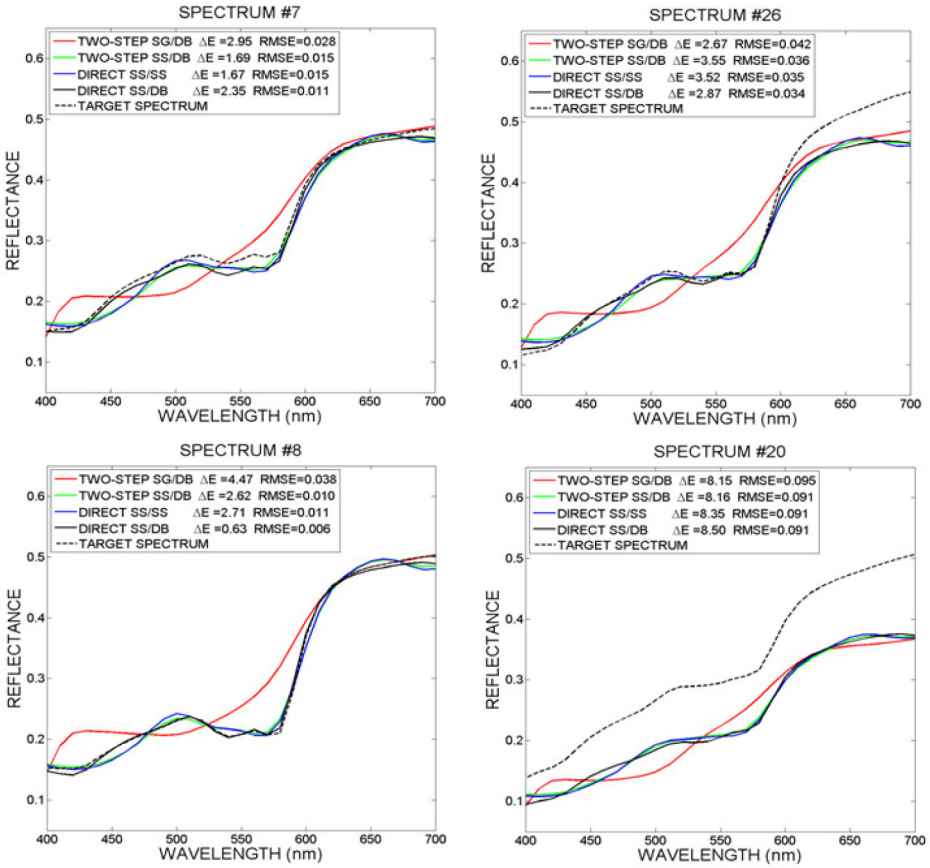

Fig. 5. Representative sample of reconstructed spectra. Upper row: two spectra (#7, #26) with errors close to the mean error of 2.7 (cf Table 1) for direct reconstruction (DIRECT SS/DB) as indicated by a solid line (spectrum #7 with a perceptual error of 2.35 and spectrum #26 with a perceptual error of 2.87). Lower left panel: spectrum #8 with the lowest perceptual error of 0.63 for the best performing method 'DIRECT SS/DB'. Lower right panel: worst reconstruction; spectrum #20 with a perceptual error of 8.5. In all cases the 'W' feature at around 560nm is recovered if present in the original spectrum.

To visualize the performane of all four methods, the reconstructed and the target spectra are plotted for three different scenarios: representative spectra (#7, #26) with a perceptual error close to the mean error (for the best method 'DIRECT SS/DB'; cf Table 1; Fig. 5), the spectrum with the best performance (#8; error = 0.63), and the spectrum with the worst performance #20; error = 8.5). The general spectral shape is captured by all methods apart from method #1 (TWO-STEP SG/DB) which consistently over/under-estimates the reflectance throughout the visible spectrum. Significant differences between the remaining three methods (cf Table 1) exist in this medium wavelength range (520-580nm) where the visual system exhibits its highest sensitivity. Crucially, the reflectance spectrum in this wavelength range is diagnostic of the health of individuals (see discussion for further details) and the accurate reflectance reconstruction in this wavelength range is of vital importance for heath-related applications. Skin spectra often exhibit a 'W' feature (see Fig. 5) between 520 and 580nm reflecting the oxygen saturation of haemoglobin in the blood flow [24]. Only the improved direct method is able to reconstruct this important feature of human skin spectra (please see the appendix for plots and associated errors of all 34 reconstructed spectra).

Figure 6 plots the perceptual error against the root-mean-square error (RMSE); the RMSE is based on the spectral reflectances and does not make any assumptions about the visual

system. Across all four methods the variance explained is 65%; all correlations are significantly different from zero ($p<0.001$).

The correlation is strongest for the best-performing reconstruction method (DIRECT SS/DB); in this case the RMSE explains 75% of the variance in the perceptual errors, compared to 47% for the worst-performing reconstruction algorithm (TWO-STEP SG/DB). This strong association for the new reconstruction method therefore allows a farily accurate prediction of the perceptual error from a spectral-based error measure.

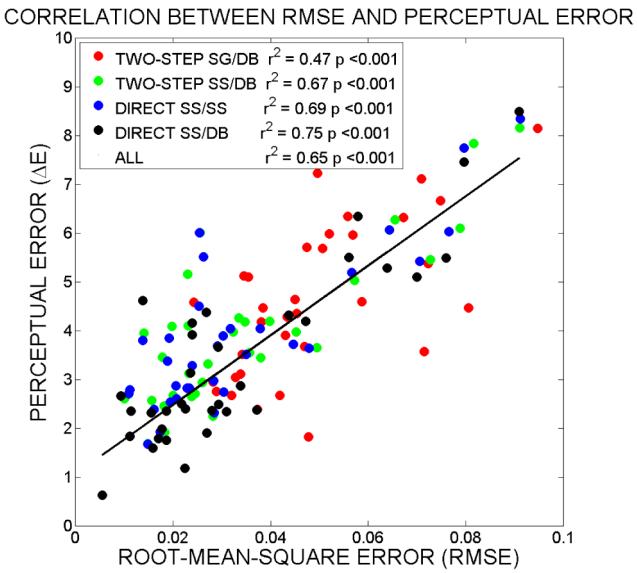

Fig. 6. Correlation between the RMSE based on spectral information and the perceptual error reflecting the visible differences between the original and the reconstructed spectra. The highest correlation between these two measures is obtained for the best-performing method (DIRECT SS/DB).

4. Discussion

Our aim was to improve on existing methods of reconstructing spectral reflectance functions for skin from RGB camera images. The best spectral reconstruction performance, i.e. the smallest mean colour difference between the predicted and the measured spectra, is achieved by the new direct method (mapping RGB to spectra) using the skin database to derive the PCA basis functions and a new Silicon Skin Chart for calibration. The mean resulting error (averaged across all illumination conditions) for this method is $2.7\Delta E^*_{ab}$, a cumulative error reduction of 1. 7 ΔE^*_{ab} in comparison to the conventional method.

Each of the three modifications leads to an incremental increase in performance and will be discussed in turn.

(1) In the two-step method, using the Silicon skin chart for the calibration step (step #1) reduces the average error by $0.7\Delta E^*_{ab}$ (cf. Table 2, rows 1 and 2). The choice of training colours used to derive the camera characterization model clearly affects the accuracy of the spectral reconstruction method; this is likely due to the fact that the gamut covered by the Silicon skin colour chart (cf. Figure 1, columns 2 and 3) is similar to the gamut of human skin whereas the X-rite ColorChecker includes 14 skin colours but is not skin-specific. Since the camera characterization model is optimized for the training colours, it will make more accurate predictions if training and test colours cover a similar gamut. In the evaluation stage, human skin images are used and the error due to the camera characterization will increase the overall reconstruction error.

(2) The most significant performance improvement is achieved by mapping the camera RGB values directly into the skin spectral reflectance functions, and thereby bypassing the mapping from RGB into the XYZ tristimulus values. Using the direct method (Table 1, row 4) compared to the 2-step method (row 2) leads to an error reduction of about 1 ΔE^*_{ab}. Since the same calibration chart (Silicon) and database are used, the improvement must be due to the methods. In the direct method, only a single optimization is performed, namely optimizing the coefficients of the polynomial regression (cf. Figure 3). In contrast, both 2-step methods first estimate the XYZ values and subsequently the spectra are reconstructed from the estimated XYZ hence introducing two sources of errors. The reflectance reconstruction based on the true XYZ has already a reconstruction error; the reconstruction based on the estimated XYZ adds an additional error. Furthermore, the reflectance r based on the estimated XYZ is aiming for the reconstructed reflectance having the same XYZ as the estimated XYZ, hence it cannot be expected the reconstructed reflectance is the same as the original or has the same characteristics as the original. Both of these shortcomings are addressed in the direct method which minimizes the difference between the estimated and the original reflectance.

(3) In the direct method, when the silicon skin colours are used for the principal component analysis (Table 1, row 3) instead of the skin database (row 4), the error increases on average by 0.6 ΔE^*_{ab}. Although the Silicon skin colour chart is optimized to mimic the colour gamut of natural skin (Fig. 1), there are significant differences in their respective spectral reflectance functions (compare Fig. 2 with Fig. 4) which results in slightly different basis functions. Since the reflectance recovery methods are tested with real skin images, a larger error occurs when the Silicon spectra are used for the derivation of the principal components. In addition to minimizing the perceptual error, the use of the human skin database instead of the silicon spectra also results in reconstructed spectra with the main characteristics of human skin spectra which is important for biomedical applications (see last section of the discussion for details)

(4) In addition to the spectral differences between the Silicon colour charts and the human skin database, the number of spectra used for the PCA also differ significantly: the silicon skin chart consists of 90 samples, whereas the skin database contains 4392 spectra. We have tested the effect of the sample size by using a randomly selected subset (n = 90) of the human skin spectra. We find that the errors are virtually identical to the errors obtained for n = 4392 (Illuminant A: 3.1 (8.5); 2.6 (8.6); D50: 2.4 (8.4); F02: 3.0 (8.6); F11: 2.6 (8.6); D65: 2.7 (8.5)). We therefore conclude that it is not the number of samples, but the spectral similarity between the database and the test spectra that accounts for the better performance.

(5) Neither for the new nor for the conventional method, is there a limitation on the requirement of the illumination (see reference 9 as well) for the imaging.

Advantages of the new skin-specific reconstruction algorithm for biomedical applications

Precise reconstruction of skin spectra from camera RGB inmages is important since spectral reflectance functions contain information not available in colorimetric skin values. Spectral reflectance data contain the necessary information to model skin appearance under different illumination conditions and also allow the extraction of information about skin chromophores, such as oxygen saturation of haemoglobin. Figure 7 (left panel) shows the absorbance of haemoglobin (http://omlc.org/spectra/hemoglobin), either in the oxygenated (thick red line) or the de-oxygenated state (thick green line); the locations of the peak sensitivities of the short- (442nm), medium- (543nm), and long-wavelength (570nm) sensitive cones are indicated by the thin vertical lines, respectively. The right panel shows the inverse of the haemoglobin absorbance and demonstrates that the 'W' feature in the skin reflectance spectra (e.g. Figure

5, upper left panel) is a reflection of the oxygenation saturation of haemoglobin. Recovering the 'W' feature of the skin spectra is important for two reasons. Firstly, the visual system is very sensitive to colour variations in the wavelength region containing the 'W' feature (520 to 580nm) due to the location of the peak cone sensitivities. When reconstructed skin spectra are used to model skin appearance under different illumination conditions, imperfectly reconstructed spectra in this region can lead to significant appearance changes. Secondly, variations of hemoglobin (Hb) oxygenation in skin tissue are important indicators for its physiological conditions and important for clinical care. Precise estimates of Hb oxygenation requires a concise reconstruction of this 'W' feature, which is achieved by incorporating the use a human skin database in our new reconstruction method (cf. Figure 3).

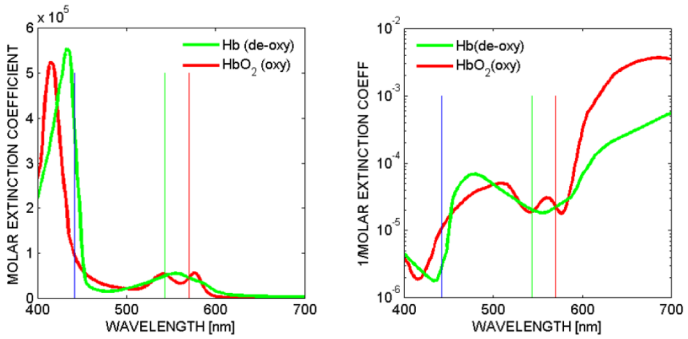

Fig. 7. On the left the absorbance as a function of wavelength is shown for the oxygenated (red line) and the de-oxygenated haemoglobin (green line). On the right panel, the inverse of the absorbance is plotted to show that the 'W' feature in the skin reflectance spectra (e.g. Figure 5, upper left panel) is a reflection of the oxygenation saturation of haemoglobin. Thin vertical lines indicate the location of the peak sensitivities of the cones in the human retina.

5. Conclusion

In conclusion, the three modifications together reduce the perceptual error to less than 3 ΔE^*_{ab} units across all five different illuminations, which is within the acceptable range for skin reproduction for maxillofacial soft tissue prostheses [25]. We conclude that our direct method using the spectral skin database to derive the basis functions is likely to be of significant value for biomedical applications, as well as for digitally rending skin under varying illumination conditions.We also show that perceptual errors associated with the spectral reconstruction algorithm can be predicted from a spectral-based RMS error and thereby increasing its practical value for health-related applications.

Acknowledgements

This work was supported by the Royal Academy of Engineering (Grant NO. RECI033), the EPSRC (EP/K040057, EP/L001012) and the National Natural Science Foundation of China (No. 61178053,61575090). We would also like to thank Spectromatch Ltd (London) for kindly providing the silicon skin colour samples. The spectral skin database will be made publicly available on pcwww.liv.ac.uk/~sophiew/skin, together with the source code.

Appendix: Reconstructed skin spectra (n=34) with associated reconstruction error (RMSE and Perceptual Error)

In this appendix, all the 34 test skin spectra and their reconsctructed spectra using the four methods: TWO-STEP SG/DB, TWO-STEP SS/DB, DIRECT SS/SS, DIRECT SS/DB are shown in Figs. 8-10 respectively. Furthermore, the RMSE and perceptual errors for each reconstrcuted spectra are also shown in the diagrams.

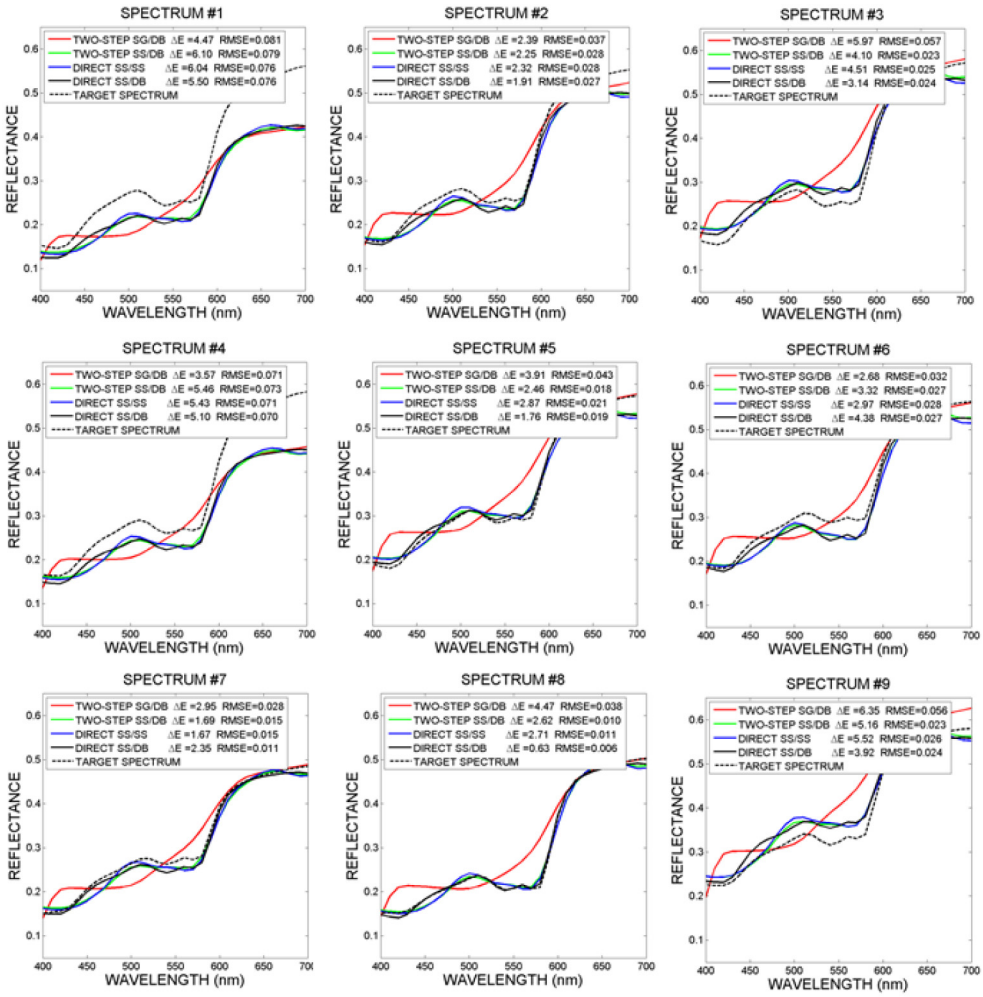

Fig. 8. Reconstructed skin spectra #1-9 with associated reconstruction error (RMSE and Perceptual Error) using each of the four methods: TWO-STEP SG/DB, TWO-STEP SS/DB, DIRECT SS/SS, DIRECT SS/DB together with the original spectra

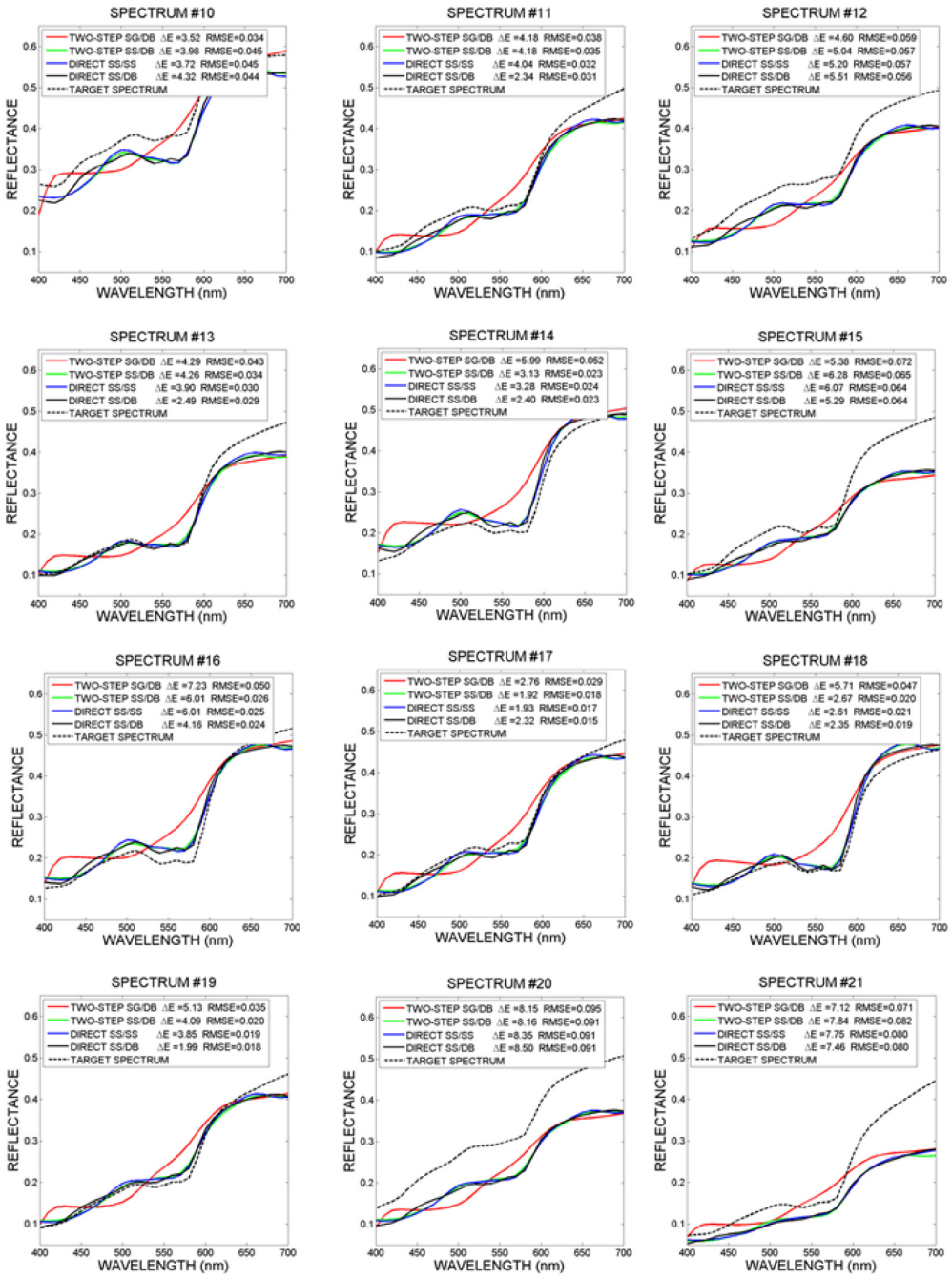

Fig. 9. Reconstructed skin spectra #10-21 with associated reconstruction error (RMSE and Perceptual Error) using each of the four methods: TWO-STEP SG/DB, TWO-STEP SS/DB, DIRECT SS/SS, DIRECT SS/DB together with the original spectra

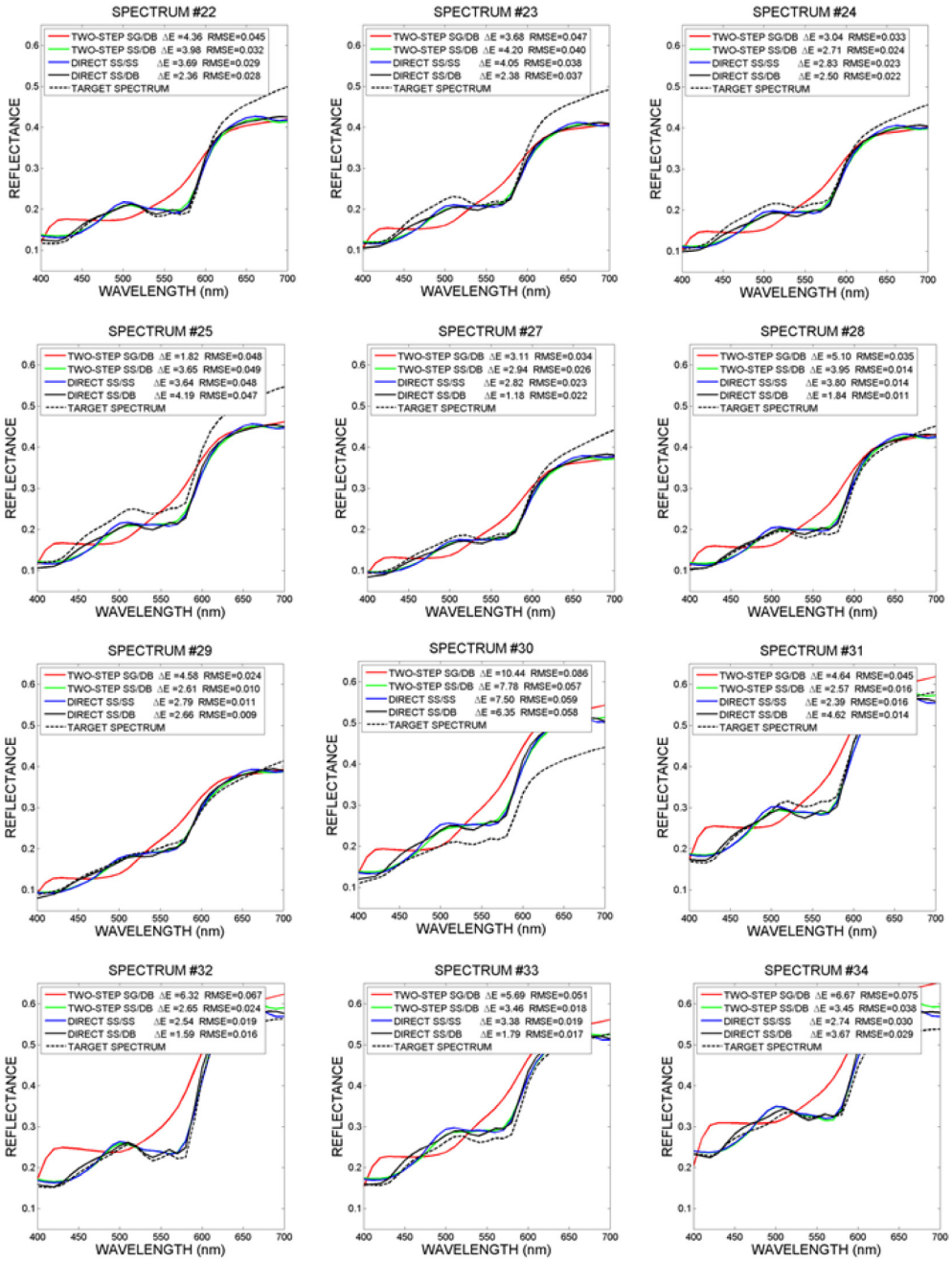

Fig. 10. Reconstructed skin spectra #22-34 with associated reconstruction error (RMSE and Perceptual Error) using each of the four methods: TWO-STEP SG/DB, TWO-STEP SS/DB, DIRECT SS/SS, DIRECT SS/DB together with the original spectra