Modelling Skin Colours for Preferred Colour Reproduction

Huanzhao Zeng and Ronnier Luo**, *Hewlett-Packard Company, Vancouver, WA, USA **University of Leeds, Leeds, UK*

Abstract

The rendering of skin colours plays a significant role for preferred colour reproduction. To study preferred skin colour enhancement, we evaluated various skin colour boundary models for skin detection, face tracking, and colour enhancement, and selected an elliptical modelling approach to model the skin colour boundary for its accuracy, efficiency, and convenience for skin colour enhancement. An image database was developed to train ellipse models and ellipsoid models. The skin boundary model was applied to guide psychophysical experiment to determine a preference skin colour centre for preferred colour enhancement.

1. Introduction

Colour rendering is an important factor for judging perceived image quality of the colour reproduction of digital images. Skin tone, as the most important category among memory colour categories, plays an important role for preferred colour reproduction [1-2]. To enhance skin colours, a skin colour region must be detected. A simple method to determine a skin colour region is to explicitly define the range in a specific colour space [3-7]. Although it has the advantage of simple decision rules that lead to constructing a very rapid classifier, it is difficulty to find adequate decision rules empirically. A different approach, non-parametric skin colour modelling, is to estimate skin colour distribution from a training data set without deriving an explicit skin colour model [8]. A colour space may be quantized to a number of bins, forming a 2-D or 3-D histogram representing with a colour distribution lookup table (LUT) or probability map that may be directly used for skin colour detection and face tracking [9-12]. The advantages of the approach are fast and easy in training, and theoretically independent of the shape of the skin distribution. The disadvantage is that a large storage space is required. Assuming skin colours are distributed around a colour centre, the skin colour distribution may be described with a Gaussian-like function. This led to the skin colour modelling using Single Gaussian Model (SGM) [13-17]. SGM assumes a unimodal distribution which may cause intolerable error in estimation and discrimination. An improved approximation can be obtained when the values are generated by one of the several randomly occurring independent sources. In this case the distribution function is a multimodal one which can be estimated using a finite number of mixed Gaussians or a Gaussian mixture model (GMM). Although the performance of the SGM is similar to GMM for low false positive detection rates, it is significantly decreased for high true positive detection rates [18].

From the human perception point of view, the shape of each equal probability distribution locus of the skin colour boundary should be smooth. Yet, this may be a difficult task for GMM. An elliptical shape should well approximate each equal probability distribution boundary. By examining skin and non-skin distributions in six colour spaces, Lee and Yoo [19] concluded that skin colour cluster can be well modelled using an elliptic shape. This is not surprised from colour and vision points of view, since the human perceptual colour tolerance in a perceptually uniform colour space can be well modelled with ellipses [20-22].

The colours of an object category distribute around its prototypical colour with a probability density function [23]. Due to various physical factors (illuminations, camera characteristics, image editing, etc.), the skin colour distribution deviates from the Gaussian distribution. However, the shape of the equal-distribution contour should be approximately elliptical. Observing that the boundary of skin colours is of nearly elliptic shape in uniform colour spaces, we adapted an elliptical boundary model, modified and expanded it to compensate the lightness dependency. The model was applied to guide psychophysical experiments to determine a preferred skin colour centre for skin colour enhancement.

2. Elliptical Boundary Model

The cluster of skin colours in a 2-D chromatic space such as CIELAB a* vs b* planes may be approximated using an elliptical shape. Let X_1, …, X_n be distinctive chromatic vectors of a pre-processed training data set and $f(X_i)=f_i$ (i=1, …, n) be the number of samples (or occurrences) of a colour with chromatic coordinate of X_i. An elliptical boundary model (X)=(X; ,) is defined as

$$\Phi(X) = [X - \Psi]^T \Lambda^{-1} [X - \Psi] \tag{1}$$

where two model parameters, and are given by

$$\Psi = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{2}$$

$$\Lambda = \frac{1}{N} \sum_{i=1}^{n} f_i (X_i - \mu)(X_i - \mu)^T \tag{3}$$

where

$$N = \sum_{i=1}^{n} f_i$$

is the total number of samples in the pre-processed training data set, and the vector

$$\mu = \frac{1}{N} \sum_{i=1}^{n} f_i X_i$$

is the mean of chromatic vectors.

Given a threshold and an input chromatic vector X of a pixel, X is classified as skin chromatic coordinates if (X) < and as non-skin chromatic coordinates otherwise. The threshold trades off correct detections by false detections. As increases, the number of correct detections increases but the number of false detections increases as well.

(X) = defines an elliptical boundary between skin and non-skin colours. The centre of the ellipse is given by and the principal axes are determined by .

In a 2-D chromatic plane, $^{-1}$ is represented in a matrix form

$$\Lambda^{-1} = \begin{pmatrix} \lambda_{00} & \lambda_{01} \\ \lambda_{10} & \lambda_{11} \end{pmatrix} \quad (4)$$

(X) may be reorganized in the following form

$$\Phi(X) = \lambda_{00}(x-x_0)^2 + (\lambda_{01}+\lambda_{10})(x-x_0)(y-y_0) + \lambda_{11}(y-y_0)^2 \quad (5)$$

By translating the origin to (x0, y0) and rotating an angle, , to rotate the x-y coordinates to the principal axes, can be derived by

$$\theta = 0.5 \arctan\left(\frac{2\lambda_{01}}{-\lambda_{00}+\lambda_{11}}\right),$$

and two parameters related to the principal axes are

$$a = \lambda_{00}\cos^2(\theta) - \lambda_{01}\cdot\sin(2\theta) + \lambda_{11}\sin^2(\theta)$$

$$b = \lambda_{00}\sin^2(\theta) + \lambda_{01}\sin(2\theta) + \lambda_{11}\cos^2(\theta)$$

The lengths of the semi-major and semi-minor axes are $\sqrt{\frac{1}{a}}$ and $\sqrt{\frac{1}{b}}$.

The relationship between $^{-1}$ and a, b, and is:

$$\begin{pmatrix} \lambda_{00} & \lambda_{01} \\ \lambda_{10} & \lambda_{11} \end{pmatrix} = \begin{pmatrix} a\cdot\cos^2(\theta)+b\cdot\sin^2(\theta) & \frac{b-a}{2}\sin(2\theta) \\ \frac{b-a}{2}\sin(2\theta) & a\cdot\sin^2(\theta)+b\cdot\cos^2(\theta) \end{pmatrix}$$

To model skin colours in a 3-D colour space, becomes a 3x3 matrix. Eq. (5) becomes

$$\Phi(x,y,z) = u_0(x-x_0)^2 + u_1(x-x_0)(y-y_0) + u_2(y-y_0)^2 + u_3(x-x_0)(z-z_0) + u_4(y-y_0)(z-z_0) + u_5(z-z_0)^2 \quad (6)$$

where $u_0 = {}_{00}$, $u_1 = {}_{01} + {}_{10}$, $u_2 = {}_{11}$, $u_3 = {}_{02} + {}_{20}$, $u_4 = {}_{12} + {}_{21}$, and $u_5 = {}_{22}$.

A database to derive the coefficients of the elliptical model is described in the following section.

3. Image Database

The image database to construct skin colour boundary composes about 2500 photos. About 60% of photos were captured between 2001 and 2008 using professional digital cameras in indoor studios where the lighting conditions were well controlled, and therefore each image had proper white balance. Most of other images were captured outdoor in recent years using various professional digital cameras.

A tool was developed to mask a portion of skin colours for each image. Fig. 1 shows a snapshot of the tool. The left window shows the original image. Once the mouse is pointed to a skin colour and clicked, the skin colour of this point is applied as the seed colour (colour centre) to grow the colour region. The span of the region is determined by a range slider that sets a colour difference threshold of each pixel to the seed colour. The pixels marked as skin colours are presented as a magenta colour displayed on the centre window. The mask values of the image are drawn as a gray-scale image shown on the right window.

Although an RGB colour space can be used to grow a region, we found that using L*a*b* colour space gave us results more correlated to human perception. Thus, CIE L*a*b* colour space was used in this study. The RGB colour values are converted to L*a*b* using the embedded ICC profile (or sRGB ICC profile if no embedded ICC profile exists) of the image.

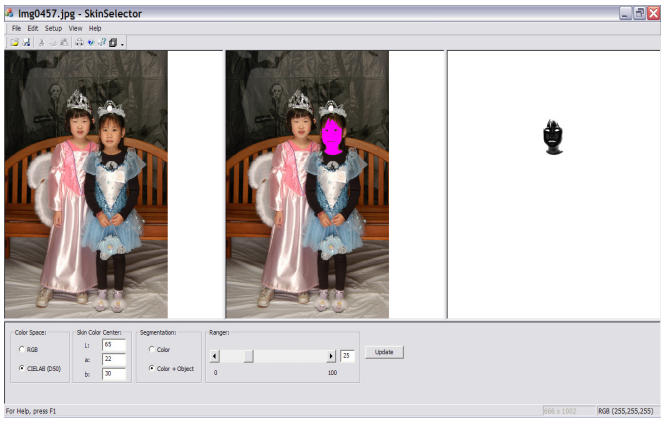

Fig. 1 *A tool to create a skin colour database*

If the segmentation method, "Colour + Object" in the tool, is selected, the regional growth is subjected to the constraint that skin colours must be connected together as a single region, which prohibits more than one isolated region. This method is used for constructing the skin database for: 1) it prohibits growing similar colours to other objects; and 2) it provides an approach for using different colour centres to mark different persons on one image.

After an object is marked, the source image (on the left window), the image with a marked skin object (on the centre window), the image of the skin mask (on the right window), and the setting parameters are saved. An image with very high resolution is re-sampled to an ideal size to avoid that the weighting for a very high resolution image is higher than that for a lower resolution image.

The marked pixels are used to analyze skin colours, and all other pixels that are not marked (white pixels in the image on the right window) are ignored. Therefore, marking all skin pixels is not necessary.

The skin colours of all images were selected mostly by a single person. Because a following step was implemented to remove lower occurrence pixels during counting pixel occurrences for each image (10% were removed in this study), and a huge amount of skin pixels (in the order of billions) were selected from diversified images, the bias from users selection should be insignificant.

After masking all images, a script reads each image (the image on the left window) and its associated mask image (image on the right window) and adds occurrences of skin colours to a 256x256x256 RGB lookup table (LUT). The reason for using 256^3 LUT is the convenience for counting occurrence of 8-bit RGB images. The number on each node of the LUT represents the occurrences of the RGB colour as a skin pixel. So the number on every non-skin node is zero. For every 8-bit/channel RGB image, each skin colour adds a count to the corresponding bin of the LUT. To remove noisy pixels and pixels that are inaccurately marked as skin pixels, a small percentage of pixels with low occurrences are excluded from counting occurrences at the time each image is processed. In this study, 10% least occurrence pixels were removed from each image.

4. Results of the Skin Colour Modelling

Fig. 2 shows the skin colour cluster of the selected skin colours of the image database under CIELAB colour space. Snapshots in two different viewing angles are shown. The shapes of constant-lightness slices are close to ellipses, but the sizes and the locations are different for different lightness levels. If an ellipse is used to fit the skin boundary on the a*-b* plane, it should be large enough to cover the dominant mid-tone skin colours, although smaller ellipses fit well for lighter and darker tones.

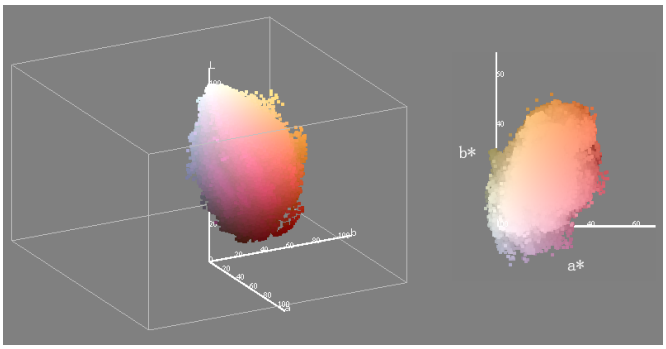

Fig. 2 *3-D gamut of all skin pixels selected from the image database*

As described in the previous section, a skin object was marked for each image. Although each skin colour has a mask value between 0 and 255 to indicate the likelihood to be a skin colour, the skin mask was used as a binary mask for the modelling, i.e., each colour was treated as a skin colour or a non-skin colour, and its likelihood value (colour difference to the central skin colour) was not applied for the modelling.

The RGB colour at each node of the RGB histogram LUT is converted to CIE L*a*b* colour space and the white point is adapted to D50. The LUT is used to train the elliptical model, where the count in each bin of the LUT is the occurrence, f(X), in Eq. (3), and the a*b* value of each bin location is the colour, X.

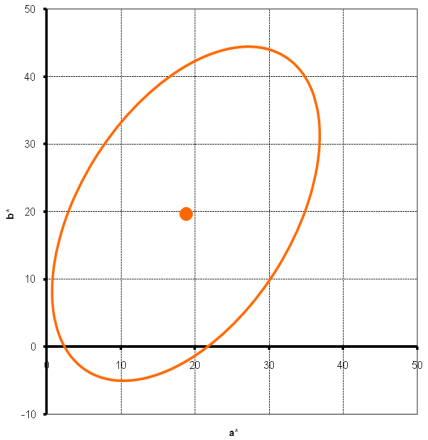

Fig. 3 *Skin colour boundary in a*-b* coordinate*

Fig. 3 shows the modelled ellipse in a*-b* coordinates to cover 95% of skin colours (equivalent to about 99% coverage rate with occurrence weighting). The centre coordinates are (19, 20), together with the ellipse parameters [a, a/b, θ] of [27, 1.8, -62°] (negative-degree means counter-clockwise). For practical applications, the principle axes may be decreased or increased proportionally (equivalent to adjusting of the elliptical model).

To study the dependency of ellipse on L*, the skin pixels in the database was divided into different bucket, each contains data of a certain L* range. We divided L* into ten bucket, each occupying an L* of 10 units. Ellipses in a*-b* coordinates fitted for each bucket are shown in Fig. 4. Each ellipse is fitted to cover 90% of data points within the bucket. There is no ellipse at the bucket of L* within [0, 10]. The upper left one is for the L* bucket of [10, 20] and the last one is for L* bucket of [90, 100]. L* increases orderly from left to right and top to bottom.

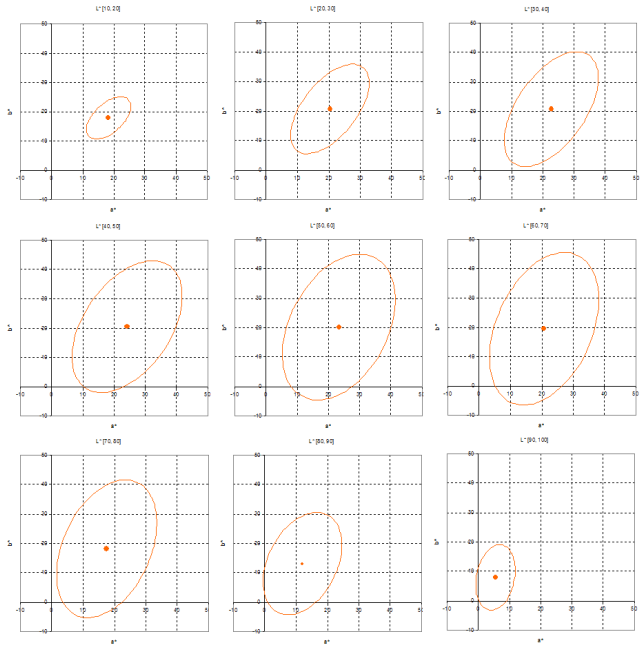

Fig. 4 *Skin colour ellipses in different constant-lightness buckets*

To model the L*-dependency of the skin ellipses mathematically, the training data set was divided into twenty buckets for L* from 0 to 100, with an L* interval of 5 for each bucket. The skin colour centres, the sizes of ellipse (sizes of principal axes), and the orientations of ellipses varying according to different lightness levels were model. Figures 5-6 show examples of the chroma and hue angle of the ellipse centre vary as functions of lightness (a smooth fitting curve was plotted in each figure). The ellipse equation for each lightness level was developed in the same manner.

Fig. 6 shows that hue angles in dark and light skin colours do not follow the global trend. It was suspected that skin colours with L* close to white was influenced significantly by lighting conditions and white-balances of cameras and dark skin colours are not as reliable; therefore we did not attempt to fit the abrupt changes in both ends.

Using a 3-D ellipsoid to fit the skin colour boundary simplifies the modelling. Fig. 7 shows an ellipsoid to fit the skin colour cluster in CIE L*a*b* colour space. The ellipsoid centre is (59, 19, 20), the principal axis parameters [a, a/b, a/c] are [38, 1.4, 2.5], and the unit vectors of three principal axes relative to the

centre are (0.97, -0.14, -0.19), (0.24, 0.44, 0.87), and (0.04, 0.89, -0.46).

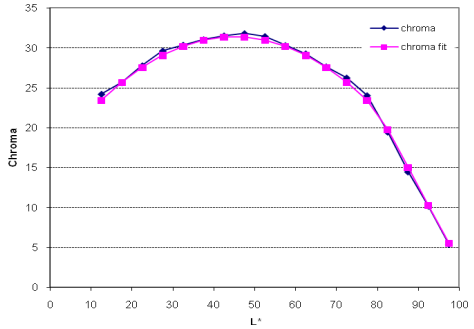

Fig. 5 *Chroma of skin colour centres*

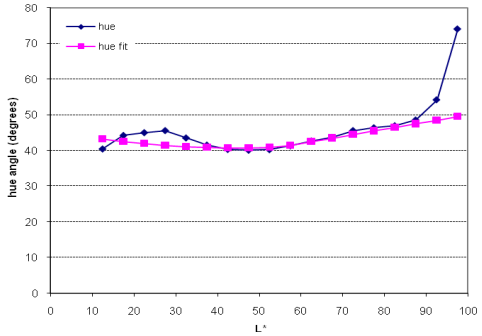

Fig. 6 *Hue angle of skin colour centres*

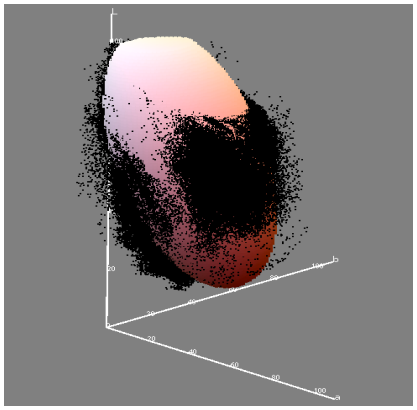

Fig. 7 *An ellipsoid to cover 90% of the skin colours*

We use all combinations of 8-bit sRGB colours (256x256x256 colours) to test the detection rate and assume that each colour occurs equally frequently. The skin colour histogram generated from the image database is applied to the 256^3 colours. With the above 2-D ellipse to model skin colours ignoring lightness dependency, the true positive detection rate (TP) is 90% and the false positive detection rate (FP) is 1.1%. With a 3-D ellipsoid the size to yield 90% true positive rate, the false positive detection rate is 1.1% as well. The result shows that the ellipsoid model and the ellipse model perform about the same at TP = 90%.

The size of the ellipse or ellipsoid may be increased to improve TP by the sacrifice of higher FP. For example, increasing the size of the ellipsoid to yield 95% TP will increase FP to about 5%.

5. Apply the Skin Colour Model for Preferred Colour Reproduction

The skin colour boundary model developed above was applied to study preferred skin colour reproduction. We conducted an experiment to determine a skin colour centre for preferred skin colour enhancement. The evaluation was conducted in HP's Perceptual Science Lab in Vancouver, Washington. The room lighting was designed to simulate a typical dim surround condition. Five HP L2335 Active Matrix TFT 23-inch LCD displays were used to display images. The monitors have a native resolution of 1920 x 1200 pixels and a pixel pitch of 0.258mm. Each display's white point was adjusted to D65.

Observers were from a pool of HP employees at the Vancouver site for psychophysical experiments. All of them satisfactorily completed colour vision discrimination testing and observer orientation training. They were experienced at evaluating image quality because of involvement in the evaluations that were conducted regularly. They were not considered experts, but were believed to predict the general consumer response. Periodic external evaluations have validated the calibration of the HP internal testing with real customers. The observers were diverse in gender, and ages were between 25 and 50.

In the psychophysical experiment, the lightness-independent elliptical model was applied to simplify the experimental design. A set of nine skin colour centres was predetermined to process images. Fig. 8 shows the nine colour centres in CIE a*b* chromatic plane. All nine points are within the skin colour ellipse derived from our Halloween image database, and the centre point is the skin elliptical centre.

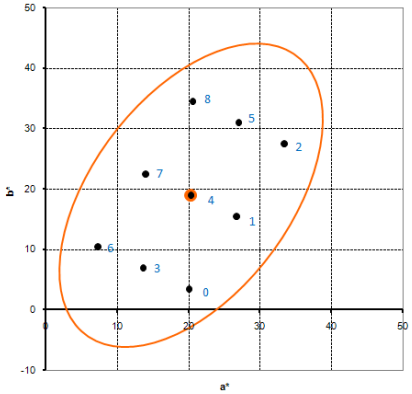

Fig. 8 *Nine predetermined skin colour centres*

Twelve sRGB images for the experiment were carefully selected to cover different skin types (see Fig. 9). The skin colours of each image were processed in a way that skin colours were morphed toward a skin colour centre (the adjustment gradually fades away as colours move off the skin colour boundary). First, each image was converted from the source RGB colour space to L*a*b* (adapted to D50) using the source RGB ICC profile. The 2-D elliptical skin model was applied to compute Mahalanobis distance, or (X). If is greater than 1.0, the colour was

considered non-skin colour, thus no colour adjustment was applied. Otherwise, the Mahalanobis distance was used as a weight to morph skin colours toward each of the nine skin centres. With nine predetermined skin centres, nine versions of images were created from each image for the experiment.

A software tool was used to display a pair of images at a time on the display for observer judgment (see Fig. 10). Each observer was presented with an image pair on display and was asked to indicate which rendition of the two was preferred for skin tones and to indicate the degree of preference using a scale from 0 to 10 where 0 is most preferred to the left image (most dislike to the right image) and 10 is the most preferred to the right image (most dislike to the left image), and 5 is neutral. After recording the response, the next image pair was loaded and the observer would proceed until all the samples were evaluated. Image pairs were presented to each observer in a different random order, and image placements (left/right) were also randomized. Four images were evaluated each day by nineteen observers. It took three days to complete the evaluation. Observers in different days are not all the same.

Fig. 9 *Images for the experiment*

Fig. 10 *A snapshot of a display screen*

The experimental result from all images is shown in Fig. 11. The scores on the vertical axis are the preference strengths. A negative number means dislike. A larger number means stronger preference. The preference trends are very similar across most of images. Overall, there is a strong preference for skin colour locations #4 and #7. The location #0 and #2 are most dislike and their scores are very close. The next dislike locations are the pair of #6 and #8 and the pair of #3 and 5.

We analyzed some experimental data to study inter-observer preferences. Since most observers did not judge all images, we could only select a few images to analyze observer variations. The first day's observation data for the judgments of the first four images by nineteen observers were analyzed. Fig. 12 shows the result. Each blue dot is the colour centre of an observer's judgement. The thick orange ellipse encompasses these colour centres. Each of other ellipses shows the preference region of an observer's judgement. The figure shows that although the preference hue angles are very close with each other, some observers preferred less chromatic skin colours than others.

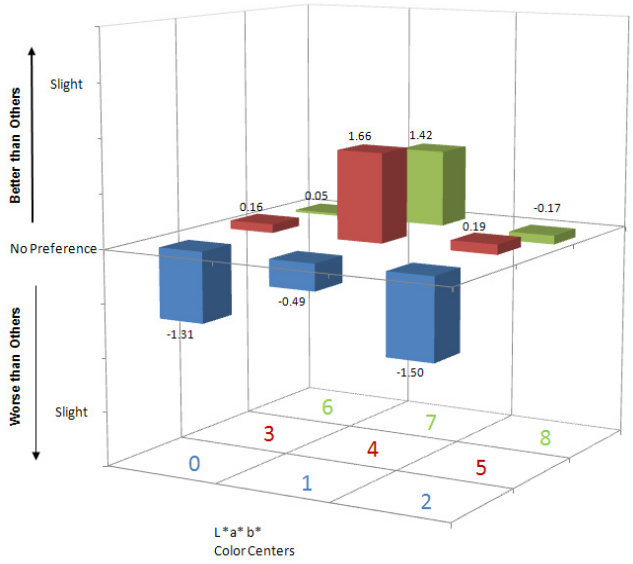

Fig. 11 *Overall preference scores*

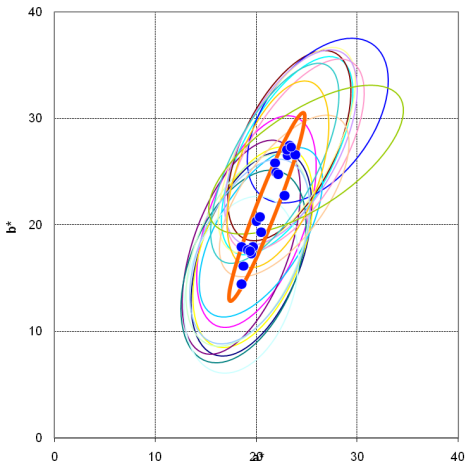

Fig. 12 *Observer judgments for the first four images*

Fig. 13 shows preferred colour regions of each of the 12 images. The pink ellipse encompasses skin colours. Orange ellipses, light gray ellipses, and black gray ellipses are the preferred colour regions derived from Caucasian face images, oriental face images, and face images with very dark skin colours (African face images and an Asian face image). The figure shows that the preferred skin colour region for reproducing dark skin colours seems to be slightly less chromatic than that for medium and light skin colours, and preferences for medium and light skin colours show no difference. In addition, the shapes and orientations of the ellipses and the skin colour boundary are very similar.

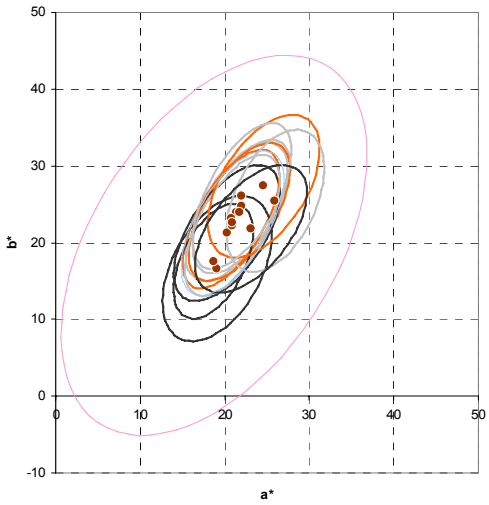

Fig. 13 *Preferred skin colour regions of all images*

6. Conclusions

An image database was built to study skin colours of digital images. It was used to model the elliptical boundary of skin colours. The statistical skin colour centre and elliptical boundary was found from billions of pixels that were cropped from more than two thousand images. Lightness-independent and lightness-dependent ellipses in a*-b* coordinates were modelled. An ellipsoid modelling in CIE L*a*b* colour space was developed to simplify the modelling for lightness dependency.

The skin colour modelling was applied to guide psychophysical experiments to study skin colour preference for preferred skin colour reproduction. The results reveal a strong preference at an 'ideal' skin colour centre. An approximate preferred skin colour centre was found in this experiment. Follow-up experiments were conducted to identity the preferred colour centre more accurately and to verify the result. The result will be reported separately.

References

1. Fernandez, S. and Fairchild, M (2001), "Preferred Colour Reproduction of Images with Unknown Colorimetry", IS&T/SID 9th Colour Imaging Conference, 274-279.
2. Braun, K.M. (2006) "Memory Color Enhancement Algorithm", International Congress of Imaging Science, 227-229.
3. Chai, D. and Ngan, K.N. (1999) "Face segmentation using skin-color map in videophone applications", IEEE Transactions on Circuits and Systems for Video Technology, 9(4): 551-564.
4. Gomez, G. and Morales, E.F. (2002) "Automatic Feature Construction and a Simple Rule Induction Algorithm for Skin Detection", Proc. ICML Workshop on Machine Learning in Computer Vision, 31-38.
5. Hsu, R.-L., Abdel-Mottaleb, M., and Jain, A.K. (2002) "Face detection in color images", IEEE Transactions on Pattern Analysis and Machine Intelligence, 24 (5): 696-706.
6. Kovac, J., Peer, P., and Solina, F. (2003), "2D versus 3D colour space face detection", IEEE 4th EURASIP Conference focused on Video/Image Processing and Multimedia Communications, Vol. 2, 449-454.
7. Mahmoud, T.M. (2008) "A New Fast Skin Color Detection Technique", Proc. World Academy of Science, Engineering and Technology, Vol. 33, 518-522.
8. Vezhnevets, V., Sazonov, V., and Andreeva, A. (2003) "A Survey on Pixel-Based Skin Color Detection Techniques", http://graphics.cs.msu.ru.
9. Chen, Q., Wu, H., and Yachida, M. (1995) "Face detection by fuzzy pattern matching", IEEE 5th International Conference on Computer Vision, 591-596.
10. Zarit, B.D., Super, B.J., and Quek, F.K.H. (1999) "Comparison of Five Color Models in Skin Pixel Classification", International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, 58-63.
11. Sigal, L., Sclaroff, L., and Athitsos, V. (2000) "Estimation and prediction of evolving color distributions for skin segmentation under varying illumination", IEEE Conference on Computer Vision and Pattern Recognition, Vol. 2, 152-159.
12. Brown, D.A., Craw, I., and Lewthwaite, J. (2000) "A SOM Based Approach to Skin Detection with Application in Real Time System", Proc. International Conference on Pattern Recognition, Vol. 1, 1056-1059.
13. Yang, M.-H. and Ahuja, N. (1998) "Detecting human faces in color images", IEEE International Conference on Image Processing, Vol. 1, 127-130.
14. Menser, B. and Wien, M. (2000) "Segmentation and tracking of facial regions in color image sequences", Proc. SPIE: Visual Communications and Image Processing, Vol. 4067, 731-740.
15. Hsu, R.-L., Abdel-Mottaleb, M., and Jain, A.K. (2002) "Face detection in color images", IEEE Transactions on Pattern Analysis and Machine Intelligence, 24 (5): 696-706.
16. Park, D.-S., Kwak, Y., Ok, H., and Kim, C.Y. (2006) "Preferred skin color reproduction on the display", J. Electronic Imaging, 15(4): 041203.
17. Almohair, H.K., Ramli, A.R., Elsadig, A. M., and Hashim, S.J. (2007) "Skin Detection in Luminance Images using Threshold Technique" International Journal of the Computer, the Internet and Management, Vol. 15(1): 25-32.
18. Caetano, T.S., Olabarriaga, S.D., and Barone, D.A.C. (2002) "Performance evaluation of single and multiple-Gaussian models for skin color modelling", IEEE Brazilian Symposium on Computer Graphics and Image Processing XV, 275-282.
19. Lee, J.Y. and Yoo, S.I. (2002) "An Elliptical Boundary Model for Skin Color Detection", Proceedings of the International Conference on Imaging Science, Systems, and Technology.
20. Macadam, D.L. (1942), "Visual Sensitivities to Color Differences in Daylight," J. Opt. Soc. Am. 32, 247-273.
21. Wyszecki, G., Fielder, G.H. (1971). "New Color-Matching Ellipses". J. Opt. Soc. Am. 61 (9): 1135–1152.
22. Luo, M.R., Rigg B., (1986) "Chromaticity-discrimination ellipses for surface colour", Color Res. Appl. 11(1): 25-42.
23. Yendrikhovskij, S.N., Blommaert, F.J.J., Ridder, H. (1999) "Colour Reproduction and the Naturalness Constraint", Colour Res. Appl. 24(1): 54-67.

Acknowledgement

I would like to thank David Berfanger and Timothy Mauer for their critical support for designing the psychophysical experiment and analyzing results.