Journal of Imaging Science and Technology® 55(3): 030201-1–030201-12, 2011.

Skin Color Modeling of Digital Photographic Images

Huanzhao Zeng▲
Hewlett-Packard Company, Vancouver, Washington 98683
E-mail: huan_zeng@hp.com

M. Ronnier Luo▲
University of Leeds, Leeds, United Kingdom

Abstract. *Three elliptical skin color models are presented for skin color detection. The first one is to model the skin color cluster using a single ellipse ignoring the lightness dependency. It is simple and efficient, and the skin color detection accuracy may be adequate for many applications. In the second model, the skin color ellipse is adapted to different lightness levels to better fit the shape of the skin color cluster. The model is more complex to train, and the computation efficiency is lower, but the skin color detection accuracy is considerably higher. In the third method, an ellipsoid is trained to fit the skin color cluster. It is almost as simple to train as the first model, but the skin color detection accuracy is higher. Having skin color detection accuracy almost as high as the second model, this model is easier to train and may be more efficient in computation.* © *2011 Society for Imaging Science and Technology.*
[DOI: 10.2352/J.ImagingSci.Technol.2011.55.3.030201]

INTRODUCTION

Color rendering is an important factor for judging the perceived quality of the color reproduction of digital images. Skin tone, as the most important category among memory colors, plays an important role in the preferred color reproduction.[1,2] Various skin color detection models have been presented in the past. A simple method is to explicitly define the range of colors in a specific color space.[3–7] In general, this method is computationally efficient and low in hardware cost, yet the accuracy may be compromised. Another method is to estimate skin color distribution from training data without deriving an explicit skin color model.[8–11] A skin probability map is constructed and may be quantized and represented as a lookup table (LUT). The probability of a color that is not located on a node may be quantized to the closest node or be computed through interpolation. While the method is fast in training and is theoretically independent of the shape of the skin color distribution, a large storage space may be required.

With the assumption that skin colors spread around a skin color center due to variations in physical conditions (e.g., skin types, capturing conditions, etc.), the skin color distribution may be approximated with a Gaussian-like function. The idea leads to the proposal of a single Gaussian model (SGM), which is formulated by a multivariate normal distribution function.[12–15] Although modeling is reasonably accurate under strict conditions, SGM may cause intolerable error in estimation and discrimination of skin colors captured in complex environments. A better approximation can be obtained using a Gaussian mixture model (GMM) which mixes a finite number of Gaussian functions.[16–18] GMM may be more appropriate than SGM if high correct detection rates are desired.[19] However, it is more complex to train and more expensive in terms of computation.

Storring et al.[20] and Fredembach et al.[21] combined standard red-green-blue (RGB) bands and near-infrared bands to detect human skin. Their results demonstrate an improved robustness over pure RGB based approaches. The approach may be generalized for the skin color detection of multispectral images, yet it is not appropriate for general consumer imaging.

From the human perception point of view, the shape of each equal probability distribution locus of a skin color boundary should be smooth. An elliptical shape may well approximate each equal probability distribution boundary. Sanger et al. applied an ellipse distribution function to express skin colors for face detection.[22] Lee and Yoo[23] concluded that the skin color cluster can be well modeled using an ellipse. This is similar to the modeling of the human perceptual color tolerance in a perceptually uniform color space (UCS) in which the visual color tolerance can be well modeled with ellipses.[24–26] Additional evidence to support elliptical modeling is that colors of an object category distribute around its prototypical color with a probability density function.[27] Due to various physical disturbances (illuminations, camera characteristics, image editing, etc.), the skin color distribution deviates from a Gaussian distribution. However, the shape of equal-distribution contours should be approximately elliptical. Since the boundary of skin colors is of nearly elliptic shape in uniform color spaces, an elliptical boundary model is adapted, modified, and expanded to compensate the lightness dependency in this study.

This article is organized in the following order: Elliptical modeling is presented first; next is the training of models; discussion of training results from different ethnic types, different image databases, and different color spaces follows; and the last section is the conclusion.

▲IS&T Member.

Received Jun. 2, 2010; accepted for publication Feb. 2, 2011; published online Apr. 4, 2011.
1062-3701/2011/55(3)/030201/12/$20.00.

ELLIPTICAL SKIN BOUNDARY MODEL

The cluster of skin colors may be approximated using an elliptical shape. Let $X_1,\ldots,X_n$ be distinctive colors (a vector with two or three coordinates) of a skin color training data set and $f(X_i)=f_i$ $(i=1,\ldots,n)$ be the occurrence counts of a color, X_i. An elliptical boundary model[23] $\Phi(X)=(X,\Psi,\Lambda)$ is defined as

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

where Ψ and Λ 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 $N=\Sigma_{i=1}^{n}f_i$ is the total number of occurrences in a training data set and $\mu=(1/N)\Sigma_{i=1}^{n}f_iX_i$ is the mean of color vectors.

Given a threshold ρ and an input color X of a pixel, X is classified as a skin color if $\Phi(X)<\rho$ and as a nonskin color otherwise. The threshold ρ trades off correct detections by false detections. As ρ increases, the correct detection rate increases; however, the false detection rate increases as well. $\Phi(X)=\rho$ defines an elliptical boundary between skin and nonskin colors. The center of ellipse is given by ψ and the principal axes are determined by Λ.

Ellipse Skin Color Modeling

Ignoring the lightness coordinate, the cluster of skin colors may be modeled using a single ellipse. In a two-dimensional (2D) chrominance space, X is expressed as $X=\binom{x}{y}$, and Λ^{-1} is represented in a matrix form,

$$\Lambda^{-1}=\begin{pmatrix}\lambda_{00} & \lambda_{01}\\ \lambda_{10} & \lambda_{11}\end{pmatrix}. \tag{4}$$

$\Phi(X)$ can be reorganized in the following form:

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

or

$$\Phi(x,y)=u_0(x-x_0)^2+u_1(x-x_0)(y-y_0)+u_2(y-y_0)^2, \tag{5b}$$

which is similar to the form used by Shen and Berns.[28]

An angle θ to rotate the x-y coordinates to the principal axes can be computed[29] by

$$\theta=0.5\arctan\left(\frac{2\lambda_{01}}{-\lambda_{00}+\lambda_{11}}\right). \tag{6}$$

The two parameters related to the principal axes are

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

$$b=\lambda_{00}\sin^2(\theta)+\lambda_{01}\sin(2\theta)+\lambda_{11}\cos^2(\theta). \tag{7}$$

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

The relationship between Λ^{-1} and a, b, and θ is

$$\begin{pmatrix}\lambda_{00} & \lambda_{01}\\ \lambda_{10} & \lambda_{11}\end{pmatrix}=\begin{pmatrix}a\cos^2(\theta)+b\sin^2(\theta) & \frac{b-a}{2}\sin(2\theta)\\ \frac{b-a}{2}\sin(2\theta) & a\sin^2(\theta)+b\cos^2(\theta)\end{pmatrix}. \tag{8}$$

To adapt an ellipse to different lightness levels, a set of ellipses for different lightness levels may be modeled. The detail will be described in a later section.

Ellipsoid Skin Color Modeling

To consider the lightness dependency of the shape of skin cluster, the cluster of skin colors in a lightness-chrominance color space may be modeled with an ellipsoid. In a three-dimensional (3D) color space, X is expressed as

$$X=\begin{pmatrix}x\\ y\\ z\end{pmatrix},$$

and Λ^{-1} is represented in a matrix form,

$$\Lambda^{-1}=\begin{pmatrix}\lambda_{00} & \lambda_{01} & \lambda_{02}\\ \lambda_{10} & \lambda_{11} & \lambda_{12}\\ \lambda_{20} & \lambda_{21} & \lambda_{22}\end{pmatrix}. \tag{9}$$

$\Phi(X)$ in Eq. (1) can be reorganized as

$$\begin{aligned}\Phi(x,y,z)=&\lambda_{00}(x-x_0)^2+(\lambda_{01}+\lambda_{10})(x-x_0)(y-y_0)\\&+(\lambda_{02}+\lambda_{20})(x-x_0)(z-z_0)+\lambda_{11}(y-y_0)^2\\&+(\lambda_{12}+\lambda_{21})(y-y_0)(z-z_0)+\lambda_{22}(z-z_0)^2.\end{aligned} \tag{10}$$

According to Eq. (3),

$$\Lambda=\frac{1}{N}\sum_{i=1}^{n}f(x_i,y_i,z_i)\times\begin{pmatrix}(x_i-x_0)^2 & (x_i-x_0)(y_i-y_0) & (x_i-x_0)(z_i-z_0)\\ (x_i-x_0)(y_i-y_0) & (y_i-y_0)^2 & (y_i-y_0)(z_i-z_0)\\ (x_i-x_0)(z_i-z_0) & (y_i-y_0)(z_i-z_0) & (z_i-z_0)^2\end{pmatrix}. \tag{11}$$

Comparing Eq. (10) to Eq. (11), $\lambda_{01}=\lambda_{10}$ and $\lambda_{21}=\lambda_{12}$. The ellipsoid function (10) can be written as

$$\Phi(x,y,z) = u_0(x-x_0)^2 + u_1(x-x_0)(y-y_0) + u_2(y-y_0)^2 + u_3(x-x_0)(z-z_0) + u_4(y-y_0)(z-z_0) + u_5(z-z_0)^2, \quad (12)$$

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

A general approach to find the three principal axes of the ellipsoid is to translate the origin of the coordinate to the center of the ellipsoid, and then to rotate three coordinates to overlap with three ellipsoid principal axes. After translating the origin of the coordinates to the center of the ellipsoid, the coordinate $X(xyz)$ becomes $X'(x'y'z')$, where

$$x' = x - x_0,$$

$$y' = y - y_0,$$

$$z' = z - z_0.$$

Denote a 3×3 rotation matrix as M_r; $X'=M_rX''$. Equation (1) can be rewritten as

$$\Phi(X'') = X''^T M_r^T \Lambda^{-1} M_r X''. \quad (13)$$

We denote a 3X3 $M=M_r^T\Lambda^{-1}M_r$, $\Phi(X'')=X''^TMX''$. To rotate X' to X'', M_r must be such that M becomes a diagonal matrix. A three-dimensional rotation may be specified with three Euler angles, α, β, and γ, to rotate around the x, y, and z axes. However, α, β, and γ cannot be derived and represented with simple forms so that M becomes a diagonal matrix. An exhaustive search approach was applied to solve the problem. Since two points on the ellipsoid surface intercepted by the longest principal axis have the longest distance to the center, the longest axis can be searched by finding those two points. Due to the symmetric behavior, it is not necessary to search the entire gamut, and only a point needs to be found (the other point is the mirror from the center). With exhaustive search within a portion of the ellipsoid, a point that has the longest distance to the center is found. A vector that connected the center and this point is the longest principal axis.

Similarly, two points on the ellipsoid surface intercepted by the shortest principal axis have the shortest distance to the center. This property was applied to find the shortest axis. Since the third principal axis is perpendicular to the other principal axes, its direction is the cross product of the vectors of the other two axes, and the axis passes through the ellipsoid center. These properties are used to find the third principal axis.

CONSTRUCTING AN IMAGE DATABASE TO TRAIN SKIN MODELS

The aim of the skin color detection determines how to collect training data. For example, if skin detection is to be used for a specific lighting condition, the training data set should be created under the same lighting condition. If skin detection is for general purposes, the training data should be collected from various images that cover different capturing conditions and skin types. Since the main purpose of this study is to develop skin color models for the preferred skin color enhancement of general digital photographic images, images captured using different digital cameras under various conditions were collected for training skin color models.

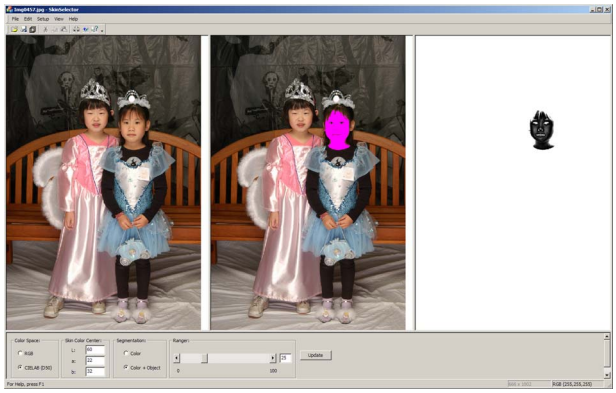

Figure 1. A tool to label skin colors.

An image database, the so-called Halloween database, is used to train skin color models and is composed of about 2500 digital images that cover Caucasian, Asian, and African facial tones. About 60% of images were captured between 2001 and 2008 using different 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 the past few years using various professional and consumer digital cameras.

A tool, *SkinSelector*, was developed to label skin colors in each image. Figure 1 shows a snapshot of the tool. The left window shows the original image. Once the mouse is pointed to a skin color and clicked, the skin color of this point is applied as the seed color (color center) to grow the color region. The span of the region is determined by a range slider that sets a color difference threshold of each pixel to the seed color. A color difference value is scaled to an 8 bit mask value between 0 and 255. A seed color has a color difference of zero and is corresponded to a mask value of zero. A color difference that equals or is greater than the threshold set by the range slider corresponds to a mask value of 255. All other color differences scale to a range of 0–255 accordingly. A mask value of 255 corresponds to nonskin colors. Pixels selected as skin colors are marked pink on the center window. The mask values of the image are represented as a grayscale image shown on the right window.

Although a RGB color space can be used to grow a region, it was found that results produced in CIELAB color space[48] were more closely correlated with the human visual perception. Therefore, CIELAB color space was selected in this study. RGB color values are converted to $L^*a^*b^*$ using the embedded ICC profile[49] of each image (or *s*RGB ICC profile if no embedded ICC profile exists).

If the segmentation method, "Color+Object" in the tool, is selected, the regional growth subjects to the con-

straint that skin colors must be connected together as a single object, which prohibits more than one isolated region. This method is used to construct the skin database for: (1) it prohibits growing similar colors to other objects and (2) it enables labeling skins of different persons with different color centers on an image.

After an object is labeled, the source image (on the left window), the image with a labeled skin object (on the center window), the image of the skin mask (on the right window), and the setting parameters are saved. An image with very high resolution is resampled to a size of about 2 Mpixels to avoid the possibility that the image is weighted higher than a lower resolution image.

Labeled pixels are used to analyze skin colors, and all other pixels that are not labeled (white pixels in the image on the right window) are ignored. As a result, labeling all skin pixels is not necessary.

After labeling all images, a script reads each image (the image on the left window) and its associated labeled image (image on the right window) and adds occurrences of skin colors to a $256 \times 256 \times 256$ RGB LUT. The reason for using 256^3 LUT is the convenience for counting occurrences of 8 bit RGB images. The number on each node of the LUT represents the occurrences of the RGB color as a skin pixel. So the number on every nonskin node is zero. Each skin color from an 8 bit per channel RGB image adds an occurrence count to the corresponding bin of the LUT. To remove noisy pixels and pixels that may be inaccurately selected as skin pixels, a small percentage of pixels with lowest occurrences are excluded from counting occurrences at the time each image is processed. In this study, the 10% least occurrence pixels were removed from each image.

The skin colors of the Halloween image database were labeled mostly by one person. The bias from the user selection should be insignificant for two reasons: First, during counting skin pixel occurrences from each image, a process was implemented to remove a small percentage of labeled skin pixels whose color histograms were below a threshold; and, second, a large number of skin pixels (in the order of billions) were collected from diverse images. Nevertheless, a second image database, the Royal Photographic Society (RPS) database, was created to verify the bias of skin labeling from different users and the dependency on training data sets. It is a collection of 626 photographic images from different sources, including indoor and outdoor images and covering different ethnic types. All images were sampled to a uniform resolution of 1200×1800. Again, SkinSelector was used to label skin pixels. The skin labeling was mostly done by a different person. A comparison of training results using these two databases is presented in the Discussion section.

RESULTS OF THE SKIN COLOR MODELING

Figure 2 shows the cluster of the selected skin colors of the Halloween database in CIELAB color space. The right one is the projection in a^*-b^* coordinates, i.e., the top-down view of the cluster. The shapes of constant-lightness slices are close to ellipses, but the sizes and locations of ellipses at different lightness are different. If an ellipse is used to fit the skin boundary on the a^*-b^* plane, it should be large enough to cover the dominant midtone skin colors, although smaller ellipses fit well for lighter and darker tones.

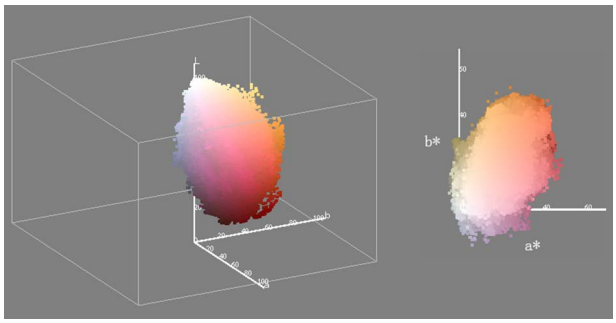

Figure 2. 3D gamut of all labeled skin pixels of Halloween database.

As described in the previous section, each labeled skin pixel in the database has a mask value between 0 and 255 to encode skin color likelihood (the higher the value, the lower the likelihood). However, mask values are binarized to skin or nonskin color for the skin color modeling in this study.

The RGB color at each node of the RGB skin occurrence LUT is converted to CIE $L^*a^*b^*$ color space, and the white point is adapted to D50 using the linear Bradford transformation which is the white adaptation algorithm used to create the official *s*RGB ICC profile.[30,31] The LUT is used to train elliptical models. The count in each bin of the LUT is the occurrence f_i in Eq. (3), and a^*b^* or $L^*a^*b^*$ of each bin location is the color X_i.

Lightness-Independent Ellipse Model (Single-Ellipse Model)

By projection of all colors to a^*-b^* coordinates (i.e., ignoring each color's lightness value), an ellipse is trained. Although the accuracy to fit skin color in an ellipse is sacrificed for simplicity and efficiency, it is adequate for some applications. Figure 3 shows the modeled ellipse in a^*-b^* coordinates to cover 95% of the labeled skin colors of Halloween database. The center coordinates are (19, 20), together with the ellipse parameters $[a, a/b, \theta]$ of [27, 1.8, −62°], where a and b are the semimajor and semiminor axes, and θ is the orientation angle of the major axis (negative degree means counterclockwise). For practical applications, the principle axes may be increased or decreased proportionally (equivalent to adjusting ρ of the elliptical model) to optimize the skin color boundary.

Lightness-Dependent Ellipse Model (Multiellipse Model)

The centers, sizes, and orientations of 2D chrominance ellipses of the skin color cluster at different lightness levels are different. Training ellipses separately on different lightness levels should improve the skin color modeling accuracy. To train lightness-dependent ellipses, the labeled skin colors of Halloween database were classified into many subsets, each containing pixels within a bin of lightness. In this study, the full range of L^* from 0 to 100 was divided into ten buckets, each occupying an L^* increment of 10 units. The training

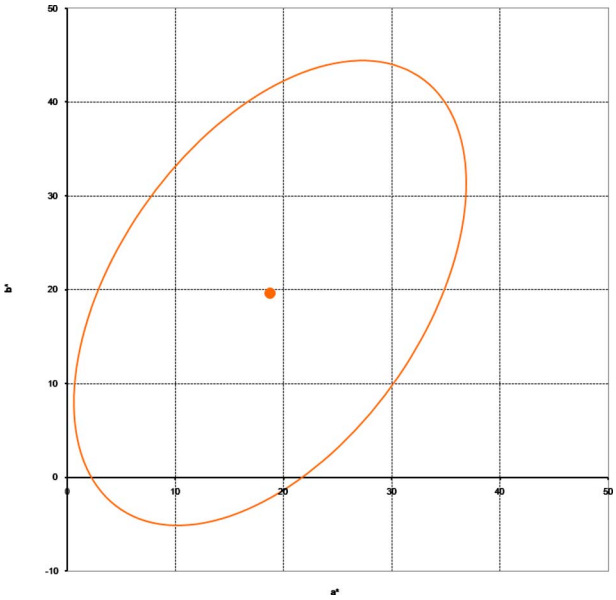

Figure 3. The trained skin color ellipse in CIELAB a^*-b^* coordinates.

data set was sorted into these ten L^* buckets. Ellipses in a^*-b^* coordinates trained for each bucket are shown in Figure 4. Each ellipse is fitted to cover 90% of skin colors 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 the L^* bucket of [90, 100]. The L^* bucket increases in order from left to right and from top to bottom in the figure. The skin color centers, sizes of principal axes, and orientations of ellipses as functions of lightness are described in the following subsections.

Skin Color Center

The chroma of each ellipse center is plotted in Figure 5. There are no data in the first bin where L^* is in the range of 0–10. A curve to fit the trained points is plotted as well. The curve is fitted with the equation $C^* = -0.000\,04L^{*3} - 0.0013L^{*2} + 0.4226L^* + 16.848$. Hue angles of ellipse centers are plotted in Figure 6. Since they are close to a constant, a line is fitted with a constant hue angle of 47.35° averaged from all hue angles.

Orientations and Sizes of Ellipses

The orientations (θ) of the trained major axes (negative means clockwise rotation) are plotted in Figure 7. The angle of the last bin for the highlight region does not follow the global trend. Since the result may be affected by lighting and white balance, this last point was ignored from curve fitting. The orientations of the major axes were fitted with a straight line by an equation: $\theta = 49 - 0.23L^*$.

The length of the trained semimajor axis in each L^* level and its fitting curve are plotted in Figure 8 where $\Phi(X)$ was set to 1 ($\rho = 1$) as the skin boundary. The semimajor axes were fitted with a polynomial equation: $a = 0.000\,004L^{*3} - 0.0127L^{*2} + 1.3331L^* - 5.0139$.

The length of the trained semiminor axis in each L^* level and its fitting curve are plotted in Figure 9. The points were fitted with a polynomial equation:

$$b = 65.452\left(\frac{L^*}{100}\right)^3 + 67.657\left(\frac{L^*}{100}\right)^2 - 5.2756\left(\frac{L^*}{100}\right) + 5.5431.$$

An Alternative Formulation of the Ellipse Model

Based on Eq. (5b), skin colors are bounded within the region of

$$u_0(x - x_0)^2 + u_1(x - x_0)(y - y_0) + u_2(y - y_0)^2 \leq \rho.$$

In this study, ρ was set to 1 to train u_0, u_1, and u_2 so that ellipses covered 90% of skin colors in the database; ρ may then be reduced or increased to optimize the skin region. The trained values of u_0, u_1, and u_2 are plotted in Figure 10. In general, each data set is smooth in the midtone area and is not smooth in shadow and highlight. Since the abrupt changes in shadow may be the result of noise and the abrupt changes in highlight may be the result of lighting variation and white imbalance, no attempt was made to fit the rapid changes at both ends of the lightness scale. All curves were to fit midtone areas accurately and extended to both ends smoothly, and the abrupt behaviors at both ends were ignored. Accordingly, u_0, u_1, and u_2 were fitted with the following equations:

$$u_0 = 0.005 + 0.006\left(\frac{L^* - 60}{30}\right)^2,$$

$$u_1 = 0.0218\left(\frac{L^*}{100}\right)^3 - 0.0678\left(\frac{L^*}{100}\right)^2 + 0.0684\left(\frac{L^*}{100}\right) - 0.0258,$$

$$u_2 = 0.0205\left(\frac{L^*}{100}\right)^3 - 0.0055\left(\frac{L^*}{100}\right)^2 - 0.0184\left(\frac{L^*}{100}\right) + 0.0108.$$

Ellipsoid Skin Color Model

Instead of modeling lightness-dependent ellipses, modeling an ellipsoid to fit the skin color boundary considerably simplifies the modeling and training process. Figure 11 shows a trained ellipsoid that covers 90% of the skin colors (black dots) in CIELAB color space. It should be noted that skin colors that are not within the ellipsoid mostly have very low occurrences. The ellipsoid center is (59, 19, 20); the principal axis parameters $[a, a/b, a/c]$ are [38, 1.4, 2.5], where a, b, and c are semiprincipal axes; and the unit vectors of three principal axes relative to the center are (0.97, −0.14, −0.19), (0.24, 0.44, 0.87), and (0.04, 0.89, −0.46). The matrix Λ is

$$\begin{bmatrix} 1404.6 & -110.2 & -125.8 \\ -110.2 & 349.1 & 223.7 \\ -125.8 & 223.7 & 656.0 \end{bmatrix}.$$

The u_i ($i = 0, 1, 2, 3, 4, 5$) coefficients are (0.000 73, 0.000 36, 0.003 71, 0.000 16, −0.002 46, 0.001 96), with $\rho = 1$.

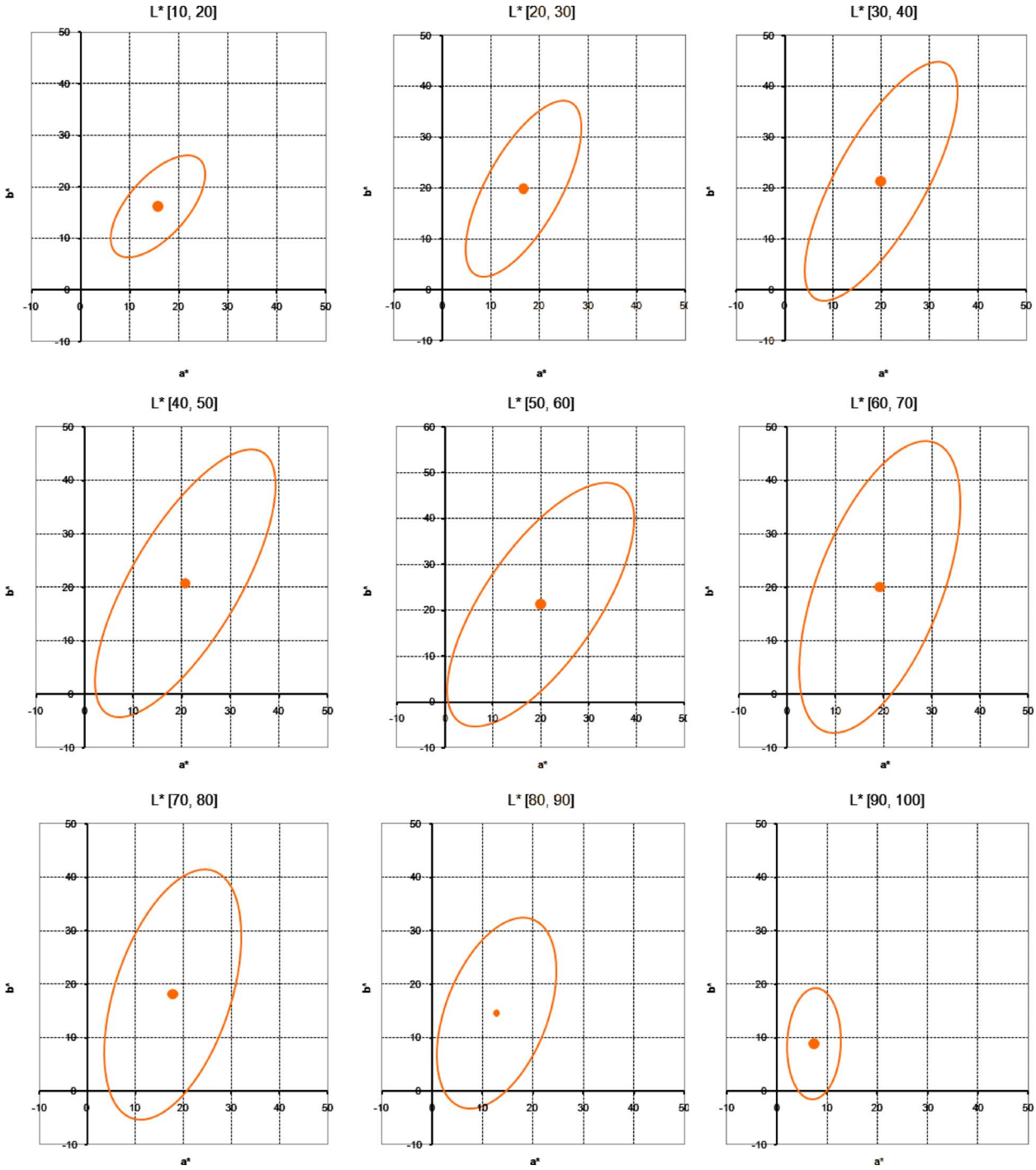

Figure 4. Skin color ellipses in different constant-lightness buckets.

Constant-lightness slices of the ellipsoid covering 90% of the skin colors are shown in Figure 12. There is no ellipse at $L^*=15$. Ellipses with L^* from 25 to 95 at intervals of 10 lightness units are plotted to compare with the ellipses modeled with the lightness-dependent ellipses shown in Fig. 4. The largest ellipse is at $L^*=65$ in both models ($L^*=65$ in Fig. 12 is comparable with the L^* bucket of [60, 70] in Fig. 4). Sizes of ellipses reduce gradually as L^* increases or decreases in both models. The orientations of the ellipses are very similar, and their eccentricities are similar as well in both models.

Skin Color Detection Accuracy

The skin color detection accuracy of a skin model is typically evaluated using true positive (TP) detection rates and false positive (FP) detection rates.[32] TP is the ratio of the number of skin pixels detected as skin pixels over the total skin pixels. FP is the ratio of the number of nonskin pixels detected as skin pixels over the total nonskin pixels. Increasing TP typically forces an increase in FP as well. In other words, to increase the likelihood that a true skin pixel is detected as a skin pixel, a nonskin pixel is more likely to be falsely detected as a skin pixel. Optimization of a skin detector should

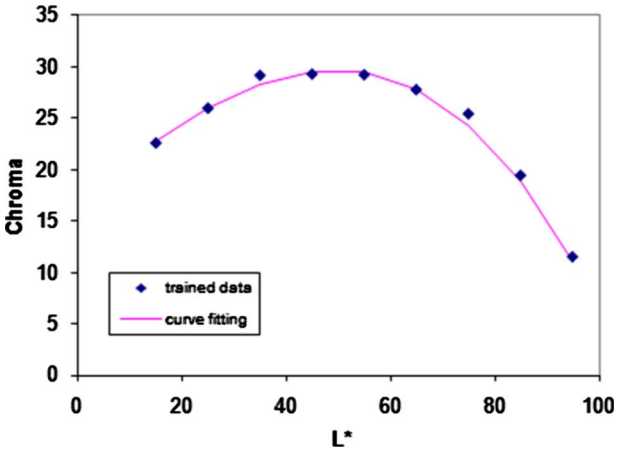

Figure 5. Chroma of the skin color centers.

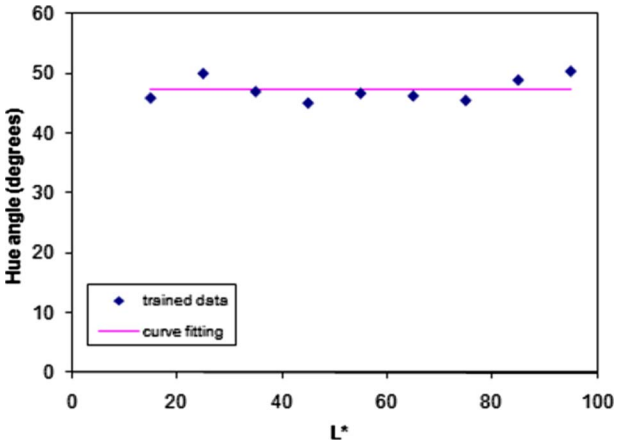

Figure 6. Hue angles (deg) of the skin centers.

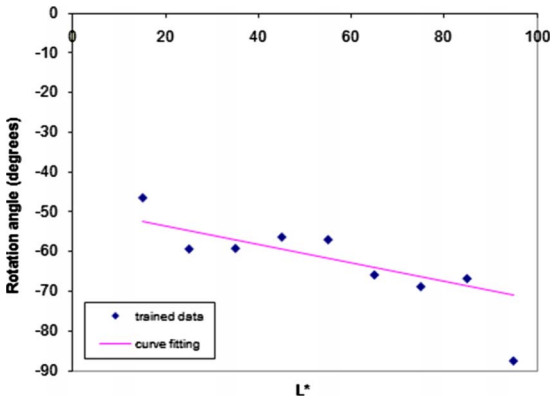

Figure 7. Orientations of the major axes.

achieve a TP as high as possible for a given FP. The relationship between TP and FP is plotted as a curve, the receiver operating characteristic (ROC) curve. The curve is very useful for determining a proper skin detection threshold for a given trade-off between TP and FP.

To verify detection rates on an image, all skin pixels of the image must be labeled. Since the Halloween image database used for skin color modeling does not have all skin

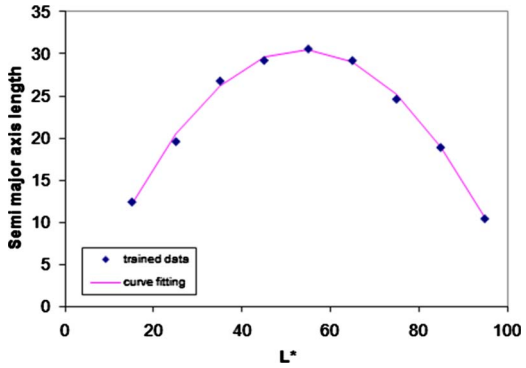

Figure 8. Semimajor axes of skin ellipses.

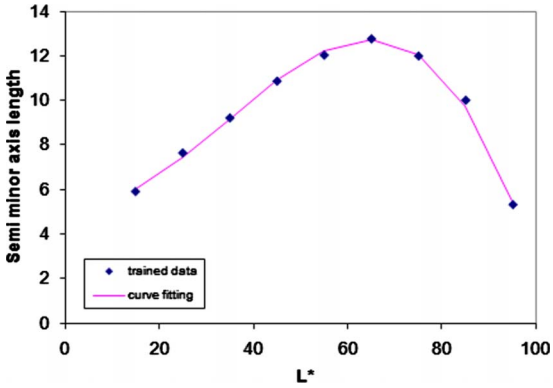

Figure 9. Semiminor axes of skin ellipses.

pixels labeled, it cannot be used to compute skin detection rates. A different image database that consists of 106 images was constructed to analyze the skin color detection accuracy of the three elliptical models. The 106 images cover different skin types and different capturing conditions. The skin pixels of each image were labeled manually using Adobe Photoshop. Each skin color model was applied to original images to detect skin colors, and the corresponding labeled images were applied to verify whether the skin detection for each pixel is correct. TP and FP were computed using all tested images.

By changing ρ of a skin model, a set of FP versus TP curves was obtained. Figure 13 shows the ROC curves of the three elliptical models. The figure shows that increasing TP is at the cost of increasing FP, as expected.

Figure 13 also shows that the single-ellipse model (the lightness-independent skin model) has the lowest detection accuracy in general, and the multiellipse model (lightness-dependent skin model) has slightly higher detection accuracy than the ellipsoid model. Because a fixed ellipse is applied to cover skin colors at different lightness levels in the single-ellipse model, a larger portion of dark colors and highlight colors that are not skin colors must be covered in order to reach the same TP as the other two models. Therefore, its FP is higher. As the FP reaches a very high value, the TP differences among the three models diminish. It demonstrates that if a high FP is acceptable, optimizing a skin color model is not critical; instead, choosing a skin color model

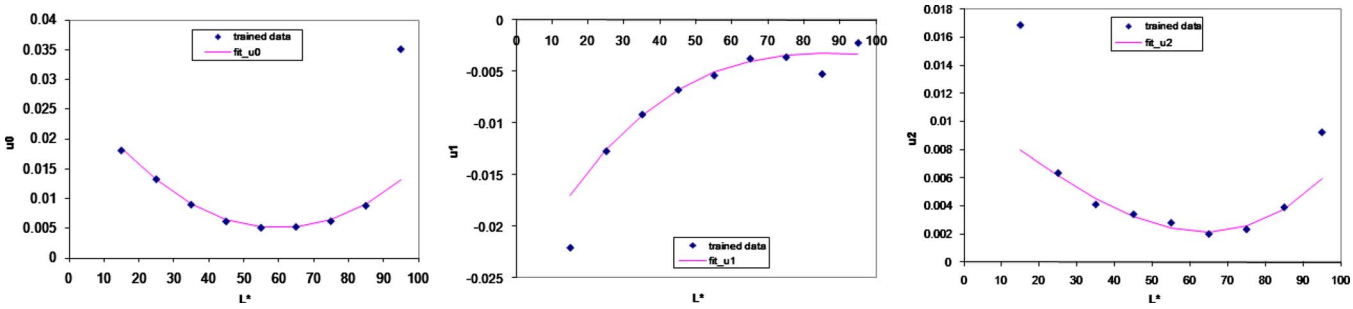

Figure 10. u_0, u_1, and u_2 of skin color ellipses.

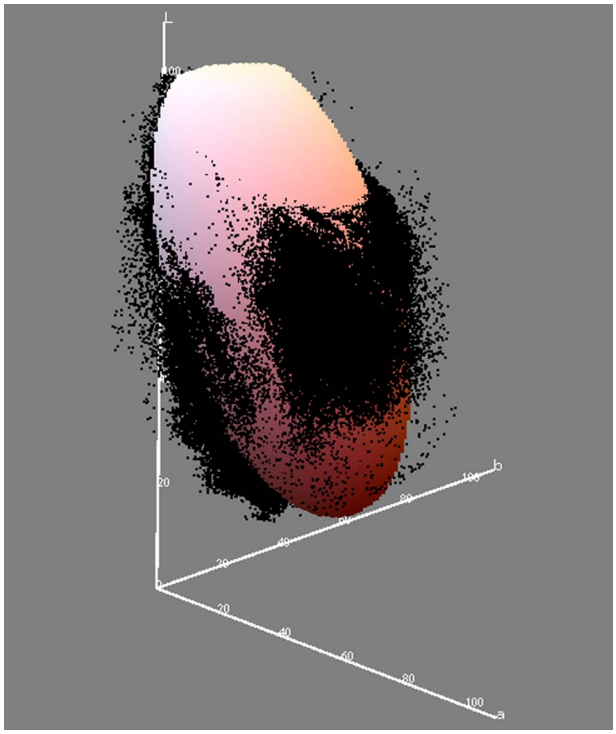

Figure 11. An ellipsoid to cover 90% of skin colors.

with high computation efficiency and low hardware cost may be more important.

Since the multiellipse model is trained at each lightness level, it should theoretically achieve a higher detection accuracy than the ellipsoid model. It is thus surprising that the detection accuracy of the ellipsoid model is so close to that of the multiellipse model, as shown in Fig. 13. Because the majority of test images have midtone skin colors, the total number of dark and light skin pixels is much smaller than the number of midtone skin pixels, and therefore the detection accuracy of dark and light skin pixels may be negligible. If both models were well optimized for midtone skin colors, the differences of detection accuracies in light and dark tones may have little influence on the overall detection accuracy.

The ellipsoid model fits skin clouds well for the midtones but not so well for light and dark tones, while the multiellipse model should fit skin color clouds well for every lightness level. To verify this hypothesis, seven dark skin images from the 106 images were chosen to compute detection rates of the three models. Their ROC curves were plotted in Figure 14. The result, namely, that the difference of detection rates between the ellipsoid and multiellipse models is larger in this case than when using all the mages was expected. Because the multiellipse model was trained on each lightness bucket, its detection accuracy is the highest. Since the skin color boundary parameter of the single-ellipse model, ρ, is adjusted to fit dark skin colors of the test images, its detection accuracy is close to that of the multiellipse model.

DISCUSSION

Knowledge of how different factors influence training results of elliptical modeling should be helpful for training models accurately and evaluating models confidently. Three important factors that will be evaluated in this section are skin types, image database, and color space.

Skin Color Modeling of Different Skin Types

The RPS database includes three subsets: a Caucasian set composed of 302 Caucasian images, an Asian set composed of 285 Asian images, and an African set composed of 28 African images. Each set was used to train a lightness-independent skin model and an ellipsoid model for each skin type. A comparison of the three ellipses is shown in Figure 15. Although the Caucasian skin color region is shifted slightly toward less chromatic and less yellowish color region, the Caucasian and Asian skin color regions are very similar. The African skin color region has a higher mean chroma, a larger chroma variation, and a smaller hue variation. Its hue range is in between those of the other two. The center coordinates, together with the semimajor axis (a), the eccentricity (a/b), and the orientation of the major axis (θ) (negative θ means counterclockwise) of three ellipses are listed in Table I.

An ellipsoid was trained for each skin type to cover 95% of the skin colors. Figure 16 shows a side by side comparison of Caucasian, Asian, and African skin color ellipsoids in CIELAB color space and their projection on a^*-b^* coordinates. The lightness ranges of Caucasian and Asian skin colors are about the same, while the African skin color region is slightly darker than the other two. The result is consistent with the 2D ellipse modeling. The Asian skin color region is slightly more yellowish and slightly more chromatic than the Caucasian skin colors, and the African skin color region is more chromatic than the other two skin color types.

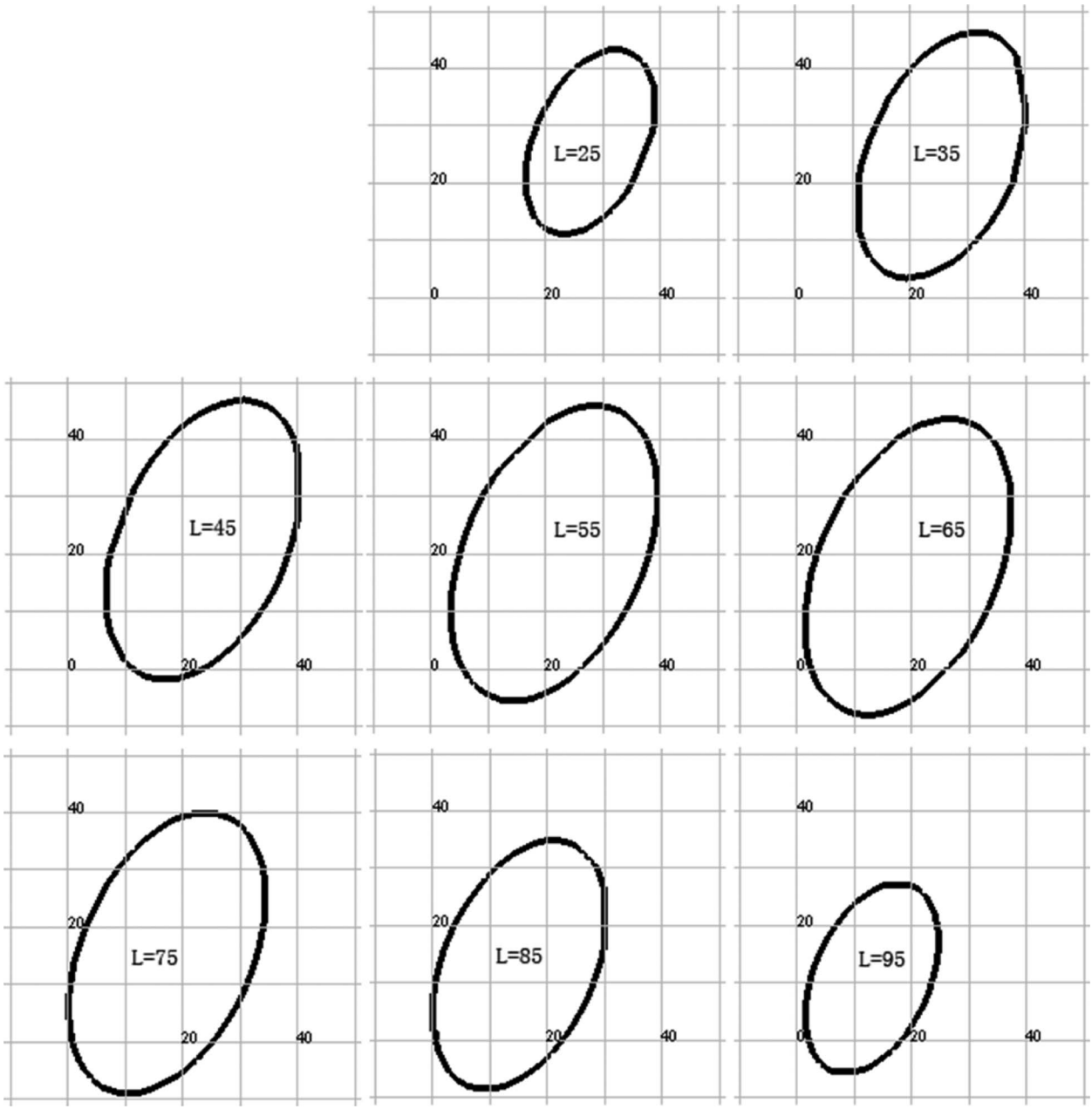

Figure 12. Constant-lightness slices of the ellipsoid covering 90% of skin colors.

Training with Different Image Databases

To study how training results are influenced by different image databases, the lightness-independent ellipse trained earlier using the Halloween database was compared with an ellipse trained using the RPS database. In Figure 17, the Halloween and RPS ellipses were trained using the same configuration parameters. The centers of the two ellipses are almost the same (the circular orange dot is the center of the RPS ellipse, and the square green dot is the center of the Halloween ellipse), the eccentricities of two ellipses are very close, and the orientations of two ellipses are about the same. However, the ellipse trained using the RPS database is larger. However, this result occurred because the person who labeled the skin pixels in the RPS image database did so more aggressively. By controlling the threshold value ρ, we can increase or decrease proportionally the size of an ellipse. With $\rho=1.25$, the Halloween ellipse is expanded to the adjusted-Halloween ellipse which is very close to the RPS ellipse. The result demonstrates that the results trained using these two different databases are very consistent, and the training result can be made independent of skin color labeling by different persons.

Skin Color Modeling with Different Color Spaces

Various color spaces (e.g., RGB, *r-g*, YC_bC_r, HSV/HIS/HSL, YUV, YIQ, $L^*u^*v^*$, $L^*a^*b^*$, etc.) have been used to define skin color gamut for skin color detection, face detection, or skin color enhancement.[33–45] Zarit et al.[9] investigated five color spaces ($L^*a^*b^*$, Fleck HS, HSV, *r-g*, and YC_rC_b) for skin detection. Their result shows that the goodness of a skin

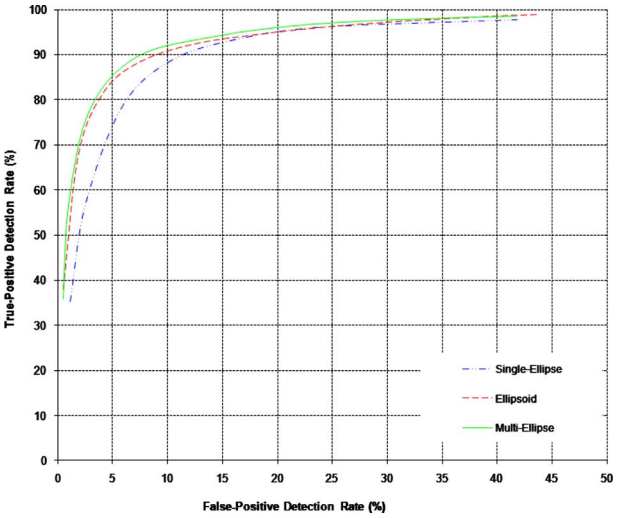

Figure 13. ROC curves of three skin elliptical models.

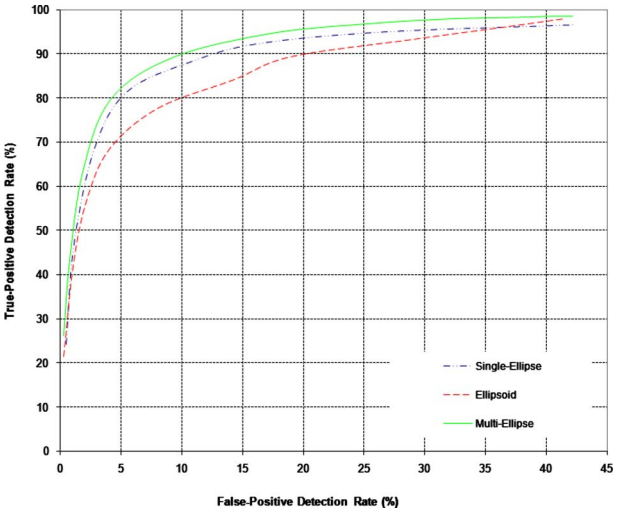

Figure 14. ROC curves of three elliptical skin models tested on dark skin images.

model depends on the color space used. Albio et al.[32] theoretically proved that "separability of the skin and nonskin is independent of the color space chosen." This disagrees with practical experience. Shin et al. evaluated skin detection using RGB color space and other eight color spaces, normalized RGB, CIE *XYZ*, CIE $L^*a^*b^*$, HIS, SCT, YC_rC_b, YIQ, and YUV, and concluded that the RGB color space provided the best separability between skin and nonskin.[46] This result may, however, only confirm that Shin's skin detection method works best in RGB color space.

In summary, a RGB color space may be more suitable for histogram based models; for a RGB LUT to store the trained skin colors can be used directly to process RGB pixels without additional color transformation. Using a luminance-chrominance color space for skin color detection reduces the interaction between luminance and chrominance, and therefore simplifies the process. If the dependence on luminance (or lightness) is ignored, skin color detection using chrominance (e.g., *r-g*, C_bC_r, a^*b^*, or u^*v^*) further simplifies the process and improves the efficiency. However, the detection rate may be compromised.

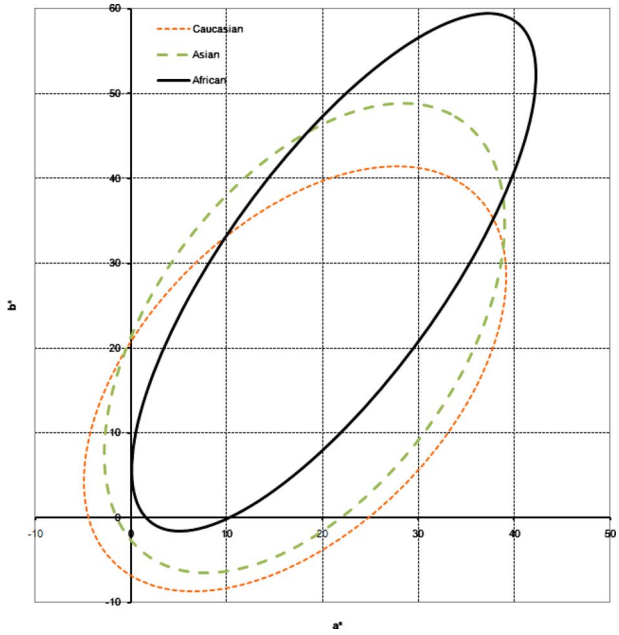

Figure 15. Caucasian, oriental, and African skin ellipses in CIELAB a^*-b^* coordinates.

Table I. Comparison of ellipse coefficients for three different skin types.

	Skin center	a	a/b	θ
Caucasian	(17, 16)	29	1.7	–52
Asian	(18, 21)	31	1.9	–60
African	(21, 29)	35	3.0	–58

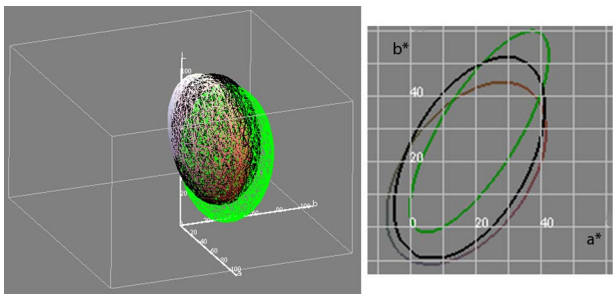

Figure 16. Caucasian (color), oriental (black), and African (green) skin color ellipsoids in CIELAB color space (left) and their projection in a^*-b^* coordinates (right).

The present study has aimed for the preferred color enhancement of digital images. Although CIELAB color space, a profile connection color space in ICC color management, was chosen for the workflow, a more uniform color space, CAM02-UCS,[47] has also been under our consideration. Therefore, skin color modeling in CIELAB and CAM02-UCS was compared.

Since CAM02-UCS is more uniform than CIELAB color space, the skin color boundary of an elliptical model in

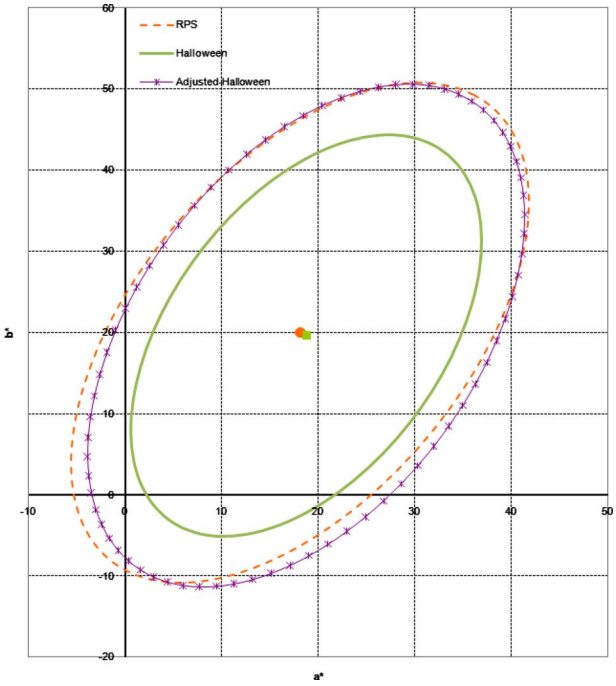

Figure 17. Skin color ellipses trained using two different databases.

CAM02-UCS may be closer to circular than in CIELAB color space. To verify whether this is true, a skin ellipsoid was trained in CAM02-UCS. The CIE *XYZ* values computed from *s*RGB colors were transformed to CAM02-UCS for elliptical modeling. The scene luminance was set to 500 lx, and the surround viewing condition was set to average.

The ellipsoid modeling result shows that the longest axis is almost parallel with the lightness axis in both CIELAB color space and CAM02-UCS. This implies that the longest axis primarily models the lightness dependency. The other two axes, b and c, primarily model the chrominance dependency. The ratio of these two axes, b/c, can be viewed as the eccentricity of an ellipsoid projected on the chromaticity axes; $b/c=1$ means that the ellipse is a circle. We found that b/c is 0.56 for the ellipsoid in CIELAB color space and is 0.73 for the ellipsoid in CAM02-UCS. The b/c closer to unity in CAM02-UCS evinces that the distribution of skin colors in CAM02-UCS is slightly more uniform than in CIELAB color space.

Skin color detection accuracies in CIELAB and CIECAM02-UCS were compared as well. The ROC curves of ellipsoid modeling in CIELAB and CIECAM02-USC were generated using the database used above to study skin detection accuracy. The results plotted in Figure 18 illustrate that the skin detection accuracy in CIECAM02-UCS is higher than that in CIELAB color space. Improved uniformity of skin colors in CAM02-UCS may be the reason that the skin detection accuracy is more accurate in this color space.

CONCLUSIONS

Skin color distributions were estimated using three elliptical models. To model skin colors with a single ellipse is simple

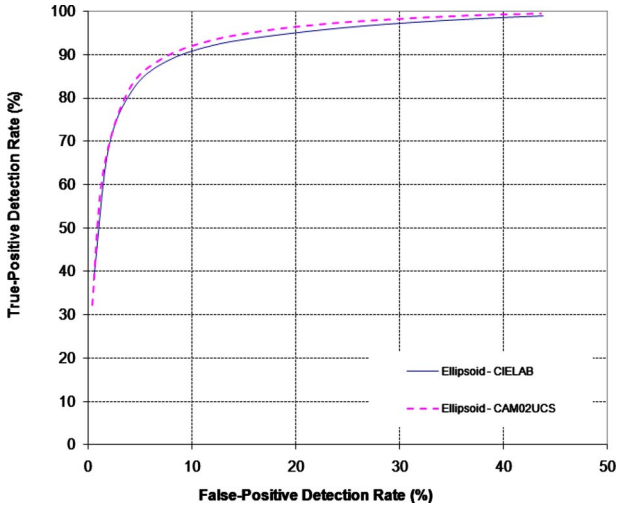

Figure 18. ROC curves of the ellipsoid modeling in CIELAB and CAM02-UCS color spaces.

in training and is efficient in computation. To cover high chroma skin colors in the midtone region, a large enough ellipse must be determined, although smaller ellipses better fit light and dark skin colors. To improve the skin color detection accuracy, a lightness-dependent ellipse model was derived to adjust skin color ellipses to fit skin colors in different lightness levels. However, formulation of lightness-dependent ellipses is complex, and the computation of skin color boundaries is accordingly less efficient. A third model, an ellipsoid skin color model, represents a compromise among modeling complexity, computation efficiency, and detection accuracy. Unlike single-ellipse modeling, it adapts skin gamut boundary to different lightness levels. Although the gamut adaptation to different lightness is not as accurate as that of the lightness-dependent ellipse model, the ellipsoid modeling is simpler to train and more efficient in computation.

The consistent results of skin color ellipses trained with two different databases verify that the method of constructing databases is reliable for skin color modeling. A separate training of Caucasian, Asian, and African skin colors demonstrates that the Caucasian skin color gamut and the Asian skin color gamut are very similar; the Asian skin colors are slightly more yellowish and slightly more chromatic than the Caucasian skin colors. The lightness ranges of the Caucasian and Asian skin types are about the same. Compared to the other two skin types, the African skin color region is slightly darker, its center is more chromatic, its chroma variation is higher, and its hue range is in between those of the other two skin types. The results of the skin color ellipsoids trained in CIELAB and CAM02-UCS color spaces reveal that CAM02-UCS is slightly more uniform in the skin color area. With ellipsoid modeling, the skin color detection in CAM02-UCS is slightly more accurate than in CIELAB color space as well.

ACKNOWLEDGMENTS

The authors would like to thank Tuo Wu and Shilin Guo who helped collecting images and labeling skins for RPS image database.

REFERENCES

[1] S. R. Fernandez and M. D. Fairchild, "Preferred color reproduction of images with unknown colorimetry", *Proc. IS&T/SID Ninth Color Imaging Conference* (IS&T, Springfield, VA, 2001) pp. 274–279.

[2] K. M. Braun, "Memory color enhancement algorithm", *Proc. ICIS'06: Int'l. Congress of Imaging Science* (IS&T, Springfield, VA, 2006) pp. 227–229.

[3] D. Chai and K. N. Ngan, "Face segmentation using skin-color map in videophone applications", IEEE Trans. Circuits Syst. Video Technol. **9**, 551–564 (1999).

[4] G. Gomez and E. F. Morales, "Automatic feature construction and a simple rule induction algorithm for skin detection", *Proc. ICML Workshop on Machine Learning in Computer Vision* (ICML, Broken Arrow, OK, 2002) pp. 31–38.

[5] R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color images", IEEE Trans. Pattern Anal. Mach. Intell. **24**, 696 (2002).

[6] J. Kovac, P. Peer, and F. Solina, "2D versus 3D color space face detection", *IEEE Fourth EURASIP Conference Focused on Video/Image Processing and Multimedia Communications* (IEEE, Piscataway, NJ, 2003) Vol. **2**, pp. 449–454.

[7] T. M. Mahmoud, "A new fast skin color detection technique", *Proc. World Academy of Science, Engineering and Technology* (WASET, Blackpool, UK, 2008) Vol. **33**, pp. 518–522.

[8] Q. Chen, H. Wu, and M. Yachida, "Face detection by fuzzy pattern matching", *IEEE Fifth International Conference on Computer Vision* (IEEE, Piscataway, NJ, 1995) pp. 591–596.

[9] B. D. Zarit, B. J. Super, and F. K. H. Quek, "Comparison of five color models in skin pixel classification", *IEEE International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems* (IEEE, Piscataway, NY, 1999) pp. 58–63.

[10] J. Brand and J. S. Mason, "A comparative assessment of three approaches to pixel-level human skin-detection", *IEEE 15th International Conference on Pattern Recognition* (IEEE, Piscataway, NY, 2000) Vol. **1**, pp. 1056–1059.

[11] L. Sigal, L. Sclaroff, and V. Athitsos, "Estimation and prediction of evolving color distributions for skin segmentation under varying illumination", *IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, Piscataway, NY, 2000) Vol. **2**, pp. 152–159.

[12] M.-H. Yang and N. Ahuja, "Detecting human faces in color images", *IEEE International Conference on Image Processing* (IEEE, Piscataway, NY, 1998) Vol. **1**, pp. 127–130.

[13] B. Menser and M. Wien, "Segmentation and tracking of facial regions in color image sequences", Proc. SPIE **4067**, 731–740 (2000).

[14] H. K. Almohair, A. R. Ramli, A. M. Elsadig, and S. J. Hashim, "Skin detection in luminance images using threshold technique", International Journal of the Computer, the Internet and Management **15**, 25–32 (2007).

[15] D.-S. Park, Y. Kwak, H. Ok, and C. Y. Kim, "Preferred skin color reproduction on the display", J. Electron. Imaging **15**, 041203 (2006).

[16] M. J. Jones and J. M. Rehg, "Statistical color models with application to skin detection", Int. J. Comput. Vis. **46**, 81–96 (2002).

[17] Q. Huynh-Thu, M. Meguro, and M. Kaneko, "Skin-color extraction in images with complex background and varying illumination", *IEEE Workshop on Applications of Computer Vision* (IEEE, Piscataway, NY, 2002) pp. 280–285.

[18] R. Hassanpour, A. Shahbahrami, and S. Wong, "Adaptive Gaussian mixture model for skin color segmentation", *Proc. World Academy of Science, Engineering and Technology* (WASET, Blackpool, UK, 2008) Vol. **31**, pp. 1–6.

[19] T. S. Caetano, S. D. Olabarriaga, and D. A. C. Barone, "Performance evaluation of single and multiple-Gaussian models for skin color modeling", *Proc. IEEE Brazilian Symposium on Computer Graphics and Image Processing XV* (IEEE, Piscataway, NY, 2002) pp. 275–282.

[20] M. Storring, H. J. Andersen, and E. Granum, "A multispectral approach to robust human skin detection", *Proc. CGIV: Second European Conference on Colour Graphics, Imaging and Vision* (IS&T, Springfield, VA, 2004) pp. 110–115.

[21] C. Fredembach, N. Barbuscia, and S. Susstrunk, "Combining visible and near-infrared images for realistic skin smoothing", *Proc. IS&T/SID 17th Color Imaging Conference* (IS&T, Springfield, VA, 2009) pp. 242–247.

[22] D. Sanger, Y. Miyake, H. Haneishi, and N. Tsumura, "Algorithm for face extraction based on lip detection", J. Imaging Sci. Technol. **41**, 71 (1997).

[23] J. Y. Lee and S. I. Yoo, "An elliptical boundary model for skin color detection", *Proc. International Conference on Imaging Science, Systems, and Technology* (CSREA, Las Vegas, NV, 2002) pp. 579–584.

[24] D. L. Macadam, "Visual sensitivities to color differences in daylight", J. Opt. Soc. Am. **32**, 247–273 (1942).

[25] G. Wyszecki and G. H. Fielder, "New color-matching ellipses", J. Opt. Soc. Am. **61**, 1135–1152 (1971).

[26] M. R. Luo and B. Rigg, "Chromaticity-discrimination ellipses for surface colour", Color Res. Appl. **11**, 25–42 (1986).

[27] S. N. Yendrikhovskij, "Image quality: Between science and fiction", *Proc. IS&T's PICS Conference* (IS&T, Springfield, VA, 1999) pp. 173–178.

[28] S. Shen and R. S. Berns, "Evaluating color difference equation performance incorporating visual uncertainty", Color Res. Appl. **34**, 375–390 (2009).

[29] H. Zeng and M. R. Luo, "Modelling skin colours for preferred colour reproduction", *Proc. IS&T/SID 17th Color Imaging Conference* (IS&T, Springfield, VA, 2009) pp. 175–180.

[30] M. R. Luo and R. W. G. Hunt, "A chromatic adaptation transform and a colour inconstancy index", Color Res. Appl. **23**, 154–158 (1998).

[31] M. Nielsen and M. Stokes, "The creation of the sRGB ICC profile", *Proc. IS&T/SID Sixth Color Imaging Conference* (IS&T, Springfield, VA, 1998) pp. 253–257.

[32] A. Albio, L. Torres, and E. J. Delp, "Optimum color spaces for skin detection", *Proc. IEEE International Conference on Image Processing* (IEEE, Piscataway, NJ, 2001) Vol. **1**, pp. 122–124.

[33] S. Birchfield, "Elliptical head tracking using intensity gradients and color histograms", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (IEEE Computer Society, Los Alamitos, CA, 1998) pp. 23–25.

[34] J. Cai, A. Goshtasby, and C. Yu, "Detecting human faces in color images", *Proc. IEEE International Workshop on Multi-Media Database Management Systems* (IEEE, Piscataway, NJ, 1998) pp. 124–131.

[35] L. Jordao, M. Perrone, J. P. Costeira, and J. Santos-Victor, "Active face and feature tracking", *Proc. IEEE International Conference on Image Analysis and Processing* (IEEE, Piscataway, NJ, 1999) pp. 572–576.

[36] D. Chai and A. Bouzerdoum, "A Bayesian approach to skin color classification in YCbCr color space", *IEEE TENCON* (IEEE, Piscataway, NJ, 2000) Vol. **2**, pp. 24–27.

[37] R. Qian, "System for detecting skin-tone regions within an image", US Patent 6,332,033 (2001).

[38] B. Martinkauppi, "Face colour under varying illumination—Analysis and applications", Ph.D. dissertation, University of Oulu, 2002.

[39] S. L. Phung, A. Bouzerdoum, and D. Chai, "A novel skin color model in YCbCr color space and its application to human face detection", *Proc. IEEE International Conference on Image Processing* (IEEE, Piscataway, NJ, 2002) Vol. **1**, pp. 289–292.

[40] F. Tomaz, T. Candeias, and H. Shahbazkia, "Improved automatic skin detection in color images", *Proc. Seventh Digital Image Computing: Techniques and Applications* (IET, Sydney, 2003) pp. 419–427.

[41] F. Boussaid, D. Chai, and A. Bouzerdoum, "On-chip skin detection for color CMOS imagers", *Proc. IEEE International Conference on MEMS, NANO and Smart Systems* (IEEE, Piscataway, NJ, 2003) pp. 356–361.

[42] F. Gasparini and R. Schettini, "Skin segmentation using multiple thresholding", Proc. SPIE **6061**, 60610F (2006).

[43] F. Gasparini, S. Corchs, and R. Schettini, "Recall or precision-oriented strategies for binary classification of skin pixels", J. Electron. Imaging **17**, 023017 (2008).

[44] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-color modelling and detection methods", Pattern Recogn. **40**, 1106–1122 (2007).

[45] A. Choudhury, M. Rogers, and B. Gillam, "A novel skin tone detection algorithm for contraband image analysis", *Proc. IEEE Third International Workshop on Systematic Approaches to Digital Forensic Engineering* (IEEE, Piscataway, NJ, 2008) pp. 3–9.

[46] M. C. Shin, K. I. Chang, and L. V. Tsap, "Does color space transformation make any difference on skin detection?" *IEEE Sixth Workshop on Applications of Computer Vision* (IEEE, Piscataway, NJ, 2002) pp. 275–279.

[47] M. R. Luo, G. Cui, and C. Li, "Uniform colour spaces based on CIECAM02 colour appearance model", Color Res. Appl. **31**, 320–330 (2006).

[48] R. W. G. Hunt, *Measuring Colour*, 3rd Edition (Fountain Press, England, 1998), pp.53–72.

[49] International Color Consortium website, www.color.org, accessed 2010. model.