ADAPTIVE SKIN DETECTION UNDER UNCONSTRAINED LIGHTING CONDITIONS USING A BIGAUSSIAN MODEL AND ILLUMINATION ESTIMATION

JIAN-HUA ZHENG, CHONG-YANG HAO, YANG-YU FAN AND XIAN-YONG ZANG

Electronic & Information Engineering Institute of Northwestern Polytechnical University, Xi'an, 710072, CHINA
e-mail: zheng_jianh@163.com; zheng_jianhua@126.com;

ABSTRACT

An algorithm is proposed to improve the performance of skin detection algorithms under poor illumination conditions. A hybrid skin detection model is addressed to solve these problems by combining two Gaussian models of skin under normal conditions and bright illumination. According to the distribution of the combined models, the algorithm automatically evaluates the skin segmentation result of an adaptive threshold algorithm based on a Gaussian model by estimating the illumination conditions of image. If the estimation result shows that the illumination condition is very different from the normal one, the skin color of the original image needs compensation, and then the algorithm feeds the compensated image back to the Gaussian model for finer skin detection. The experimental results show that our algorithm can cope with a complex illumination change and greatly improve skin classification performance under inferior illumination conditions.

Keywords: adaptive procedure, Bigaussian model, compensation, illumination estimation, skin detection.

INTRODUCTION

Skin color is an important visual cue for face detection, face recognition, visual tracking, video and image comprehension. A lot of algorithms in the above areas use a skin color detection algorithm as a post-processing step to separate skin regions from the background of a scene and treat the skin regions as candidate faces for detecting and tracking. Therefore, precise and reliable skin region detection and segmentation under different conditions is the key factor to improve the performance of these algorithms in face detecting and tracking.

Most existing skin detection algorithms work well in a normal environment, but are not reliable in the case of unpredictable and drastically changing real-world environments. Under a distinct change of illumination, skin color shows up as too bright or too dark, or exists with highlight regions somewhere on the forehead, cheekbone and arms, or with shadow regions somewhere over the face. Even in varying colored light conditions, the colors that skin shows are very different from the original skin color. In the above-mentioned cases, most existing skin detection algorithms often detect nothing or fragmented skin regions, which will seriously influence the performance

of a face detecting and tracking algorithm based on the face color cue. As a result, skin detection under the above conditions has become a 'hot' issue in recent years. In order to deal with a dynamically changing environment, a robust algorithm is proposed in this paper by automatically evaluating and adjusting the skin segmentation result based on a hybrid skin detection model.

SKIN COLOR DETECTION UNDER UNCONSTRAINED ILLUMINATION

Conventional skin detection algorithms consider that the distribution of human skin color of different people is clustered in a chromatic color space and can be represented by a Gaussian model. By measuring the probability that each pixel belongs to the skin cluster, skin colors and non-skin colors are separated. The research shows that luminance may vary across a person's face due to the ambient lighting and is not a reliable measure in separating skin from non-skin regions. It follows that skin colors of different people differ much less in color than in brightness. In other words, skin colors of different people are very close. So luminance can be removed from the skin color representation in the chromatic color space, and a 'pure' color space, without the luminance component,

is used to modeling skin color; for instance, the Cb-Cr and I-Q space instead of the YCbCr and YIQ space. But under significantly changing illumination, skin colors of different people, even of one person, differ a lot in color. This changes the skin color distribution and weakens the skin clustering effect in the chromatic space. In order to segment human skin regions from non-skin regions based on color, we need a reliable skin color model that is adaptable to people of different skin colors and to different lighting conditions.

To reduce illumination disturbance, a normalized color space is used in the representation of the human skin model in many proposed approaches; for instance, the image is transformed from RGB space to r-g space, where r = R/(R+G+B), g = G/(R+G+B), or from YCbCr space to NCb-NCr space, where NCb = Cb/(Y+Cb+Cr), NCr = Cr/(Y+Cb+Cr). This can reduce the brightness dependence in the chromaticity coordinates r and g, or NCb and NCr. This technique is the simplest solution for unpredictable illumination conditions. But it can only achieve a minor improvement in skin detection results and is successful only if the lighting conditions do not change too dramatically.

Under varying color lighting conditions, skin color may change not only in the luminance component, but also in the chrominance components. For instance, captured skin color deviates towards bluish if the face is illuminated by fluorescent light, or towards reddish under tungsten light. One solution for skin detection under colored lighting conditions is to correct the image color first, which is also called color constancy. But most color constancy algorithms are constraint approaches. The priori conditions these algorithms rely on, such that the average color in the scene must be gray, and that the illumination change must be global, are rigorous only sometimes for real-life cases.

For skin detection under unconstrained lighting conditions, a single Gaussian model is not sufficient to model the distribution of human skin color. Adaptive skin selection approaches model the distribution of skin color as multiple fixed Gaussian models and adaptively select the best one of the Gaussian models as the skin model in measuring the skin-like color probability. Phung *et al*. (2002) argue that the decision boundary between skin and non-skin in the chrominance plane is reduced for low and high luminance. Therefore, in their approach, the luminance component is taken into account in skin and non-skin classification. The distribution of skin color is modeled with three Gaussian clusters that correspond approximately to three levels of luminance: low,

medium and high. If the minimum distance from a pixel to the three Gaussian clusters is below a certain threshold, the pixel is classified as skin. Wong *et al*. (2003) separate the Gaussian model into six groups along the luminance axis, any one of which has its own skin decision boundary. A pixel in the image is classified as skin if it belongs to any one of the six groups. Their approaches constitute a successful solution to skin detection under poor or strong lighting condition. But in real life, the manner of the illumination change is more complex.

The skin model mixture approach also models the distribution of skin color as multiple Gaussian components, but the skin decision is based on the mixture contribution of these Gaussian components. Tang *et al*. (2000) use Gaussian models under two normalized color spaces, r-g and NCb-NCr. The image is first transformed into the two normalized color spaces, and next, skin color similarities of each normalized image pixel are measured in the r-g space and in the NCb-NCr space, respectively. The final skin color similarity of each pixel is obtained by combining the two similarities in the two normalized color space. In the Gaussian mixture model (Raja *et al*., 1998; Yang and Ahuja, 1999), the conditional density for an image pixel belonging to skin is modeled as a mixture with multiple component densities. Each component is Gaussian with its own mean and covariance matrix, and the mixture parameters decide the contribution of each component to the skin similarity. Expectation Maximization (EM) provides an effective maximum likelihood algorithm for fitting the mixture. The Gaussian mixture, in fact, models the different skin tones as different Gaussian clusters, which enhances the ability of the skin model adapting to the change of skin tones under different illuminations.

In order to cope with a wide variation in illumination, we require a dynamically adaptive skin model. The parameters of the skin model are learned and updated online to reflect the changing color of skin under varying illumination. Model updating (Sahbi *et al*., 2001) is achieved by recursively adapting the mean, covariance matrix and prior probabilities of each Gaussian cluster using pixels from the detected face region. Since the face is generally oval in shape, pixels from an oval region on the face tracking result may be taken as training pixels. But this incurs a risk, because not all pixels in the oval region can be safely treated as skin, especially when tracking failures occur. Therefore, avoiding color data from the erroneous frames to be used to adapt the skin model, tracking failures have to be detected for frames-selective adaptation (Raja *et al*., 1998). The skin locus (Soriano

et al., 2000; 2003), which is a chromatic constraint using the knowledge of the range of skin color under the normalized color coordinates, is added to geometric constraints for selecting training pixels to update the skin model. In general, the dynamically adaptive skin model is used in face detection in video sequences or real-time face tracking, and the training data for updating the model is sampled from the observed frames. It usually assumes that the lighting conditions must change smoothly over time and that the updated skin model based on the previous frames must reflect the lighting change in the next frame correctly. Therefore, it is not suitable for face detection in still images obtained under unconstrained lighting conditions.

Unlike the previous work, we use two Gaussian skin models, namely one Gaussian model as the ground truth model for skin detection and segmentation, and a second one for adaptive skin segmentation result evaluation. The adaptive skin detection scheme based on this model can cope with complex illumination changes, and the adaptive method is suitable for face detection in both still images and live video.

BIGAUSSIAN SKIN DETECTION MODEL

The selection of a color model to represent human skin color is important for face detection in a color image. The YCbCr color space is suitable for real-time application and is used in many image and video standards. But in order to show the efficiency of the proposed approach, the YCbCr space is used instead of normalized color spaces like r-g and NCb-NCr.

The distribution of skin color can be represented by a Gaussian model $G(\boldsymbol{m},\boldsymbol{C})$, with the mean

$$\boldsymbol{m} = E(\boldsymbol{x}) = (\overline{C}_b, \overline{C}_r)^T, \tag{1}$$

where $\boldsymbol{x} = (C_b, C_r)^T$ is the chrominance vector, and the covariance

$$\boldsymbol{C} = E[(\boldsymbol{x}-\boldsymbol{m})(\boldsymbol{x}-\boldsymbol{m})^T] = \begin{pmatrix} \sigma_{Cb,Cb} & \sigma_{Cb,Cr} \\ \sigma_{Cr,Cb} & \sigma_{Cr,Cr} \end{pmatrix}. \tag{2}$$

Selecting training pixels from manually segmented skin regions can access the ground truth skin model parameters. But the model parameters may vary for individually selecting training data. Under changing illumination, the color of skin pixels have also changed and distributed outside of the skin color cluster according to the normal skin color model, and the conventional skin detection algorithms can be unstable. If selecting distinct skin color pixels under

those conditions into the training set when accessing the distribution model parameters, the skin model is trained and adapted to accept the skin pixels under those conditions as skin. But it will weaken the skin color clustering effect in the chromatic space and make the performance of skin detection algorithms worse, because many non-skin pixels are incorrectly classified as skin pixels. Consequently, a strongly clustering skin color model leads to an increasing skin false rejection and a decreasing skin false alarm. On the contrary, a weakly clustering skin color model leads to a decreasing skin false rejection and an increasing skin false alarm. That is to say, there are constraint relations between the degree of skin color clustering and the performance of the skin detection algorithm. A better clustering skin model can be evaluated as that having both smaller skin errors and meanwhile having a higher correct skin decision.

In order to investigate skin color distribution, skin pixels obtained under different lighting conditions from persons of different ethnicities - Asian, Caucasian and African - are used. We draw one point for each skin sample in YCbCr color space according to the luminance and chrominance values of each skin sample, and finally we get the skin color cloud shown in Fig. 1, where Fig. 1a represents the distribution of human skin colors under normal illumination. From Fig. 1a, we can clearly find that the skin distribution boundary in the chrominance plane varies along the luminance axis. The middle part of the skin color cloud has a larger boundary; on the contrary, the low and high parts have a smaller boundary. This implies that different skin decision boundaries should be applied for different lighting conditions. But unlike the previous work that splits the skin distribution into three parts (Phung *et al.*, 2002) or six parts (Wong *et al.*, 2003) along the luminance axis, we divide skin color samples into two training sample sets for normal and special environments. To determine the distribution of human skin color under normal conditions, skin samples are extracted from images obtained under normal illumination and selected equally between different ethnicities. Skin samples under special conditions are gathered from bright skin regions in images obtained under strong lighting conditions, but with the exclusion of samples from the highlighted skin regions. Fig. 1a and Fig. 1c represent the skin color cloud under normal and strong lighting conditions. The projection of the skin color cloud on the Cb-Cr plane, which is illustrated in Fig. 1b and Fig. 1d, can be considered to represent the largest skin distribution boundary on the Cb-Cr plane.
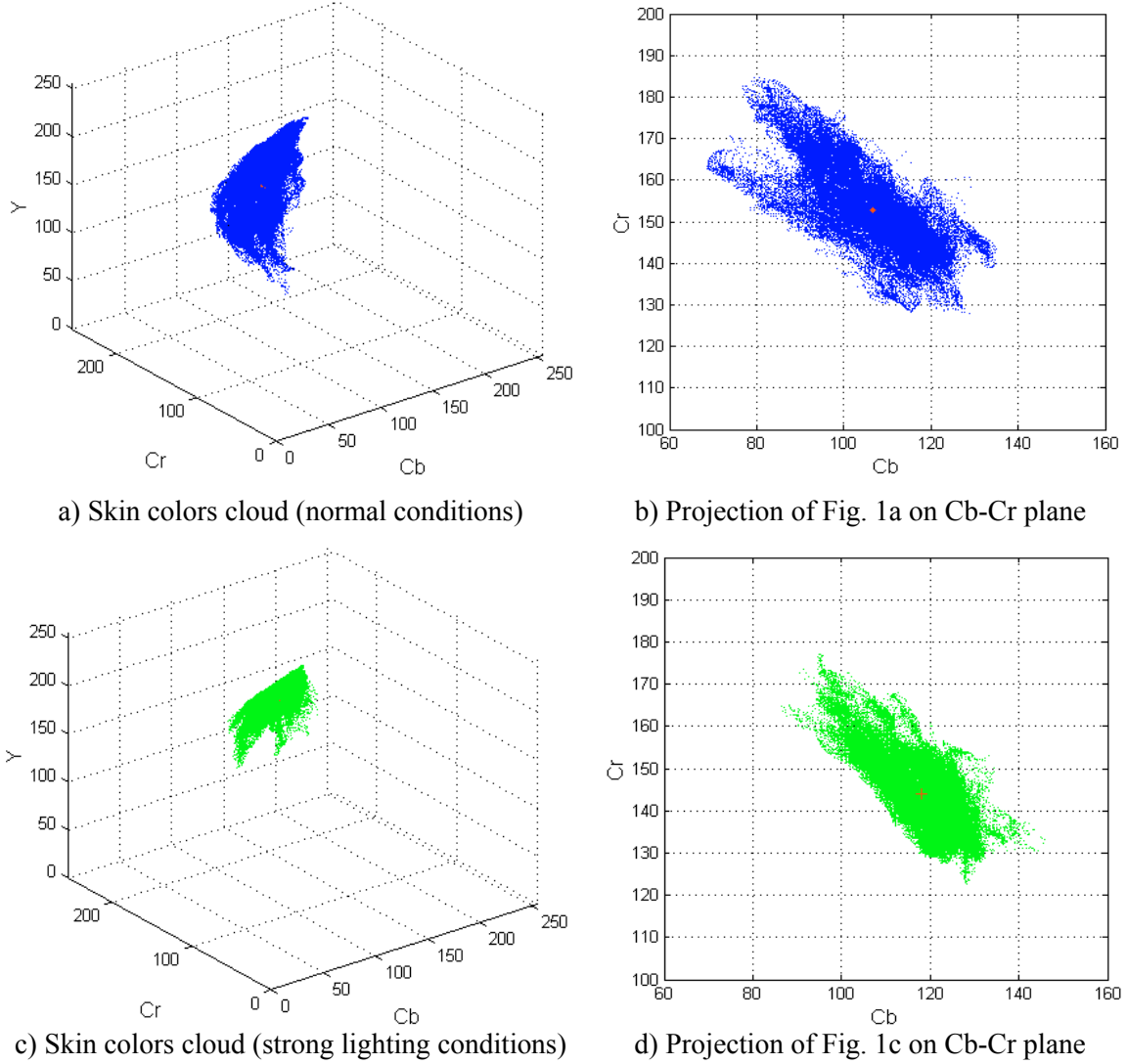
a) Skin colors cloud (normal conditions)

b) Projection of Fig. 1a on Cb-Cr plane

c) Skin colors cloud (strong lighting conditions)

d) Projection of Fig. 1c on Cb-Cr plane

Fig. 1. *Skin colors cloud of training samples for normal and special conditions in YCbCr space*

.

As discussed in section 2, the ability of a skin model adapting to illumination changes can be strengthened if the skin samples from unconstrained lighting conditions are modeled as separate skin models. Therefore, each of the skin color sample sets is modeled as a Gaussian model in the proposed approach. Thus, we get two Gaussian skin models. The formal one under normal conditions can be represented by $G_s(\boldsymbol{m}_s, \boldsymbol{C}_s)$, standing for the standard Gaussian skin model. The latter one, obtained under special conditions, can be represented by $G_w(\boldsymbol{m}_w, \boldsymbol{C}_w)$, standing for the special skin model. By combining the two clustering Gaussian models, we get the Bigaussian skin detection model. Fig. 2 shows the distribution of the model fitted by our data. In the following algorithm, only the standard skin model is used as the ground truth model in skin detection and segmentation, whereas the special skin model is used as a reference to take part in the automatic estimating and adjusting steps in the adaptive algorithm.
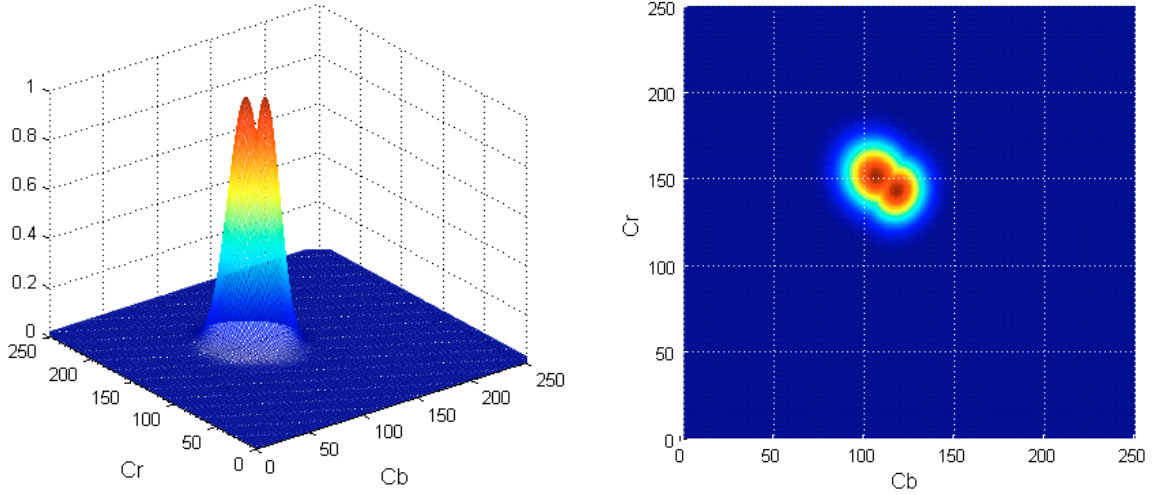
Fig. 2. *Bigaussian skin detection model (left) and its projection on chrominance plane (right).*

ILLUMINATION ESTIMATION AND COMPENSATION

By comparing Fig. 1b and Fig. 1d, it can be found that the skin distribution projection shifts on the Cb-Cr plane noticeably, which is due to the different lighting conditions when obtaining the skin color samples. This implies that the change of lighting conditions affects not only the luminance component, but also the chrominance components of skin pixels. The shift orientation and the shift range can be used to estimate the environmental illumination.

Once the skin color regions have been detected from a test image, we can estimate the illumination conditions in the image based on the detected skin regions. Since the illumination conditions under which we have obtained the standard Gaussian model are known and treated as normal standard conditions, we use the chromaticity shift of detected skin regions on the Cb-Cr plane to check the illumination diference between the normal conditions and conditions derived from the test image. If the illumination in the test image is estimated to be close to normal conditions, the skin segmentation result based on the standard Gaussian model may be good enough, and most of the skin pixels in the test image could have been detected. On the contrary, if the illumination is estimated to be very different from normal conditions, the illumination change may affect the skin pixels chrominance values. Therefore, the skin segmentation result could be unreliable and needs improvement.

In the proposed algorithm, four statistics derived from the skin models are used. They are the chrominance component mean values $\boldsymbol{m}$ and covariance $\boldsymbol{C}$ in the Gaussian distribution model, standard deviation $\boldsymbol{S}$ of chrominance components and luminance mean value $\overline{Y}$. The vector that characterizes the skin cluster center is defined as $\boldsymbol{v} = (\overline{Y}, \boldsymbol{m})^T$, the components of which are the mean luminance and chrominance values. For the standard skin model $G_s(\boldsymbol{m}_s, \boldsymbol{C}_s)$ and the special skin model $G_w(\boldsymbol{m}_w, \boldsymbol{C}_w)$, we have $\boldsymbol{v}_s = (\overline{Y}_s, \boldsymbol{m}_s)^T$ and $\boldsymbol{v}_w = (\overline{Y}_w, \boldsymbol{m}_w)^T$, where the subscripts s and w denote the corresponding elements of the normal and special skin models, respectively.

Suppose that skin color regions have been detected from an image based on the standard skin model and Gaussian probability metrics. We now evaluate how seriously the illumination change affects the chromaticity of skin pixels in the image and decide whether or not the detected result needs improvement. As the majority of pixels in detected skin regions can be safely treated as reliable ground true skin pixels, we can consider the detected skin regions as known skin in the image. For simplification, we use the mean values of the skin regions to characterize the whole detected skin region. The mean values can be treated as a known skin pixel P in the image, which is sure to belong to the standard skin cluster.

So for the detected skin regions, we have the mean vector $\boldsymbol{v}_p = (\overline{Y}_p, \boldsymbol{m}_p)^T$. Three distances related to P are determined as follows:

$$d_{ps} = v_s - v_p = (\Delta \overline{Y}_{ps}, \Delta m_{ps})^T,$$
$$d_{pw} = v_w - v_p = (\Delta \overline{Y}_{pw}, \Delta m_{pw})^T, \qquad (3)$$
$$d_{sw} = v_w - v_s = (\Delta \overline{Y}_{sw}, \Delta m_{sw})^T,$$

where d_{ps} is the distance between P and the standard skin cluster center, d_{pw} is the distance between P and special skin cluster center, and d_{sw} is the distance between the two cluster centers.

On the chrominance plane, Cb-Cr, by measuring the position of the known skin pixel P in the standard skin cluster region, we can estimate the illumination and color deviation of the original image. If P is out of the maximum inscribed circle of the standard skin cluster, which is centered in the standard skin cluster center and has the maximum radius of all inscribed circles of the standard skin cluster, the detected skin results are affected by the illumination conditions of the image and need improvement; otherwise, the results are accepted as the final best results.

The evaluation criteria can be represented as

$$\tilde{d}_{ps} > R \,, \tag{4}$$

where R is the radius of the maximum inscribed circle of the standard skin cluster, and $\tilde{d}_{ps}$ is the Euclidean distance between the known skin pixel P and the standard skin cluster center on the chrominance plane Cb-Cr. So,

$$\tilde{d}_{ps} = \left| d_{ps}(C_b, C_r) \right| = \sqrt{d_{ps}^2(C_b) + d_{ps}^2(C_r)}$$

$$= \sqrt{\Delta m_{ps}^2(C_b) + \Delta m_{ps}^2(C_r)} \,, \tag{5}$$

where $\Delta m_{ps}(C_b)$ and $\Delta m_{ps}(C_r)$ are the C_b and C_r components of chrominance vector Δm_{ps}, respectively, and $|\cdot|$ denotes the absolute value.

For fast evaluation, we need a simpler evaluation equation than Eq. 4, hence we zoom in and zoom out the left part value of Eq. 4.

For $\left| \Delta m_{ps}(C_b) \right| \geq 0$ and $\left| \Delta m_{ps}(C_r) \right| \geq 0$, we have:

$$\tilde{d}_{ps} \leq \left| \Delta m_{ps}(C_b) \right| + \left| \Delta m_{ps}(C_r) \right| \,, \tag{6}$$

and $$\tilde{d}_{ps} \geq \max \left\{ \left| \Delta m_{ps}(C_b) \right|, \left| \Delta m_{ps}(C_r) \right| \right\}. \tag{7}$$

As P is a reliably known skin pixel, it can be considered as a special point in the standard skin cluster marked as $x = P$. So, we derive as follows:

$$\left| \Delta m_{ps}(x|_{x=P}) \right|^2 = E \left| \Delta m_{ps}(x) \right|^2 = E \left| m_s - m_p \right|^2$$

$$= E \left| m_p - m_s \right|^2 = E \left| x|_{x=P} - E_s(x) \right|^2 = D_s(x) = S_s^2(x).$$

$S_s(x)$ is the standard deviation of x, so $S(x) > 0$.

Then we have:

$$\left| \Delta m_{ps}(C_b) \right| = S_s(C_b) \text{ and } \left| \Delta m_{ps}(C_r) \right| = S_s(C_r).$$

So, we zoom out the left part of (6) as follows:

$$\sqrt{\Delta m_{ps}^2(C_b) + \Delta m_{ps}^2(C_r)} = \sqrt{S_s^{\,2}(C_b) + S_s^{\,2}(C_r)}$$

$$> \sqrt{2 S_s(C_b) \times S_s(C_r)} \geq \sqrt{2} \times \sqrt{\min^2 \left\{ S_s(C_b), S_s(C_r) \right\}}$$

$$> \min \left\{ S_s(C_b), S_s(C_r) \right\} \geq R \,. \tag{8}$$

From Eq. 6 and Eq. 8, we have the simplified evaluating equation

$$\begin{cases} |\Delta m_{ps}(C_b)| + |\Delta m_{ps}(C_r)| > \min[S_s(C_b), S_s(C_r)] \\ \Delta m_{ps}(C_b) < 0 \text{ and } \Delta m_{ps}(C_r) > 0 \end{cases}, \tag{9}$$

where S_s is the chrominance standard deviation of the standard skin cluster.

If the evaluation is based on the special skin cluster, we obtain another form of simplified evaluating equation, namely

$$\begin{cases} \max \left\{ |\Delta m_{pw}(C_b)|, |\Delta m_{pw}(C_r)| \right\} > \min[S_w(C_b), S_w(C_r)] \\ \Delta m_{pw}(C_b) > 0 \text{ and } \Delta m_{pw}(C_r) < 0 \end{cases}, \tag{10}$$

where S_w is the chrominance standard deviation of the special skin cluster.

If Eq. 9 or Eq. 10 is satisfied, the illumination conditions of the test image are estimated as having seriously affected the chromaticity components of skin pixels. Therefore, we must compensate for the illumination change to improve the skin detection results.

Illumination compensation is done on the whole image by adjusting chromaticity components so as to minimize the chromaticity shift. But it could be a risk to accept non-skin pixels as skin pixels because the chromaticity components of non-skin pixels are changed. Therefore, the compensation values must be controlled within a certain degree so that we can safely accept skin-like color pixels as skin.

Because the illumination conditions under which we obtain the special skin color model are known, the chromaticity shift of special skin color can be a reference to access the safe range for compensation. Thus, the adjusting values in the YCbCr color space are estimated using the distance of the two skin clusters d_{sw}. The adjusting factor is defined as

$$\boldsymbol{\eta} = (\eta_Y, \eta_{Cb}, \eta_{Cr})^T = \begin{pmatrix} 1 - \boldsymbol{d}_{sw}(\overline{Y}) \times \omega \\ 1 + |\boldsymbol{d}_{dw}(C_b)| \times \omega \\ 1 + |\boldsymbol{d}_{sw}(C_r)| \times \omega \end{pmatrix}, \quad (11)$$

where ω is the scale factor ($\omega = 1\%$), and the adjusting factor of the chrominance pair components (C_b, C_r) is determined by

$$\boldsymbol{\eta}(C_b, C_r) = (\eta_{Cb}, \eta_{Cr})^T, \quad (12)$$

and the adjusting factor of luminance is determined by

$$\boldsymbol{\eta}(Y) = \eta_Y. \quad (13)$$

In addition, since P is a reliable skin pixel of the image and belongs to the standard skin cluster, the chromaticity shift of P could be a safe range for compensation. Therefore, the adjusting value vector of the chrominance pair components is defined as

$$\boldsymbol{\delta} = \boldsymbol{\eta}(C_b, C_r) \cdot \Delta \boldsymbol{m}_{ps}, \quad (14)$$

where the safety compensation range is determined by

$$|\boldsymbol{\delta}| = \boldsymbol{\eta}(C_b, C_r) \cdot |\Delta \boldsymbol{m}_{ps}|, \quad (15)$$

and the compensation orientation is determined by

$$arctg \angle \boldsymbol{\delta} = \frac{\Delta m_{ps}(C_r)}{\Delta m_{ps}(C_b)}. \quad (16)$$

In fact, Eq. 14 always tries to minimize the chromaticity shift along the opposite directions of the shift. As a result, it has no constraint to illumination change and shows robustness for skin detection.

ADAPTIVE ALGORITHM PROCEDURE

Based on the Bigaussian skin color model, the adaptive algorithm uses the illumination estimation and compensation techniques discussed above to evaluate the skin detection result automatically. If necessary, the algorithm adjusts the original input image and feeds it back to the Gaussian model adaptive threshold algorithm for fine skin region segmentation, shown in Fig. 3. Therefore, the basic skin segmentation method that our algorithm uses is also the Gaussian model adaptive threshold algorithm. The present adaptive algorithm using the Bigaussian model has four steps and is described in detail as follows.
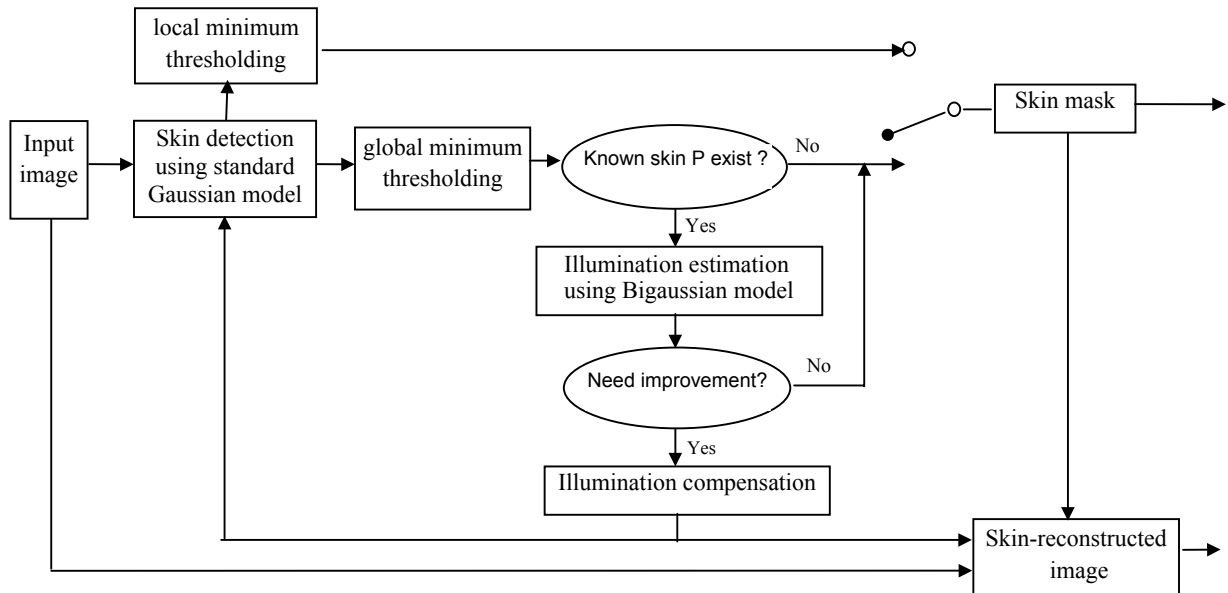

Fig. 3. *Structure of the adaptive algorithm.*

1. For the input image, we use Eq. 17 to get the likelihood value of each pixel belonging to the standard skin cluster

$$P(i, j) = P(C_b, C_r)$$
$$= \exp[-0.5(\boldsymbol{x} - \boldsymbol{m})^T \boldsymbol{C}^{-1} (\boldsymbol{x} - \boldsymbol{m})] , \qquad (17)$$

where $\boldsymbol{x} = (C_b, C_r)^T$ is the chrominance vector.
So we transform the original image to a gray scale image. Each value of a pixel in this gray scale image is the corresponding likelihood value of the original image pixel. By using the adaptive threshold technique, we get a binary image clearly indicating the skin region and non-skin region.

2. We apply the binary image to the original image and obtain the skin regions of the original image. If the total area of the detected skin regions is up to 0.5% of the whole image, the detected skin regions can be safely treated as reliable ground true skin in the image, and the mean values of the skin regions can be treated as a known reliable skin pixel P. Otherwise, the known skin pixel P does not exist; the algorithm considers the skin detection result in the first step as good enough and outputs the binary image as the final skin mask. If the known skin pixel P exists, we estimate the illumination conditions of the input image by measuring the position of P in the standard skin cluster region on the Cb-Cr chrominance plane. No matter which one of Eq. 9 or 10 is satisfied, the skin detection result in the first step needs improvement.

3. If Eq. 9 or Eq. 10 is satisfied, we compensate the illumination change using Eq. 14 and apply the adjusting value to the input image. Then we feed the adjusted image back to step (1) in order to segment skin regions again.

4. If the adjusted image needs to be output, the luminance component must be adjusted, too. Otherwise, this step is bypassed. The binary skin mask output in step (3) is applied to the adjusted image and the original image. Thus, we get a skin-reconstructed image.

The adaptive procedure of the algorithm is illustrated in Fig. 4. In Fig. 4a, a highlight exists in the face and arms as the girl is under bright lighting conditions. Therefore, the skin-likelihood values measured by the Gaussian model, in Fig. 4b, are too small to safely accept the corresponding pixels as skin. When using the adaptive thresholding technique, the result is not good in Fig. 4c. By automatic evaluation, the detected skin regions are treated as a reliable known skin pixel P in Fig. 4h, and its position in the Cb-Cr plane is measured in Fig. 4i. P (up triangle) is outside the red circle that indicates the decision boundary we get from the evaluation criteria Eq. 9. Therefore, Eq. 9 is satisfied and the whole image is compensated. We see that the skin-likelihood values in Fig. 4e are enhanced, and the skin detection results in Fig. 4f achieve a notable improvement.

From Fig. 3, we clearly see that the algorithm adapts to different lighting conditions by feeding the illumination variation from the known standard illumination conditions back to the basic segmentation algorithm. Since the compensation always tries to minimize the chromaticity shift due to the illumination change, the algorithm adapts to different manners of variation in illumination conditions.

For live video, the skin detection results of the current frame can be used to estimate current illumination conditions, and the compensation and skin segmentation can be done on the next frame. Therefore, the feedback becomes feedforward.
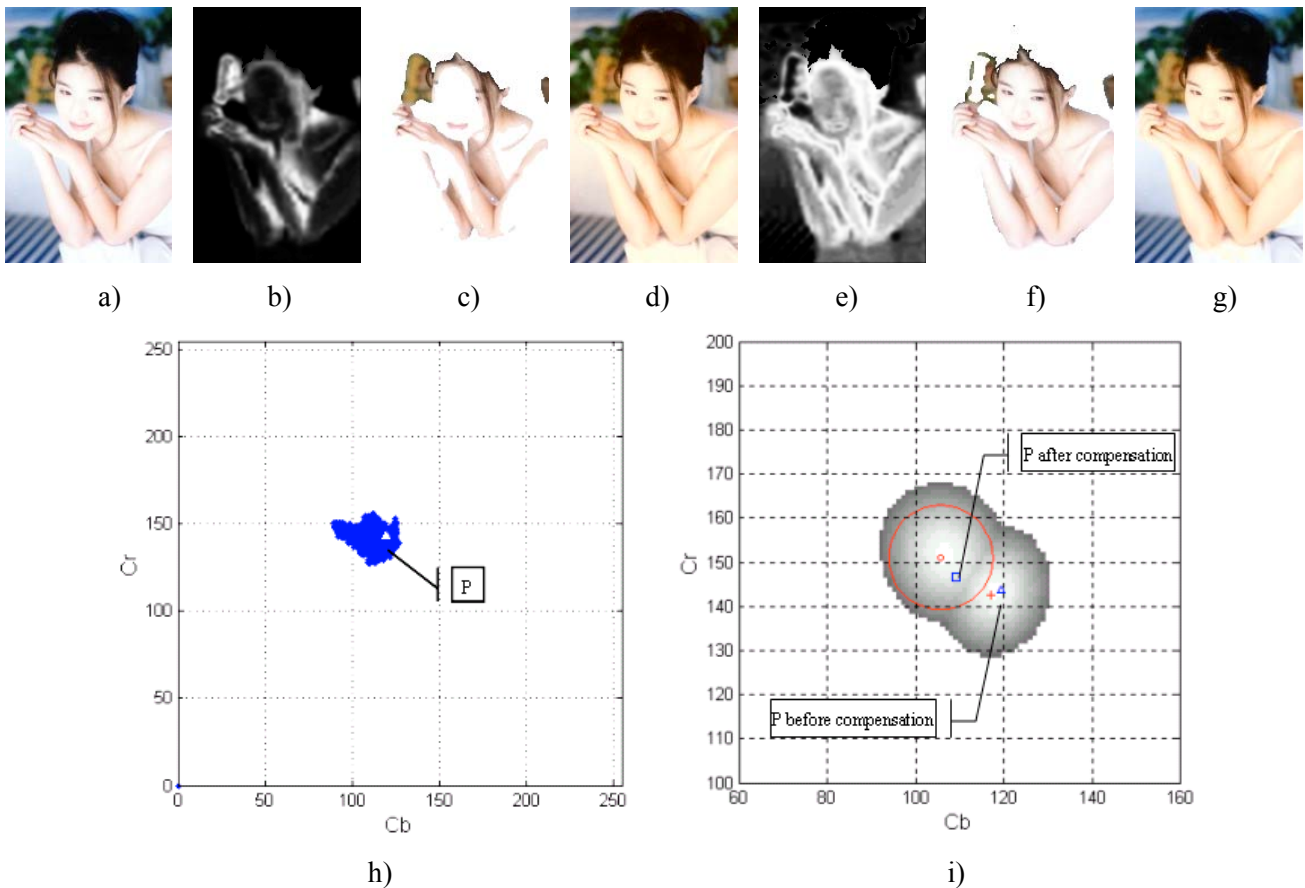
a) b) c) d) e) f) g)

h) i)

Fig. 4. *Illustration of the adaptive procedure. a) Original image. b) Gray scale image that indicates the skin likelihood of original image. c) Skin detection result obtained from Fig. 4b using standard Gaussian model and global minimum thresholding technique. d) Illumination compensated image using Bigaussian Model. e) Skin likelihood of Fig. 4d. f) Final skin detection result obtained from Fig. 4e using standard Gaussian model and Local minimum thresholding technique. g) Skin-reconstructed image. h) The detected skin regions in Fig. 4c are projected on chrominance plane Cb-Cr and the mean values can be treated as a known skin pixel P. i) Illumination estimation and compensation using Bigaussian model. The model with part upon the global minimum threshold we used in Fig. 4c is projected on Cb-Cr plane. The large circle indicates the decision boundary we get from the evaluation criteria Eq. 9.*

THRESHOLDING IN THE ALGORITHM

Since different people with different skin colors have different likelihoods, a fixed threshold value cannot be found. An adaptive thresholding process is required to achieve the optimal threshold value T_{OPT} for each run. Fig. 5 represents a typical ROC (Receiver Operating Characteristics) for skin color detection for different thresholds. In Fig. 5, T_H and T_L determine the threshold search range, in which image pixels with skin-likelihood higher than T_H are classified as skin, while pixels with skin-likelihood lower than T_L are classified as non-skin.
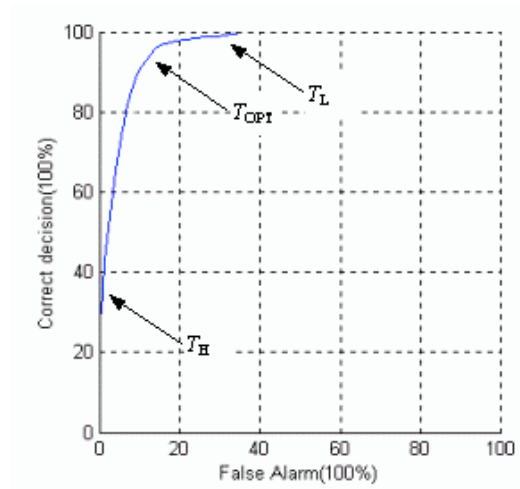

Fig. 5. *ROC (Receiver Operating Characteristics) for skin color detection for different thresholds.*

If we step the threshold value down from T_H to T_L, the segmented regions increase. As a result, the correct skin decision increases sharply, while the skin false alarm gradually increases so that other non-skin regions get included. However, the increase change in the segmented region will gradually decrease until the correct skin decision reaches a high value. If the threshold value is too small, the segmented regions have included many non-skin regions and the area of segmented regions increases sharply, consequently the skin false alarm increases sharply, too. Thus the adaptive thresholding procedure aims at searching for the optimal threshold value T_{OPT} shown in Fig. 5, at which the correct skin decision is high while the false alarm is kept at a small value. In practice, the ground truth skin for a test image is unknown. Therefore, we use the increasing change of the segmented region area to estimate the optimal threshold.
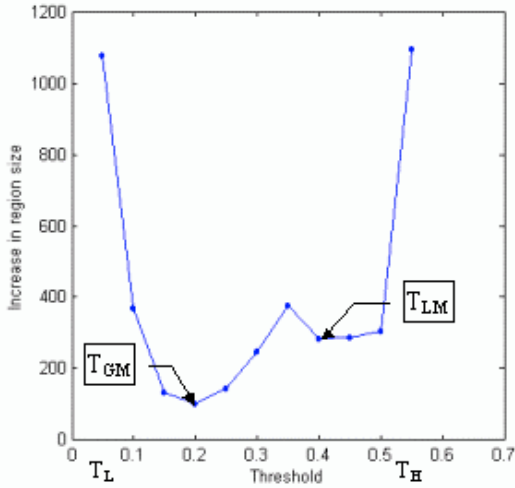

Fig. 6. *Example of two thresholding strategies.*

In our algorithm, shown in Fig. 3, two thresholding strategies are used in different segmentation steps. The global minimum threshold T_{GM} is defined as the threshold value at which the minimum increase in segmented region size is observed while stepping down the threshold value. The local minimum threshold T_{LM} is defined as the threshold value at which a local minimum of the segmented region size is first observed while stepping down the threshold value from T_H to T_L. Evidently, T_{LM} is always higher than or equal to T_{GM}. Fig. 6 illustrates the threshold searching results under the two strategies. In our algorithm, the image is segmented using the global minimum threshold T_{GM} in the first step. If the illumination estimation shows that the skin-segmented result needs improvement, the adjusted image is segmented using the local minimum threshold T_{LM}. The reason is that the chromaticity

shift has been compensated for in the adjusted image, and thus a more restricted thresholding strategy must be used so that we can safely accept skin-like color pixels as skin.

EXPERIMENTAL RESULTS

To compare skin classification performance, the standard skin model in our algorithm and the Gaussian model using the adaptive threshold algorithm use the same skin model parameters. The test images are collected from the Web, with some from the ECU face database (Phung, 2002). These images are divided into two test sets: 120 images for SET-A and 60 images for SET-B. Any one of the images in the two test sets would not be used in obtaining the skin color model. SET-A is used as a normal image test set, assuming that each image in it is obtained under normal lighting conditions. On the contrary, SET-B is used as a special image test set, in which a distinct variation in illumination is observed in the skin regions of each image, *e.g.*, a large highlight or dark regions in the skin, or skin under bright or dark illumination, or skin under colored lighting conditions. In order to obtain ground truth for a performance evaluation, the skin regions in each image are marked by hand.

In the following comparison, four different metrics (Zarit *et al.*, 1999) are used to evaluate the results of the skin detection algorithms. SE (skin error, standing for skin false rejection) is the number of skin pixels identified as non-skin, divided by the number of image pixels. NSE (non-skin error, standing for skin false alarm) is the number of non-skin pixels identified as skin, divided by the number of image pixels. S (percent of skin correct standing for a correct skin decision) is the proportion of all skin pixels identified correctly. C (percent correct) is the proportion of all image pixels (both skin and non-skin) identified correctly. Considering the correlativity of the four metrics, three derived metrics are defined as follows:

$$M_E = (SE^2 + NSE^2)^{1/2},$$

$$M_S = (SE^2 + NSE^2 + (1-S)^2)^{1/2}, \quad (18)$$

$$M_C = (SE^2 + NSE^2 + (1-S)^2 + (1-C)^2)^{1/2},$$

where M_E checks both kinds of error, while M_S and M_C evaluate the skin detection results as a whole.

Thus, we have seven metric values for each image and use the mean metric values of all test

images to indicate the skin detection performance on the test image set. For accurate evaluation, no additional processing step is used to remove skin noise due to segmentation.

Table 1 shows the performance of the two algorithms on the normal image test set, SET-A. We clearly see that the Gaussian model using the adaptive thresholding technique achieves good performance on SET-A. Both errors (SE and NSE) remain at a low level, and the correct skin decision (S) is 88.12%. On SET-A, the error level M_E and correct decision S of our algorithm increase little. This implies that the skin detection result of our algorithm is somewhat like an over-segmentation in comparison with the Gaussian model. However, false alarms of skin color can be removed by later stages of face detection. Therefore, as a whole, our algorithm achieves equivalent performance (see Ms and Mc) to the Gaussian model. In fact, for most of the test images in SET-A, our algorithm in Fig. 3 accepts the segmentation result of the Gaussian model.

On the special image test set, SET-B, the performance of the Gaussian model is unsatisfactory, even when using the adaptive thresholding technique. Due to the illumination problems, skin correct (S) is much lower than that under normal conditions, and the level of skin errors (SE) is high. This indicates that the improvement of the thresholding technique adaptive to illumination change is minor. In contrast to this behavior, our algorithm works well and can reach skin classification performance similar to that under normal conditions. The improvement of skin detection performance is very noticeable; the correct skin decision (S) increases twice, while the error level M_E keeps a value equivalent to the Gaussian model. Fig. 7 compares the metric MS of the two algorithms on each image in SET-B. We clearly see that for most of the test images, the improvement is notable. But for some images, MS values are equal to the Gaussian model, because the algorithm considers the segmentation result of the Gaussian model as a good enough and accepts it, or that the colors of skin in the image are very different from the normal ones and the algorithm accepts the skin segmentation result of Gaussian model for safe detection reasons. Fig. 7 also shows that the simplified evaluation criteria Eq. 9 and Eq. 10 are effective in estimating the variation in illumination.

Table 1. *Skin detection performance on SET-A.*

	$\overline{C}$	$\overline{SE}$	$\overline{NSE}$	$\overline{S}$	$\overline{M}_E$	$\overline{M}_S$	$\overline{M}_C$
Gaussian model with adaptive thresholding	0.9120	0.0287	0.0594	0.8812	0.0775	0.1605	0.1884
Our algorithm	0.8984	0.0219	0.0797	0.9078	0.0924	0.1546	0.1905

Table 2. *Skin detection performance on SET-B.*

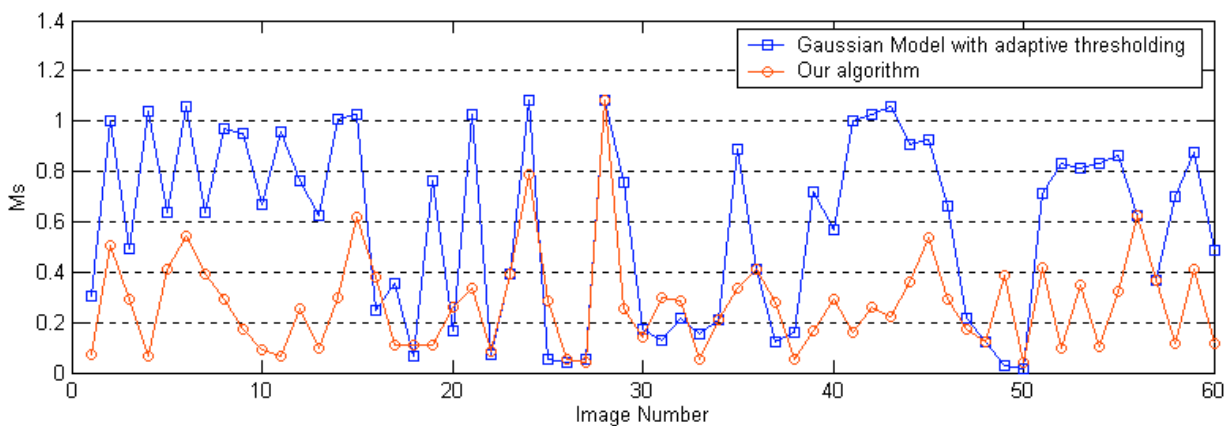|	$\overline{C}$	$\overline{SE}$	$\overline{NSE}$	$\overline{S}$	$\overline{M}_E$	$\overline{M}_S$	$\overline{M}_C$
Gaussian model with adaptive thresholding	0.7892	0.1700	0.0408	0.4618	0.1933	0.5851	0.6277
Our algorithm	0.8121	0.0590	0.1290	0.8265	0.1675	0.2741	0.3390

Fig. 7. *Skin detection performance (Ms) on SET-B.*

Fig. 8 illustrates some examples of test images in SET-B and compares the skin detection results of the two algorithms. We clearly see that there exists a wide variation in illumination conditions in the images of SET-B. Skin tones under these inferior conditions are very different from those achieved under normal conditions. Therefore, skin-likelihood values measured by the Gaussian model are small (the second column in Fig. 8). Evidently, adaptive thresholding helps little in these cases. After illumination compensation, skin tones have been reconstructed, and thus the skin likelihood values are enhanced (the fourth column in Fig. 8). With the help of the adaptive thresholding, the skin detection results of our algorithm are satisfactory. However, the detection results of our algorithm in Fig. 8b and Fig. 8e fail to locate some skin regions. As skin tones in these regions are unreliable, the algorithm has to reject them for safe detection. Additionally, the skin-reconstructed images (the last column in Fig. 8) represent a good result, and it could be a solution to color constancy based on skin tone cues.
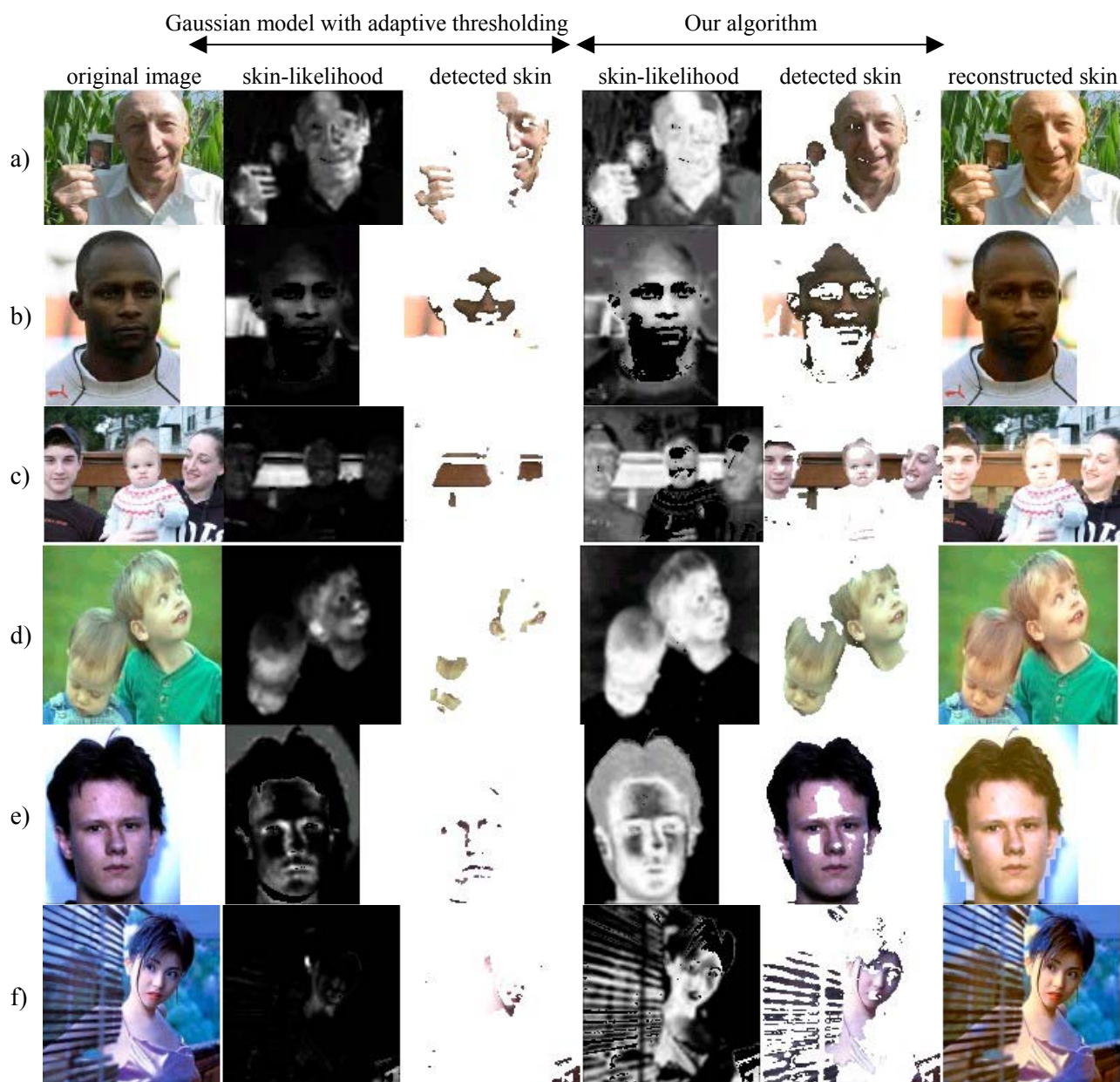

Fig. 8. *Examples of images in SET-B and comparison of the skin detection results. a) Face with shadow and highlight regions. b) Face under dark illumination. c) Faces under strong sunlight. d) Faces appear greenish. e) Face appears pinkish. f) Face appears bluish violet.*

CONCLUSION

For images with a wide variation in the illumination conditions, the improvement afforded by the adaptive thresholding technique is minor, especially for images obtained in a dynamically changing environment. Bigaussian skin detection is suitable for coping with a complex illumination change. Based on this model, an adaptive skin detection algorithm is presented. It automatically evaluates the detected skin result by estimating the difference between the illumination conditions of the image and the normal one we derived for the skin model. The variation in illumination is compensated, and the compensated image is fed back to the Gaussian model for finer skin detection. The experiment shows that our algorithm achieves a noticeable performance improvement in comparison with the adaptive threshold Gaussian model algorithm, and offers a robust solution for skin detection under varying illumination.

ACKNOWLEDGEMENTS

REFERENCE

Phung SL, Chai D, Bouzerdoum A (2002). A novel skin color model in YCbCr color space and its application to human face detection. In: Proceedings of IEEE International Conference on Image Processing, September 22-25, Rochester, U.S.A.

Phung SL (2002). ECU face detection database. Edith Cowan University, School of Engineering and Mathematics, Available at: http:// www. soem.ecu.edu.au/~sphung/face_detection/database/.

Raja Y, McKenna SJ, Gong S (1998). Tracking and segmenting people in varying lighting conditions using colour. In: Proceedings of Face and Gesture, Nara, Japan, 228-33.

Soriano M, Martinkauppi B, Huovinen S, Laaksonen M (2000). Using the skin locus to cope with changing illumination conditions in color-based face tracking. In: Proceedings of IEEE Nordic Signal Processing Symposium NORSIG2000, June 13-15, Kolmarden, Sweden, 383-6.

Soriano M, Martinkauppi B, Huovinen S, Laaksonen M (2003). Adaptive skin color modeling using the skin locus for selecting training pixels. Pattern Recognit 36(3):681-90.

Sahbi H, Boujemaa N (2001). Accurate face detection based on coarse segmentation and fine skin color adaption. In: ICISP, Agadir- Morocco.

Tang JS, Kawato S, Ohya J (2000). Face detection from a complex background. Research Report, ATR Media Integration & Communications Research Laboratories, Kyoto, Japan.

Wong KW, Lam KM, Siu WC (2003). A robust scheme for live detection of human faces in color images. Signal Processing: Image Communication 18(1):103-14.

Yang MH, Ahuja N (1999). Gaussian mixture model for human skin color and its applications in image and video databases. In: Proceedings of SPIE Conference on Storage and Retrieval for Image and Video Databases, January, San Jose.

Zarit BD, Super BJ, Quek FKH (1999). Comparison of five color models in skin pixel classification. In: Proceedings of Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, Corfu, Greece: 58-63.