

## Facial Feature Location and Pose Estimation for Color Face Image based on an Unsupervised Sphere Model and Fused Features

Wei Li

China West Normal University  
nos036@163.com

### Abstract

*For pose-varied color face image, this paper proposed a method of facial feature location and pose estimation based on an unsupervised sphere skin model and fused facial features. Firstly, an adaptive preprocessing, an established unsupervised sphere model and the holes filling technology were presented to extract the face. Then the difference feature and different local binary pattern textural feature of human face were fused to used in the hybrid integral projection technology for localizing the facial features. Finally, poses were estimated using the geometrical distribution of facial features. Experimental results show that, the method can locate facial features and classify different pose more effectively and adaptively.*

**Keywords:** *different local binary pattern, hybrid integral projection, face recognition, facial feature*

### 1. Introduction

Face recognition (FR) has been a very popular research topic in recent years. In the specific research focuses about FR, some problems such as preprocessing technology, skin model, feature location, and pose estimation are still worth studying. Many related methods had been presented in the past. For example, face detection or recognition based on skin models [1-6], facial feature extraction based on PCA (Principal Component Analysis) and LBP (Local Binary Patterns) [7], face detection based on template matching [3, 9], face detection based on Adaboost classifier [10]. In these methods, A. R. Nusirwan et al. presented a regional model with many fixed threshold in multiple color space to detect skin [1]. F. Y. Shih, *et al.*, used the Single Gaussian skin model to realize the multi-face detection [2]. Z. W. Wang, *et al.*, [3] or P. L. M. Bouttefroy, *et al.*, [4] used Gaussian mixture model to segment skin or subtract background. R. L. Hsu, *et al.*, [5] and D. Zhang, *et al.*, [6] used ellipse skin model to detect skin target. Y. Luo, *et al.*, presented a hybrid method by combined using PCA and LBP to extract facial expression feature [7]. S. Tripathi used combined skin color detector and template matching method to get skin and face objects [9]. M. Yang et al. used Adaboost-based face detection for real-time embedded systems [10]. Each of these methods has its advantages and disadvantages. Skin models had different robustness, adaptability, and applicable objects or scopes. Their parameters' estimation may be not flexible or unsupervised (fix threshold or statistical values). The feature value extraction of PCA and template matching may increase the amount of data for processing and time-consuming. Adaboost classifiers detects front face and eyes fast, which can be used for real-time systems, but it's effect greatly depended on training samples and test images. It also may spend much time on training and learning. It was only suitable for front face or small rotated face, but can't discriminate big rotated profile poses. To overcome these limitations above, this study proposed a new fast method of facial feature location and pose estimation. It mainly contains

a new preprocessing technology, an unsupervised sphere model built for face extraction, the proposed fused features containing difference and DLBP (different local binary pattern) textural features for facial feature location, the hybrid integral projection technology based on fused features, as well as pose estimation based on geometrical distribution of facial features.

## 2. Overall Design Scheme

Figure 1 described the overall design scheme. Figure 2 showed the images of scheme. As shown in Figure 1 and 2, the proposed method mainly involves four parts which are original face image collection, face extraction, facial feature location, and pose estimation. After collecting original face image, the face extraction and facial feature location are the most important parts in the scheme, which are the basis of the pose estimation. The face extraction contained preprocessing, establishing skin model and segmenting facial skin regions, as well as removing noise and filling holes. The facial feature location contained difference and calculating DLBP textural features, fusing the difference feature and textural feature, as well as facial feature location using hybrid integral projection. Based on the results of facial feature location, the pose estimation can be finish by using the geometrical distribution of facial features.

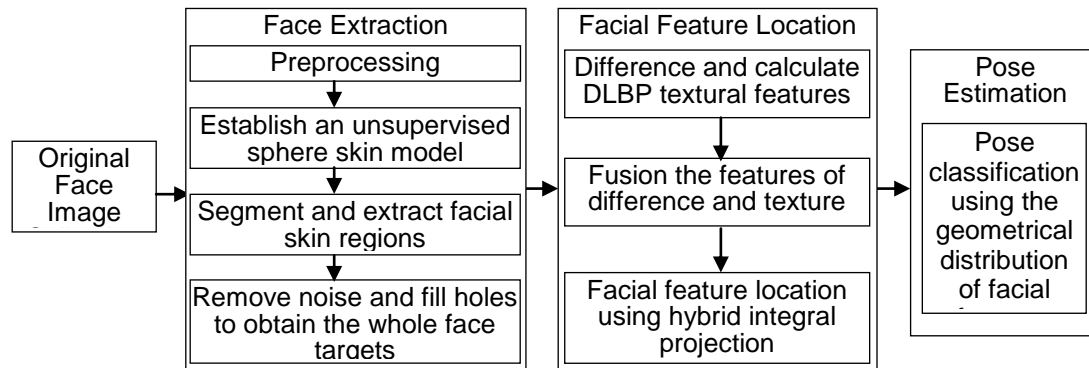


Figure 1. Description of Overall Design Scheme

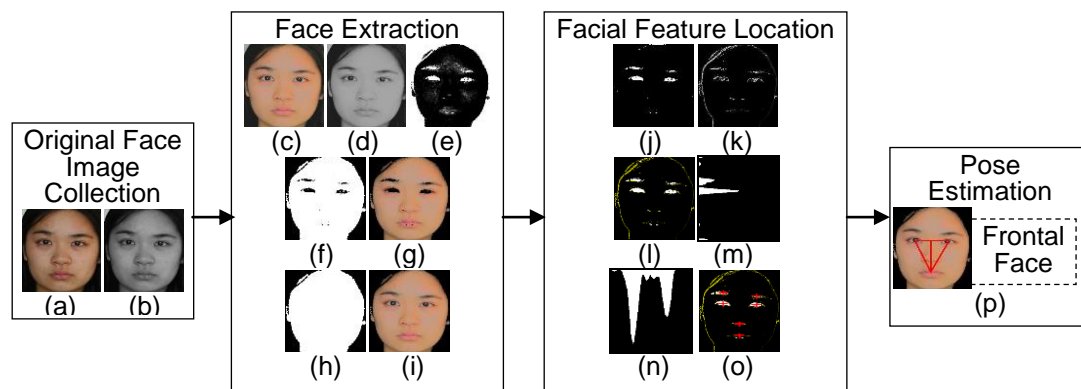


Figure 2. Images of Overall Design Scheme

(a) Original color face image, (b) Gray image of (a), (c) Preprocessed image of (a), (d) Gray image of (c), (e) Facial skin-probability face image, (f) Facial skin-segmented image, (g) Facial skin-extracted image, (h) Face-segmented image, (i) Face-extracted image, (j) Facial

feature image, (k) DLBP textural feature image, (l) Feature-fused image, (m) Horizontal hybrid projection of (l), (n) Vertical hybrid projection of (l), (o) Feature-localized image, (p) Pose-discriminated image.

### 3. Face Extraction

There are many factors which can affect the accuracy and instantaneity of face extraction, such as quality of original color face image, the established skin model, and so on. So before face extraction, the preprocessing of input face image is required and useful.

#### 3.1. Preprocessing

For original input color face image, some elements (*e. g.*, capture device, illumination, and expression) may change the brightness, contrast, facial appearance, and information of facial feature. In order to decrease these influence above, a new adaptive preprocessing is proposed, which consisted of color adjustment, brightness adjustment, color normalization, and color smoothness. Figure 3 showed the flow of preprocessing.

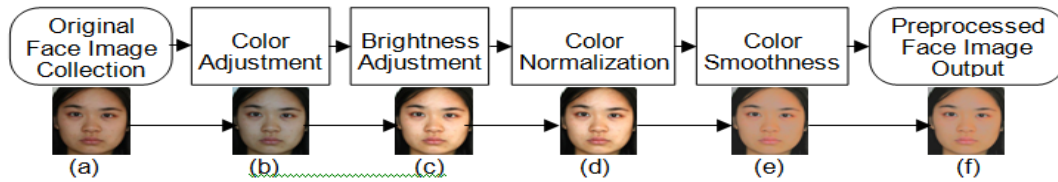


Figure 3. Flow of Preprocessing

(a) Original color face image, (b) Color-adjusted face image, (c) Brightness-adjusted face image, (d) Color-normalized face image, (e) Color-smoothed face image, (f) Preprocessed face image.

**3.1.1. Color Adjustment:** The purpose of color adjustment is to depress color offset caused by different color light source. The process of color adjustment contains the following steps:

Step 1. Transform R, G, and B values to Y, Cr, Cg, and Cb values by formula (1).

$$\begin{cases} Y_{i,j} = 0.299 \times R_{i,j} + 0.578 \times G_{i,j} + 0.114 \times B_{i,j} \\ Cr_{i,j} = 0.713 \times (R_{i,j} - Y_{i,j}) + 128 \\ Cg_{i,j} = 0.587 \times (G_{i,j} - Y_{i,j}) + 128 \\ Cb_{i,j} = 0.564 \times (B_{i,j} - Y_{i,j}) + 128 \end{cases} \quad (1)$$

Where,  $i$  and  $j$  are a point's abscissa and ordinate of image,  $R_{i,j}$ ,  $G_{i,j}$ ,  $B_{i,j}$  are values of a pixel's R, G, B components in input face image (Figure 3(a)).

Step 2. Calculate color adjustment factors ( $fR_1$ ,  $fG_1$ ,  $fB_1$ ) using respective accumulations of Cr, Cg, and Cb by formula (2).

$$\begin{cases} fR_1 = \left[ \frac{\sum_{j=0}^H \sum_{i=0}^W Cr_{i,j} + \sum_{j=0}^H \sum_{i=0}^W Cg_{i,j} + \sum_{j=0}^H \sum_{i=0}^W Cb_{i,j}}{3} \right] / \sum_{j=0}^H \sum_{i=0}^W Cr_{i,j} \\ fG_1 = \left[ \frac{\sum_{j=0}^H \sum_{i=0}^W Cr_{i,j} + \sum_{j=0}^H \sum_{i=0}^W Cg_{i,j} + \sum_{j=0}^H \sum_{i=0}^W Cb_{i,j}}{3} \right] / \sum_{j=0}^H \sum_{i=0}^W Cg_{i,j} \\ fB_1 = \left[ \frac{\sum_{j=0}^H \sum_{i=0}^W Cr_{i,j} + \sum_{j=0}^H \sum_{i=0}^W Cg_{i,j} + \sum_{j=0}^H \sum_{i=0}^W Cb_{i,j}}{3} \right] / \sum_{j=0}^H \sum_{i=0}^W Cb_{i,j} \end{cases} \quad (2)$$

Where, H and W are the height and width of image.

Step 3. Adjust the R, G, and B values of input face image using color adjustment factors ( $fR_1$ ,  $fG_1$ ,  $fB_1$ ) by formula (3).

$$R_{i,j}^{(1)} = \begin{cases} 255 & \text{if } (fR_1 \times R_{i,j} \geq 255) \\ fR_1 \times R_{i,j} & \text{else} \end{cases}, G_{i,j}^{(1)} = \begin{cases} 255 & \text{if } (fG_1 \times G_{i,j} \geq 255) \\ fG_1 \times G_{i,j} & \text{else} \end{cases}, B_{i,j}^{(1)} = \begin{cases} 255 & \text{if } (fB_1 \times B_{i,j} \geq 255) \\ fB_1 \times B_{i,j} & \text{else} \end{cases} \quad (3)$$

Where,  $R_{i,j}^{(1)}$ ,  $G_{i,j}^{(1)}$ ,  $B_{i,j}^{(1)}$  are color-adjusted R,G,B values in color-adjusted face image (Figure 3(b)).

**3.1.2. Brightness Adjustment:** The main purpose of brightness adjustment is to depress the low brightness and high brightness caused by abnormal exposure. As described in formula (4)-(6), the process of brightness adjustment contains the following steps:

Step 1. Transform color-adjusted R, G, and B values ( $R_{i,j}^{(1)}$ ,  $G_{i,j}^{(1)}$ ,  $B_{i,j}^{(1)}$ ) to Y, Cr, Cg, and Cb values by formula (4).

$$\begin{cases} Y_{i,j}^{(1)} = 0.299 \times R_{i,j}^{(1)} + 0.578 \times G_{i,j}^{(1)} + 0.114 \times B_{i,j}^{(1)} \\ Cr_{i,j}^{(1)} = 0.713 \times (R_{i,j}^{(1)} - Y_{i,j}^{(1)}) + 128 \\ Cg_{i,j}^{(1)} = 0.587 \times (G_{i,j}^{(1)} - Y_{i,j}^{(1)}) + 128 \\ Cb_{i,j}^{(1)} = 0.564 \times (B_{i,j}^{(1)} - Y_{i,j}^{(1)}) + 128 \end{cases} \quad (4)$$

Step 2. Calculate brightness adjustment factors ( $fR_2$ ,  $fG_2$ ,  $fB_2$ ) using respective accumulations of Cr, Cg, Cb, and Y by formula (5).

$$fR_2 = \frac{\sum_{j=0}^H \sum_{i=0}^W Cr_{i,j}^{(1)}}{\sum_{j=0}^H \sum_{i=0}^W Y_{i,j}^{(1)}}, fG_2 = \frac{\sum_{j=0}^H \sum_{i=0}^W Cg_{i,j}^{(1)}}{\sum_{j=0}^H \sum_{i=0}^W Y_{i,j}^{(1)}}, fB_2 = \frac{\sum_{j=0}^H \sum_{i=0}^W Cb_{i,j}^{(1)}}{\sum_{j=0}^H \sum_{i=0}^W Y_{i,j}^{(1)}} \quad (5)$$

Step 3. Adjust the R, G, and B values of color-adjusted face image using brightness adjustment factors ( $fR_2$ ,  $fG_2$ ,  $fB_2$ ) by formula (6).

$$R_{i,j}^{(2)} = \begin{cases} 255 & \text{if } (fR_2 \times R_{i,j}^{(1)} \geq 255) \\ fR_2 \times R_{i,j}^{(1)} & \text{else} \end{cases}, G_{i,j}^{(2)} = \begin{cases} 255 & \text{if } (fG_2 \times G_{i,j}^{(1)} \geq 255) \\ fG_2 \times G_{i,j}^{(1)} & \text{else} \end{cases}, B_{i,j}^{(2)} = \begin{cases} 255 & \text{if } (fB_2 \times B_{i,j}^{(1)} \geq 255) \\ fB_2 \times B_{i,j}^{(1)} & \text{else} \end{cases} \quad (6)$$

Where,  $R_{i,j}^{(2)}$ ,  $G_{i,j}^{(2)}$ ,  $B_{i,j}^{(2)}$  are brightness-adjusted values of R,G,B in brightness-adjusted face image (Figure 3(c)).

**3.1.3. Color Normalization:** The function of color normalization is to normalize the brightness-adjusted R, G, B values ( $R_{i,j}^{(2)}$ ,  $G_{i,j}^{(2)}$ ,  $B_{i,j}^{(2)}$ ) to the range of [0 255]. The method of normalization is described in formula (7).

$$R_{i,j}^{(3)} = \begin{cases} 255 \times \frac{(R_{i,j}^{(2)} - R_n^{(2)})}{(R_m^{(2)} - R_n^{(2)})} & \text{else} \\ 0 & \text{if } (R_m^{(2)} - R_n^{(2)}) \leq 0 \end{cases}, B_{i,j}^{(3)} = \begin{cases} 255 \times \frac{(B_{i,j}^{(2)} - B_n^{(2)})}{(B_m^{(2)} - B_n^{(2)})} & \text{else} \\ 0 & \text{if } (B_m^{(2)} - B_n^{(2)}) \leq 0 \end{cases}, G_{i,j}^{(3)} = \begin{cases} 255 \times \frac{(G_{i,j}^{(2)} - G_n^{(2)})}{(G_m^{(2)} - G_n^{(2)})} & \text{else} \\ 0 & \text{if } (G_m^{(2)} - G_n^{(2)}) \leq 0 \end{cases} \quad (7)$$

Where,  $R_m^{(2)}$ ,  $G_m^{(2)}$ ,  $B_m^{(2)}$  are respective maximums of  $R_{i,j}^{(2)}$ ,  $G_{i,j}^{(2)}$ ,  $B_{i,j}^{(2)}$ ,  $R_n^{(2)}$ ,  $G_n^{(2)}$ ,  $B_n^{(2)}$  are respective minimums of  $R_{i,j}^{(2)}$ ,  $G_{i,j}^{(2)}$ ,  $B_{i,j}^{(2)}$ ,  $R_{i,j}^{(3)}$ ,  $G_{i,j}^{(3)}$ ,  $B_{i,j}^{(3)}$  are color-normalized R,G,B values in color-normalized face image (Figure 3(d)).

**3.1.4. Color Smoothness:** The main purpose of color smoothness is to smooth or equal the color and light, as well as reduce interference noises caused by unequal color and light. In order to reach this purpose, the TSL and YCbCr color space were jointly considered. On one

hand, by statistical analysis in color spaces, the area percentage of skin color in TS 2D color plane (unrelated to brightness) is smaller than that of the other color plane (*e.g.*, RG, RB, GB, TL, SL, HS). That means the skin cluster is more concentrate or tightness in TS color space, and T and S components have favorable description ability of skin color. On the other hand, the Y component is relatively independent and separated from Cb and Cr components in YCbCr color space. That means the Y can be processed in YCbCr color space without affecting Cb and Cr components, but the adjusted Y component maybe affect the transformed R, G, B color components. So in this part, the rough idea of color smoothness is to adjust the Y values and limit the range of Y component in YCbCr color space by using the relationship of improved T', S', and L components. This way combined information of multiple components in two color space, so R, G, B components can be smoothed or equaled indirectly.

As described in formula (8-12), the process of color smoothness contains these steps:

Step 1. Transform color-normalized  $R$ ,  $G$ , and  $B$  ( $R_{i,j}^{(3)}, G_{i,j}^{(3)}, B_{i,j}^{(3)}$ ) to  $Y$ ,  $Cr$ , and  $Cb$  by formula (8).

$$\begin{cases} Y_{i,j}^{(2)} = 0.299 \times R_{i,j}^{(3)} + 0.578 \times G_{i,j}^{(3)} + 0.114 \times B_{i,j}^{(3)} \\ Cr_{i,j}^{(2)} = 0.713 \times (R_{i,j}^{(3)} - Y_{i,j}^{(2)}) + 128 \\ Cg_{i,j}^{(2)} = 0.587 \times (G_{i,j}^{(3)} - Y_{i,j}^{(2)}) + 128 \\ Cb_{i,j}^{(2)} = 0.564 \times (B_{i,j}^{(3)} - Y_{i,j}^{(2)}) + 128 \end{cases} \quad (8)$$

Step 2. Transform  $R_{i,j}^{(3)}, G_{i,j}^{(3)}, B_{i,j}^{(3)}$  to  $T'$ ,  $S'$  and  $L$  values according to  $T'S'L$  conversion formula (10), which is improved from traditional TSL conversion formula (9).

$$\begin{aligned} r_{i,j} &= \frac{R_{i,j}^{(3)}}{(R_{i,j}^{(3)} + G_{i,j}^{(3)} + B_{i,j}^{(3)})} - \frac{1}{3}, \quad g_{i,j} = \frac{G_{i,j}^{(3)}}{(R_{i,j}^{(3)} + G_{i,j}^{(3)} + B_{i,j}^{(3)})} - \frac{1}{3}, \quad b_{i,j} = \frac{B_{i,j}^{(3)}}{(R_{i,j}^{(3)} + G_{i,j}^{(3)} + B_{i,j}^{(3)})} - \frac{1}{3}, \\ T_{i,j} &= \begin{cases} 0, & \text{if } g_{i,j} = 0 \\ \arctan(r_{i,j}/g_{i,j}) + 1/4, & \text{if } g_{i,j} > 0, \\ \arctan(r_{i,j}/g_{i,j}) + 3/4, & \text{if } g_{i,j} < 0 \end{cases} \quad (9) \\ S_{i,j} &= \sqrt{1.8 \times (r_{i,j}^2 + g_{i,j}^2)}, \quad L_{i,j} = 0.299 \times R_{i,j}^{(3)} + 0.587 \times G_{i,j}^{(3)} + 0.114 \times B_{i,j}^{(3)} \end{aligned}$$

$$\begin{aligned} r_{i,j} &= R_{i,j}^{(3)} / (R_{i,j}^{(3)} + G_{i,j}^{(3)} + B_{i,j}^{(3)}), \quad g_{i,j} = G_{i,j}^{(3)} / (R_{i,j}^{(3)} + G_{i,j}^{(3)} + B_{i,j}^{(3)}), \quad b_{i,j} = B_{i,j}^{(3)} / (R_{i,j}^{(3)} + G_{i,j}^{(3)} + B_{i,j}^{(3)}) \\ T_{i,j}^{(1)} &= \frac{\left( \arctan(r_{i,j}/g_{i,j}) + \arctan(r_{i,j}/b_{i,j}) + \arctan(g_{i,j}/r_{i,j}) + \arctan(g_{i,j}/b_{i,j}) + \arctan(b_{i,j}/r_{i,j}) + \arctan(b_{i,j}/g_{i,j}) \right)}{2\pi}, \quad (10) \\ S_{i,j}^{(1)} &= \sqrt{(r_{i,j}^2 + g_{i,j}^2 + b_{i,j}^2)/3}, \quad L_{i,j} = 0.299 \times R_{i,j}^{(3)} + 0.587 \times G_{i,j}^{(3)} + 0.114 \times B_{i,j}^{(3)} \end{aligned}$$

Step 3. Adjust the value and limit the range of Y component in YCbCr space by formula (11).

$$Y_{i,j}^{(2)} = Y_{i,j}^{(2)} \times e^{[(T_{i,j}^{(1)} + S_{i,j}^{(1)}) - L_{i,j}]/255.1}, \quad Y_{i,j}^{(3)} = \begin{cases} 220 & \text{if } (Y_{i,j}^{(2)} > 220) \\ 30 & \text{if } (Y_{i,j}^{(2)} < 30) \end{cases} \quad (11)$$

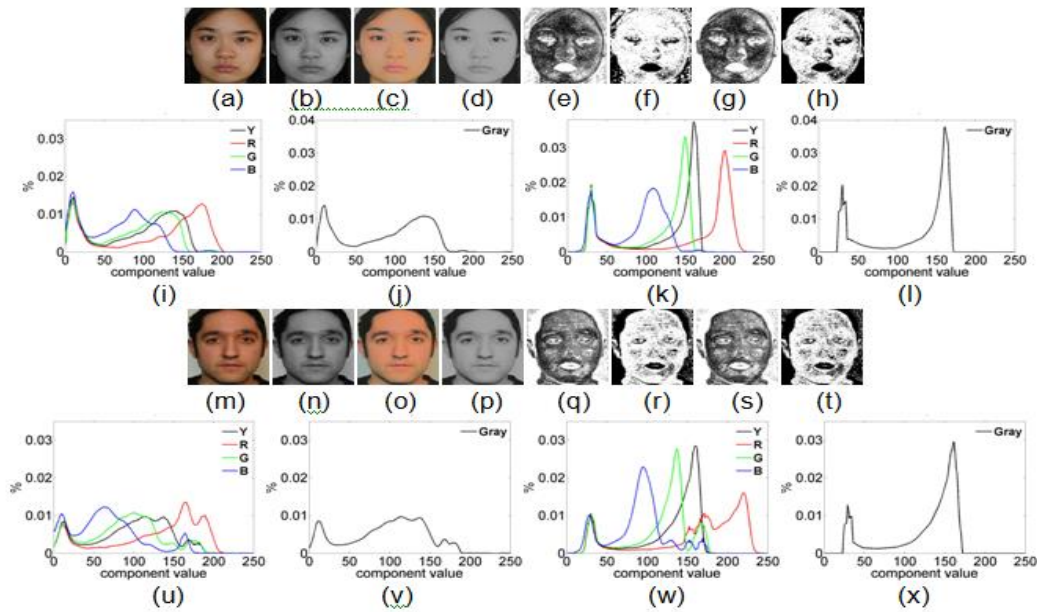
Step 4. Transform Y, Cr, and Cb values to R, G, and B values by formula (12).

$$\begin{aligned} R_{i,j}^{(4)} &= \begin{cases} Y_{i,j}^{(3)} + 1.402 \times Cr_{i,j}^{(2)} - 179.456 & \text{else} \\ 255 & \text{if } ((Y_{i,j}^{(3)} + 1.402 \times Cr_{i,j}^{(2)} - 179.456) > 255) \end{cases} \\ B_{i,j}^{(4)} &= \begin{cases} Y_{i,j}^{(3)} + 1.772 \times Cb_{i,j}^{(2)} - 226.816 & \text{else} \\ 255 & \text{if } ((Y_{i,j}^{(3)} + 1.772 \times Cb_{i,j}^{(2)} - 226.816) > 255) \end{cases} \quad (12) \\ G_{i,j}^{(4)} &= \begin{cases} Y_{i,j}^{(3)} - 0.3441 \times Cb_{i,j}^{(2)} - 0.71414 \times Cr_{i,j}^{(2)} + 135.4598 & \text{else} \\ 255 & \text{if } ((Y_{i,j}^{(3)} - 0.3441 \times Cb_{i,j}^{(2)} - 0.71414 \times Cr_{i,j}^{(2)} + 135.4598) > 255) \end{cases} \end{aligned}$$

Where,  $R_{i,j}^{(4)}$ ,  $G_{i,j}^{(4)}$ ,  $B_{i,j}^{(4)}$  are color-smoothed  $R, G, B$  values in color-smoothed face image (Figure 3(e)). The color-smoothed image is regarded as output preprocessed face image (Figure 3(f)).

**3.1.5. Analysis of Preprocessing:** By comparison of Figure 3(b) and 3(a), it's easy to find that some color offset had been reduced. By comparison of Figure 3(c), 3(d) and 3(b), it can be seen that the low brightness had been depressed. By comparison of Figure 3(e) and 3(d), the smoothed face image had more smoothed and balance color.

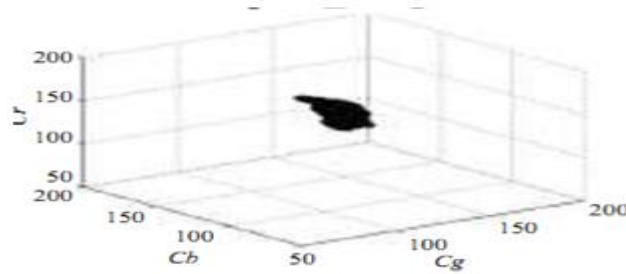
Figure 4 shows two examples of the comparison of original face image and preprocessed face image. Comparing Figure 4(a) with 4(c), Figure 4(b) with 4(d), Figure 4(m) with 4(o), or Figure 4(n) with 4(p), it's clear that the color offset, abnormal exposure, unequal color had been reduced. Face image became more smooth and clear because some noise had been removed. Comparing the Figure 4(e) with 4(g), or the Figure 4(q) with 4(s), the skin-probability face image (based on Single Gaussian Model) of preprocessed face image had the more clear description of face region and face contour than that of input face image. Comparing Figure 4(f) with 4(h), or Figure 4(r) with 4(t), the skin-segmented face image of preprocessed face image had better effect of segmentation than that of input face image. By analyzing their corresponding curves of PDD (probability density distribution) of  $R, G, B, Y$  components and gray value (Comparing Figure 4(i) with 4(k), Figure 4(j) with 4(l), Figure 4(u) with 4(w), or Figure 4(v) with 4(x)), it is also can be seen that, two peaks of every component ( $R, G, B, Y$ , Gray value) in preprocessed face image are more sharp-pointed and obvious than those in original face image. That means preprocessing can improve quality of input face image, and help for distinguishing the face and backgrounds better and easier.



**Figure 4. Comparison of Original Face Images and Preprocessed Face Images**

(a) Original color face image 1, (b) Gray image of (a), (c) Preprocessed face image of (a), (d) Gray image of (c), (e) Skin-probability face image of (a) based on SGM, (f) Skin-segmented face image of (a) based on SGM, (g) Skin-probability face image of (c) based on SGM, (h) Skin-segmented face image of (c) based on SGM, (i) PDDs of  $R, G, B, Y$  components of (a), (j)

PDD of Gray value of (a), (k) PDDs of R,G,B,Y components of (c), (l) PDD of Gray value of (c), (m) Original color face image 2, (n) Gray image of (m), (o) Preprocessed face image of (m), (p) Gray image of (o), (q) Skin-probability face image of (m) based on SGM, (r) Skin-segmented face image of (m) based on SGM, (s) Skin-probability face image of (o) based on SGM, (t) Skin-segmented face image of (o) based on SGM, (u) PDDs of R,G,B,Y components of (m), (v) PDD of Gray value of (m), (w) PDDs of R,G,B,Y components of (o), (x) PDD of Gray value of (o).



**Figure 5. The Face Skin Clustering in CgCbCr Color Space**

### 3.2. Unsupervised Sphere Model

For subsequent extraction of face targets, a robust skin model is a useful way to obtain a good skin-probability image. In existed skin models, the regional model [1] doesn't have good adaptability because it usually adopts some fixed threshold values. The accuracy and adaptability of SGM (Single Gaussian Model) [2] are also not high enough, which lead to decrease the segmentation rate. The learning process and parameter estimation of GMM (Gaussian Mixture Model) [3, 4] are time-consuming, which lead to increase the time and space complexity of the algorithm. The fixed parameters of ellipse model [5, 6] are not suitable for all types of face image, especially for the face image with color and brightness variation. To establish a simple self-adaptive robust model without high algorithm complexity and time complexity, this study proposed an unsupervised sphere skin model. By statistics and analysis of skin samples, it can be found that the face skin clustering in 3D CgCbCr color space represents the shape of sphere approximately. This sphere skin model roughly describes the shape of face skin clustering in CbCgCr color space (formula (13)), as shown in Figure 5. The samples of face skin clustering came from self-built face database, which contained preprocessed face images came from self-built multi-pose face database. To apply to all kinds of harsh conditions, such as color and brightness variation or unequal, the proposed sphere is simple but has good robustness and adaptability. That means the sphere is dynamic and movable because the coordinates (X, Y, Z) of sphere's centre are dynamic real-time estimated unsupervised parameters, as shown in formula (14).

$$\begin{cases} R_{i,j}^{(5)} = \frac{R_{i,j}^{(4)} \times 255}{(R_{i,j}^{(4)} + G_{i,j}^{(4)} + B_{i,j}^{(4)})}, G_{i,j}^{(5)} = \frac{G_{i,j}^{(4)} \times 255}{(R_{i,j}^{(4)} + G_{i,j}^{(4)} + B_{i,j}^{(4)})}, B_{i,j}^{(5)} = \frac{B_{i,j}^{(4)} \times 255}{(R_{i,j}^{(4)} + G_{i,j}^{(4)} + B_{i,j}^{(4)})} \\ Cr_{i,j}^{(3)} = 0.500 \times R_{i,j}^{(5)} - 0.4187 \times G_{i,j}^{(5)} - 0.0813 \times B_{i,j}^{(5)} + 128 \\ Cb_{i,j}^{(3)} = -0.1687 \times R_{i,j}^{(5)} - 0.3313 \times G_{i,j}^{(5)} + 0.500 \times B_{i,j}^{(5)} + 128 \\ Cg_{i,j}^{(3)} = -0.316 \times R_{i,j}^{(5)} + 0.500 \times G_{i,j}^{(5)} - 0.184 \times B_{i,j}^{(5)} + 128 \end{cases} \quad (13)$$

Where,  $R_{i,j}^{(4)}, G_{i,j}^{(4)}, B_{i,j}^{(4)}$  are R,G,B values in preprocessed face image (Figure 3(f)),  $R_{i,j}^{(5)}, G_{i,j}^{(5)}, B_{i,j}^{(5)}$  are normalized R,G,B values.

$$\begin{cases} \bar{C} = (\overline{Cr^{(3)}} + \overline{Cg^{(3)}} + \overline{Cb^{(3)}}) / 3 \\ \overline{cg} = \frac{\overline{Cg^{(3)}}}{(\overline{Cr^{(3)}} + \overline{Cb^{(3)}}) / 2}, \quad \overline{cb} = \frac{\overline{Cb^{(3)}}}{(\overline{Cr^{(3)}} + \overline{Cg^{(3)}}) / 2}, \quad \overline{cr} = \frac{\overline{Cr^{(3)}}}{(\overline{Cb^{(3)}} + \overline{Cg^{(3)}}) / 2} \\ ag = \bar{C} \times \overline{cg}, \quad ab = \bar{C} \times \overline{cb}, \quad ar = \bar{C} \times \overline{cr} \\ I_{i,j}^{(1)} = 0.596 \times R_{i,j}^{(4)} - 0.274 \times G_{i,j}^{(4)} - 0.322 \times B_{i,j}^{(4)} \\ P_{i,j}^{(1)} = \frac{(Cg_{i,j}^{(3)} - ag)^2}{(I_{i,j}^{(1)})^2} + \frac{(Cb_{i,j}^{(3)} - ab)^2}{(I_{i,j}^{(1)})^2} + \frac{(Cr_{i,j}^{(3)} - ar)^2}{(I_{i,j}^{(1)})^2} \end{cases} \quad (14)$$

Where, Sphere's radius is defined by I component, which represented color range of orange to cyan and covered range of skin color. The  $P_{i,j}^{(1)}$  is skin probability value of pixel in preprocessed face image.  $ag$ ,  $ab$ ,  $ar$  are 3D coordinates of sphere's centre, which calculated using the average ( $\bar{C}$ ) of  $\overline{Cr^{(3)}}$ ,  $\overline{Cg^{(3)}}$ ,  $\overline{Cb^{(3)}}$ , and relative coefficients ( $\overline{cg}$ ,  $\overline{cb}$ ,  $\overline{cr}$ ) of dynamic sphere's centre.  $\overline{Cr^{(3)}}$ ,  $\overline{Cg^{(3)}}$ ,  $\overline{Cb^{(3)}}$  are expectations of  $Cr_{i,j}^{(3)}$ ,  $Cg_{i,j}^{(3)}$ ,  $Cb_{i,j}^{(3)}$ , as shown in formula (15).

$$\overline{Cr^{(3)}} = \sum_{j=0}^{h-1} \sum_{i=0}^{w-1} Cr_{i,j}^{(3)} / h \times w, \quad \overline{Cg^{(3)}} = \sum_{j=0}^{h-1} \sum_{i=0}^{w-1} Cg_{i,j}^{(3)} / h \times w, \quad \overline{Cb^{(3)}} = \sum_{j=0}^{h-1} \sum_{i=0}^{w-1} Cb_{i,j}^{(3)} / h \times w \quad (15)$$

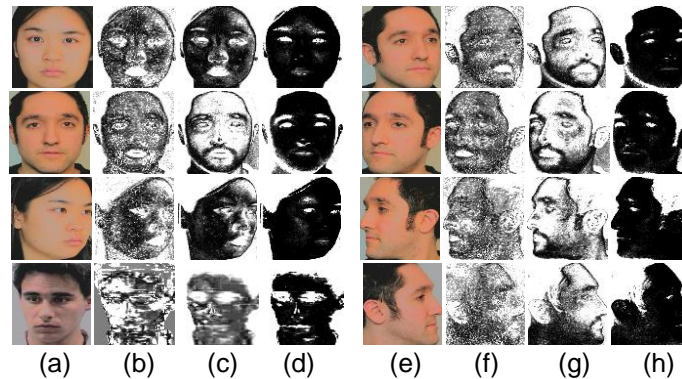
Where,  $h$  and  $w$  are height and width of preprocessed face image. The relative coefficients ( $\overline{cg}$ ,  $\overline{cb}$ ,  $\overline{cr}$ ) of dynamic sphere's centre are ratios of an expectation and the average of other expectations. These ratios reflect degree of color offset in some degree, which help to adjust the position of sphere's centre dynamically and automatically.

Based on unsupervised Sphere Model, the skin-probability face image can be acquired according to formula (16).

$$P_{i,j}^{(2)} = \begin{cases} P_{i,j}^{(1)} \times 255 & \text{if } (P_{i,j}^{(1)} \leq 1) \\ 255 & \text{else} \end{cases} \quad (16)$$

Where,  $P_{i,j}^{(2)}$  is a pixel's gray value of skin-probability face image (Figure 6(d) and 6(h)).

Figure 6 show eight comparisons of skin-probability face images based on SGM, CbCr ellipse model, and proposed unsupervised model. By comparing, for preprocessed face images (Figure 6(a) and 6(e)), the skin-probability face images (Figure 6(d) and 6(h)) based on proposed model have less noises and more smoothed face targets than those based on SGM (Figure 6(b) and 6(f)) and ellipse model (Figure 6(c) and 6(g)). That means proposed unsupervised skin model have better ability of description of skin targets.



**Figure 6. Comparisons of Facial Skin-Probability Images based on SGM, Ellipse Model and Proposed Unsupervised Model**

(a) Preprocessed face images in group 1, (b) Facial skin-probability face images of (a) based on SGM, (c) Facial skin-probability images of (a) based on ellipse model, (d) Facial skin-probability images of (a) based on proposed model, (e) Preprocessed face images in group 2, (f) Facial skin-probability images of (e) based on SGM, (g) Facial skin-probability images of (e) based on ellipse model, (h) Facial skin-probability images of (e) based on proposed model.

### 3.3. Obtain the Facial Skin Region

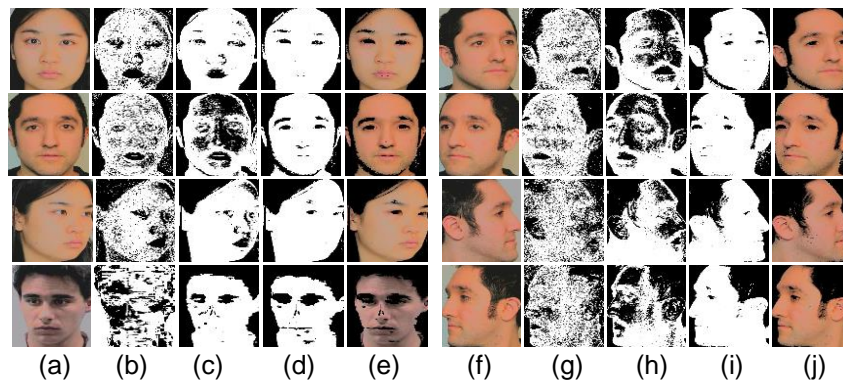
After establishing the unsupervised sphere model for preprocessed face image, the segmentation and extraction of facial skin region become easy. Using the built sphere model, according to formula (17)-(18), the segmentation and extraction of facial skin region can be finished.

$$P_{i,j}^{(3)} = \begin{cases} 255 & \text{if } (P_{i,j}^{(1)} \leq 1) \\ 0 & \text{else} \end{cases} \quad (17)$$

$$R_{i,j}^{(6)} = \begin{cases} R_{i,j}^{(4)} & \text{if } (P_{i,j}^{(1)} \leq 1) \\ 0 & \text{else} \end{cases}, \quad G_{i,j}^{(6)} = \begin{cases} G_{i,j}^{(4)} & \text{if } (P_{i,j}^{(1)} \leq 1) \\ 0 & \text{else} \end{cases}, \quad B_{i,j}^{(6)} = \begin{cases} B_{i,j}^{(4)} & \text{if } (P_{i,j}^{(1)} \leq 1) \\ 0 & \text{else} \end{cases} \quad (18)$$

Where,  $P_{i,j}^{(3)}$  is a pixel's gray value of facial skin-segmented image (Figure 7(d) and 7(i)).  $R_{i,j}^{(6)}, G_{i,j}^{(6)}, B_{i,j}^{(6)}$  are  $R, G, B$  values in facial skin-extracted image (Figure 7(e) and 7(j)).

Figure 7 show four comparisons of facial skin-segmented images based on SGM, CbCr ellipse model, and the proposed unsupervised model. By comparing, for preprocessed face images (Figure 7(a) and 7(f)), the facial skin-segmented images (Figure 7(d) and 7(i)) based on proposed model have less incorrect segmentation and more correct segmentation than those based on SGM (Figure 7(b) and 7(g)) and ellipse model (Figure 7(c) and 7(h)). That means the segmentation and extraction of facial skin region based on proposed model are more effective than those based on the other skin models. Besides, some important holes which contained information of facial features remain to be preserved in facial skin-segmented images. It is prepared for acquiring and localizing facial features.



**Figure 7. Comparisons of Facial Skin-Segmented Images based on SGM, Ellipse Model and Proposed Unsupervised Model**

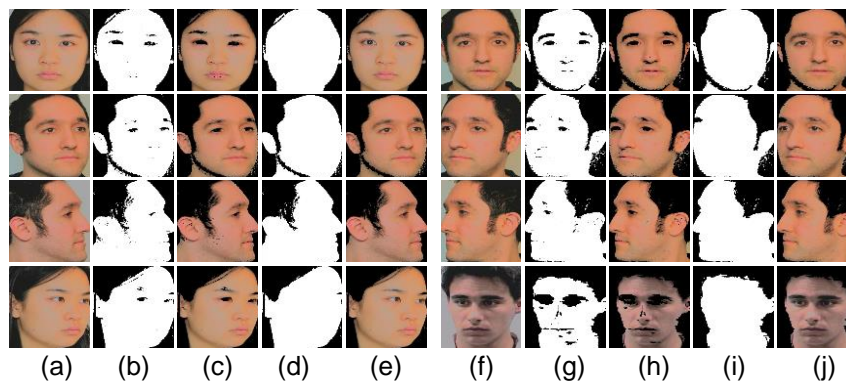
(a) Preprocessed face images in group 1, (b) Facial skin-segmented images of (a) based on SGM, (c) Facial skin-segmented images of (a) based on ellipse model, (d) Facial skin-segmented images of (a) based on proposed model, (e) Facial skin-extracted images of (a) based on proposed model, (f) Preprocessed face images in group 2, (g) Facial skin-segmented images of (f) based on SGM, (h) Facial skin-segmented images of (f) based on ellipse model,

(i) Facial skin-segmented images of (f) based on proposed model, (j) Facial skin-extracted images of (f) based on proposed model.

### 3.4. Obtain the Whole Face Targets

As shown in Figure 7, the facial skin-segmented image and facial skin-extracted image maybe contain noises caused by kinds of interference factors, such as device, environment, expression, pose, light, and so on. So their noises must be removed. While they also contained holes of facial feature, as color of facial feature (such as eyebrows, eyes, nose, and lip) is different with facial skin color. In order to acquire and locate facial features, here need to do in two aspects. One is to obtain the facial skin regions (as described in section 3.3) in facial skin-segmented image (Figure 8(b) and 8(g)) and facial skin-extracted image (Figure 8(c) and 8(h)), in which important information holes of facial features were preserved. Two is to obtain the whole face targets in face-segmented image (Figure 8(d) and 8(i)) and face-extracted image (Figure 8(e) and 8(j)), in which noises were removed and holes were filled. Note that the face-segmented image should be acquired based on facial skin-segmented image by technologies of removing noises and filling holes. Noises' removing was realized by Morphological operation ( $3 \times 3$  templates). Holes' filling adopted threshold ( $S/4$ ) of Inner contour area ( $S$ ) and was realized using Inner contour filling algorithm.

Figure 8 showed comparisons of facial skin regions and whole face regions. By comparison, it's clear that feature areas (eyebrows, eyes, nose, and lip) of facial skin region were obviously different with those of whole face region.



**Figure 8. Comparisons of Facial Skin Region and the Whole Face Regions**

(a) Preprocessed face images 1 in group 1, (b) Facial skin-segmented images of (a) based on proposed model, (c) Facial skin-extracted images of (a) based on proposed model, (d) the face-segmented images of (a) based on proposed model, (e) the face-extracted images of (a) based on proposed model, (f) Preprocessed face images 1 in group 2, (g) Facial skin-segmented images of (f) based on proposed model, (h) Facial skin-extracted images of (f) based on proposed model, (i) the face-segmented images of (f) based on proposed model, (j) the face-extracted images of (f) based on proposed model.

### 4. Facial Feature Location

After obtaining facial skin region and whole face target, the facial feature location can be realized by combined using information of difference, texture, and projection.

#### 4.1. Difference and DLBP Textural Features

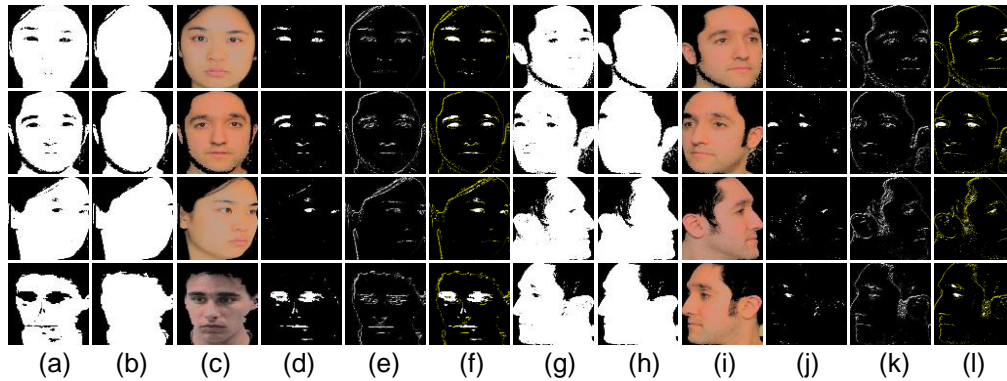
Based on facial skin-segmented image (Figure 9(a) or 9(g)) and face-segmented image (Figure 9(b) or 9(h)), the objects of facial feature in difference image (Figure 9(d) or 9(j)) can be acquired fast by spatial differencing technique. The difference image contains information of features' (eyebrows, eyes, nose, and lip) positions and part of information of face contour. It is regarded as a reference factor to help subsequent localization. Another reference factor is DLBP textural feature image.

Correspond to the face-segmented image, the face-extracted image preserved skin color of face target and removed the backgrounds, as shown in Figure 9(c) and 9(i). Based on face region of face-extracted image, the DLBP textural feature can be acquired, as shown in Figure 9(e) and 9(k). The DLBP (Different Local Binary Pattern) textural features is acquired based on LBP (Local Binary Pattern) [7] textural features (formula (19)-(20)), which is calculated according to proposed DLBP algorithm (formula (21)-(23)).

$$LBP_{p,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p, \quad (19)$$

$$s(g_p - g_c) = \begin{cases} 1, & g_p \geq g_c \\ 0, & g_p < g_c \end{cases}, \quad (20)$$

Where,  $g_c$  represents gray value of present pixel in position of  $(x_c, y_c)$ ,  $g_p$  is the eight-neighborhood pixels' gray values.  $s(g_p - g_c)$  is a sign function.



**Figure 9. Comparisons of Difference Images, DLBP Textural Feature Images, and Feature-Fused Images**

(a)Facial skin-segmented images of Figure 8(a) based on proposed model, (b) Face-segmented images of Figure 8(a) based on proposed model, (c) Face-extracted images of Figure 8(a) based on proposed model, (d) Difference images of (a) and (b), (e) DLBP textural feature images of (c), (f) Feature-fused images of (d) and (e), (g) Facial skin-segmented face images of Figure 8(f) based on proposed model, (h) Face-segmented images of Figure 8(f) based on proposed model, (i) Face-extracted images of Figure 8(f) based on proposed model, (j) Difference images of (g) and (h), (k) DLBP feature images of (i), (l) Feature-fused images of (j) and (k).

The traditional LBP value is calculated by multiply-accumulate of sign function  $s(g_p - g_c)$  and weight factors  $2^p$ . While sign function is obtained by comparing center pixel and eight-neighborhood pixels. Based on the traditional LBP algorithm, this paper proposed a

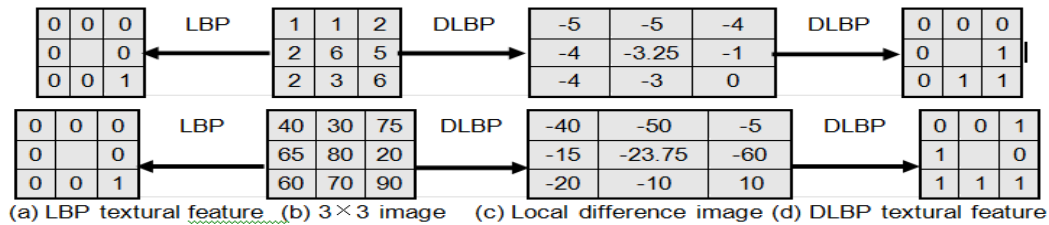
new DLBP textural algorithm, as shown in formula (21). It improved the sign function (formula (22)) by leading into an adaptive threshold T (formula (23)).

$$DLBP_{P,R} = \sum_{P=0}^{P-1} s(g_P - g_C) 2^P \quad (21)$$

$$s(x) = \begin{cases} 1, & x \geq T \\ 0, & \text{else} \end{cases} \quad (22)$$

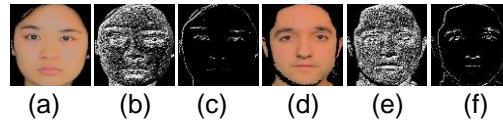
$$T = \frac{1}{P} \sum_{P=0}^{P-1} g_P - g_C \quad (23)$$

Figure 10 shows two comparisons of LBP textural feature and DLBP textural feature. From Figure 10, note that the different  $3 \times 3$  image had the same LBP textural features, but had the different DLBP textural features. This indicates the DLBP textural feature can discriminate and describe facial local feature (especially eyes and lip) better, which is help for subsequence localization.



**Figure 10. Comparison of LBP Textural Feature and DLBP Textural Feature**

Figure 11 shows two comparisons of LBP textural feature image and DLBP textural feature image. From Figure 11, it's clear that some textural feature in flat face region had been removed, and some pivotal textural feature in feature's regions (eyebrows, eyes, nose, lip, and facial contour) had been preserved.



**Figure 11. Comparison of LBP Feature Image and DLBP Textural Feature Image**

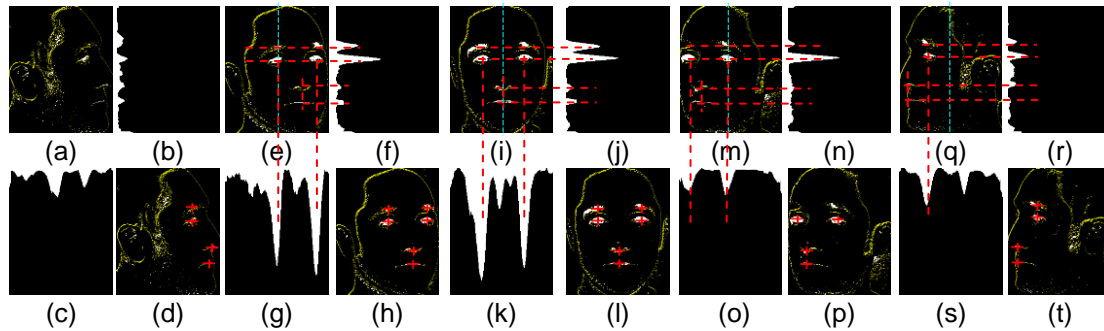
(a)(d) The face-extracted image, (b)(e) The LBP textural feature image, (c)(f) The DLBP textural feature image.

#### 4.2. Hybrid Integral Project

After obtaining the difference image and DLBP textural feature image, the feature-fused image can be acquired, as shown in Figure 9(f) and 9(l). The feature-fused image contains information of features' (eyebrows, eyes, nose, and lip) positions and DLBP textural feature. The hybrid integral project based on fused features was proposed, as is described in formula (24). Therefore the two aspects of information—DLBP textural feature's information and 2D information of difference—were combined used to help for searching for features' positions. Figure 12 shows five examples about hybrid integral project of fused feature for localizing facial features (eyebrows, eyes, nose, and lip).

$$S_v(x) = \sum_{y=y_1}^{y_2} I(x, y), \quad S_h(y) = \sum_{x=x_1}^{x_2} I(x, y) \quad (24)$$

Where,  $I(x, y)$  is pixel's gray value in feature-fused image in line  $x$  and column  $y$ .  $S_v(x)$ ,  $S_h(y)$  are vertical and horizontal integral projection functions in interval  $[x1, x2]$  and  $[y1, y2]$ , respectively. For an image with  $W$  width and  $H$  height,  $x1=0$ ,  $x2=W$ ,  $y1=0$ ,  $y2=H$ .



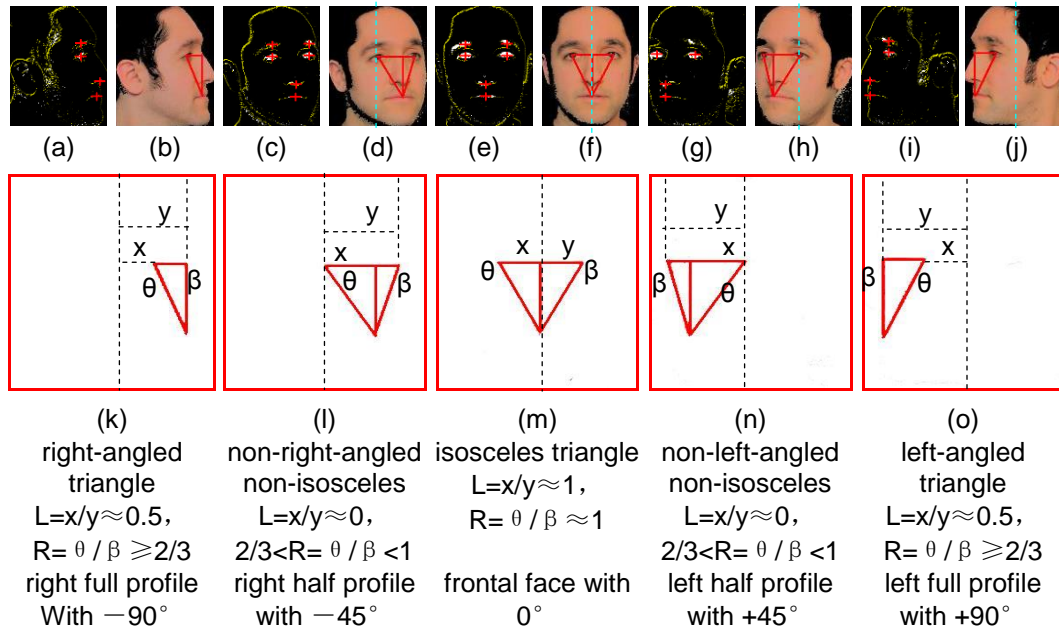
**Figure 12. Examples of Hybrid Integral Project**

(a) Feature-fused image 1 (right full profile with  $-90^\circ$ rotation), (b) Horizontal hybrid projection image 1, (c) Vertical hybrid projection image 1, (d) Feature-localized image 1, (e) Feature-fused image 2 (right half profile with  $-45^\circ$ rotation), (f) Horizontal hybrid projection image 2, (g) Vertical hybrid projection image 2, (h) Feature-localized image 2, (i) Feature-fused image 3 (frontal face with  $0^\circ$ rotation), (j) Horizontal hybrid projection image 3, (k) Vertical hybrid projection image 3, (l) Feature-localized image 3, (m) Feature-fused image 4 (left half profile with  $+45^\circ$ rotation), (n) Horizontal hybrid projection image 4, (o) Vertical hybrid projection image 4, (p) Feature-localized image 4, (q) Feature-fused image 5 (left full profile with  $+90^\circ$ rotation), (r) Horizontal hybrid projection image 5, (s) Vertical hybrid projection image 5, (t) Feature-localized image 5.

## 5. Pose Estimation

As stated above, the facial feature (eyebrows, eyes, nose, and lip) location can be realized by jointly using difference technology, DLBP textural feature, and hybrid integral project. Based on location information of facial features (Figure 13(a), 13(c), 13(e), 13(g), and 13(i)), the feature triangle which contained eyes and lip can be acquired (Figure 13(b), 13(d), 13(f), 13(h), and 13(j)). According to the feature triangle, five poses can be discriminated approximately for face image with horizontal rotation. For example, if feature triangle is a right-angled triangle, pose is right or left full profile with 90 degree rotation. If feature triangle is an isosceles triangle, pose is front face. If feature triangle is a non-isosceles triangle and also is a non-right-angled triangle, pose is right or left half profile with certain degree of  $(0\ 90)$ . The relationship (right or left) of position between triangle and central line can indicate the direction (left or right) of head's horizontal rotation.

In stage of pose estimation, two ratios ( $L$  and  $R$ ) were combined used to estimate the degree of head's rotation by limiting ratios' ranges. As shown in Figure 13, the ratio  $L$  ( $L=x/y$ ) reflects the relationship of distances ( $x$  and  $y$ ) between triangle's eyes (or triangle's right angle) and central line. The ratio  $R$  ( $R=\theta/\beta$ ) reflects the relationship of triangle's left-angle and right-angle degree ( $\theta$  and  $\beta$ ). Four parameters of triangle for calculating  $L$  and  $R$  are shown in Figure 13(k)-(o).



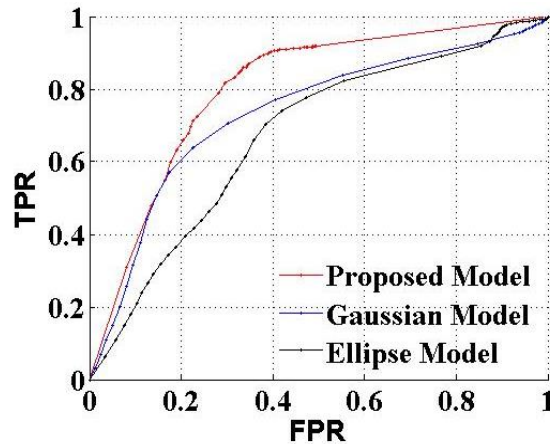
**Figure 13. Examples of Pose Estimation**

(a) Feature-localized image 1, (b) Pose-estimated image 1, (c) Feature-localized image 2, (d) Pose-estimated image 2, (e) Feature-localized image 3, (f) Pose-estimated image 3, (g) Feature-localized image 4, (h) Pose-estimated image 4, (i) Feature-localized image 5, (j) Pose-estimated image 5, (k) Pose-analysis image 1, (l) Pose-analysis image 2, (m) Pose-analysis image 3, (n) Pose-analysis image 4, (o) Pose-analysis image 5.

## 6. Experimental Results

In this section, the performance of facial skin detection based on three skin models (Single Gaussian Model, Ellipse Model, and Proposed Model) was compared, as shown in Figure 14 and Table 1. The performance of facial feature location using proposed method and other methods were compared, as shown in Table 2. The results of pose estimation were described in Table 3. Models were trained and experiments were tested by 500 face image and 500 non-face images from a mixed multi-pose database, which contained MIT CBCL face database [8] and self-built multi-pose face database.

Figure 14 shows the average ROC (receiver operating characteristic) curves of facial skin detection on multi-pose test sets, which represents ratios of TPR (True Positive Rate) and FPR (False Positive Rate) for each of three models under different thresholds. From Figure 14, the AUC (area under the curve) of ROC of proposed model is larger than AUC of ROC of Gaussian model and ellipse model. That means the proposed model gives the better performance and the color description ability than the other models on multi-pose test sets.



**Figure 14. The Average ROC Curves of Facial Skin Detection on Pose Test Sets for SGM, Ellipse Model, and Proposed Model**

Table 1 shows TPR (True Positive Rate), False Negative Rate (FNR), FPR (False Positive Rate), and TNR (True Negative Rate) of facial skin detection for three models under respective self-defined threshold on five pose test sets (full profile with  $-90^\circ$  and  $+90^\circ$ , half profile with  $-45^\circ$  and  $+45^\circ$ , frontal face with  $0^\circ$ ).

**Table 1. TPR, FNR, FPR, TNR of Three Models on Five Pose Test Sets**

Models Test Sets		Gaussian Model				Ellipse Model				Proposed Model			
Thresholds		Self-adaptive(OSTU)				Fixed Threshold(<1)				Fixed Threshold(<0.4)			
Parameters (%)		TPR	FNR	FPR	TNR	TPR	FNR	FPR	TNR	TPR	FNR	FPR	TNR
Test Sets	1( $-90^\circ$ )	69.3	30.7	30.0	67.0	63.6	36.4	10.0	90.0	99.5	0.5	24.7	75.3
	2( $-45^\circ$ )	75.5	24.5	3.0	97.0	63.0	37.0	23.5	76.5	98.4	1.6	23.8	76.2
	3( $0^\circ$ )	73.0	26.9	10.8	89.2	57.3	42.7	20.4	79.6	96.5	3.5	26.5	73.5
	4( $+45^\circ$ )	74.7	2.5	11.3	88.7	63.1	36.9	8.4	91.6	98.0	2.0	20.9	79.1
	5( $+90^\circ$ )	68.6	31.4	41.9	58.1	76.6	23.4	17.6	82.4	99.2	0.8	32.0	67.9
	Average	72.2	23.2	19.4	80	64.7	35.3	16.0	84.0	98.3	1.7	25.6	74.4

From Table 1, it's note that, when the average FPR and FNR of proposed model maintained a low values, the average TPR and TNR of proposed model can maintain a high values. This also means the performance of proposed model is better than the other models on multi-pose test sets.

Table 2 shows comparisons of the ACLR (average correct localization rate) and AFLR (average false localization rate) based on proposed method and other methods. The ACLR is obtained by using the average ratio of the total number of correctly detected facial features (eyebrows, eyes, nose, and lip) to the total number of facial features. The AFLR is obtained by using the average ratio of the total number of falsely detected facial features to the total number of facial features.

**Table 2. Comparisons of ACLR and AFLR of Proposed Method and Other Methods on Multi-pose Test Sets**

Methods		ACLR (%)				AFLR (%)			
Gaussian Model+SVM [2]		67.0				55.3			
Ellipse Model+Connected component grouping+Feature maps [5]		72.5				46.8			
Template matching [9]		64.3				55.9			
Adaboost [10]		56.2				65.6			
Proposed Method	Average	94.2				9.63			
	Eyes eyebrows nose lip	97.8	90.5	92.9	95.6	6.3	15.2	9.7	7.3

From Table 2, the proposed method has higher ACLR and lower AFLR than the other methods on multi-pose test sets in this paper. There were two main factors which decreased ACLR and increased AFLR of the other methods. As the other methods mainly focused on frontal face and frontal face with small-angle rotating, and they can't deal with left or right full profile. Therefore, the full profile test objects in multi-pose test sets became an important factor that decreased the ACLR and increased the AFLR of other methods. Besides, the other methods mainly searched for facial position roughly, as well as positions of eyes and lip. They basically don't locate the eyebrows and nose. So, the type of features is another important factor that influenced the AFLR and ACLR of other methods.

Table 3 shows the ratios of CPER (correct pose estimation rate) on five pose test sets. The CPER is calculated by using the ratio of the total number of correctly estimated poses to the total number of poses.

**Table 3. Comparisons of CPER on Five Pose Test Sets**

Test Sets	1 ( $-90^\circ$ )	2 ( $-45^\circ$ )	3 ( $0^\circ$ )	4 ( $+45^\circ$ )	5 ( $+90^\circ$ )	6 Average
CPER (%)	92.3	94.8	96.9	94.6	93.2	94.4

This method was realized by using VC6.0 and WinXP OS. The CPU is Intel(R) Core(TM) 2 Duo T6500@2.10GHZ and Memory is 2GB. The running time of algorithm depends on the normalized image size and image quality. For instance, the running time for a one-face image of size  $225 \times 275$  is 0.4s for face extraction, 0.8s for facial feature location and 0.3s for pose estimation.

There could be two situations required by this proposed method. One is the processed face image in multi-pose test sets can't have too large backgrounds. If the backgrounds are too large, it may influence the robustness of skin model. So before preprocessing, the collected original image may be clipped manually or automatically to a face image, in which the area of face is large and area of background is small. Two is the face in collected face image can only rotate in the horizontal range as the pose estimation is aiming at the horizontal rotated poses. Thus automatic algorithm for clipping backgrounds and estimation of non-horizontal rotated poses is next step of this research.

## 7. Conclusions

In this paper, a method of facial feature location and pose estimation for horizontal rotated pose face image was presented. In this method, a new preprocessing algorithm contained many correction technologies for improving image quality. For improving accuracy of skin

segmentation, a robust unsupervised sphere skin model was proposed to describe the skin color, and to lay a good foundation for subsequent processing. To fast and accurately localize the facial features, the difference features and proposed DLBP textural features were fused and technology of hybrid integral project also was combined used to search for the features' position and to accurately localize five kinds of facial features. By analyzing the geometric parameter (position, shape, and angles) of feature triangle related to eyes and lip, pose estimation algorithm was proposed to discriminate five typical horizontal poses. Experimental results show that preprocessing can help for the subsequence process effectively, the unsupervised sphere model had better robustness than the other model, the facial feature location achieved the average correct localization rate of 94.2%, and the pose estimation achieved the average correct pose estimation rate of 94.4%. Compared to other methods, this method needn't to scan all pixels in face image or all facial skin regions for projection. Therefore, it greatly saved the space and time, decreased the computational complexity. At the same time, it extended the types of localized facial features and realized fast localization and pose estimation. Besides, this method had good adaptability and stability by using adaptive preprocessing, robust unsupervised skin model, dynamic fused information of difference feature and DLBP feature, automatic technologies of hybrid integral project and fast pose estimation.

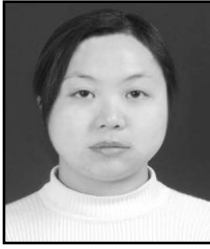
## Acknowledgements

This work was supported in part by the Scientific Research Foundation of the Education Department of Sichuan Province of China (No. 13ZB0012).

## References

- [1] A. R. Nusirwan, C. W. Kit and S. John, "RGB-H-CbCr Skin Colour Model for Human Face Detection", Proceedings of the MMU International Symposium on Information & Communications Technologies, (2006) November 16-17, Petaling Jaya, Selangor, Malaysia.
- [2] F. Y. Shih, S. X. Cheng and C. F. Chuang, "Extracting faces and facial features from color images", International Journal of Pattern Recognition and Artificial Intelligence, vol. 22, no. 3, (2008), pp. 515-534.
- [3] Z. W. Wang and S. Z. Li, "Face Recognition using Skin Color Segmentation and Template Matching Algorithms", Information Technology Journal, vol. 10, no. 12, (2011).
- [4] P. L. M. Bouttefroy, A. Bouzerdoum, S. L. Phung and A. Beghdadi, "On the Analysis of Background Subtraction Techniques using Gaussian Mixture Models", Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing, (2010) March 14-19, Dallas, TX, USA.
- [5] R. L. Hsu, M. Abdel-Mottaleb and A. K. Jain, "Face Detection in Color Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, (2002), pp. 696-706.
- [6] D. Z. Zhang, B. Y. Wu, J. B. Sun, and Q. L. Liao, "A face detection method based on skin color model", In Proceeding of the 11th joint Conference on Information Science, (2008) December 15-20, Shenzhen, China.
- [7] Y. Luo, C. M. WU and Y. Zhang, "Facial Expression Feature Extraction using Hybrid PCA and LBP", The Journal of China Universities of Posts and Telecommunications, vol. 20, no. 2, (2013), pp. 120-124.
- [8] "CBCL Face Recognition Database [Online]", Available: <http://cbcl.mit.edu/software-datasets/heisele/facerecognition-database.html>.
- [9] S. Tripathi, V. Sharma and S. Sharma, "Face Detection using Combined Skin Color Detector and Template Matching Method", International Journal of Computer Applications, vol. 26, no. 7, (2011), pp. 5-8.
- [10] M. Yang, J. Crenshaw and B. Agustine, "AdaBoost-based Face Detection for Embedded Systems", Computer Vision and Image Understanding, vol. 114, no. 11, (2010), pp. 1116-1125.

## Author



**Wei Li**, she received the B.S.C degree in computer science and technology from China West Normal University in 2004 and MSc degree in communication engineering from the Chongqing University in 2007. Since November 2009, she has been a lecturer in the Department of Communication Engineering at the China West Normal University. She specializes in pattern recognition and image processing. She is also interested in biomedical image analysis and embedded technology. She is a member of IEEE CS and China Computer Federation.