**TELKOMNIKA**, Vol. 11, No. 10, October 2013, pp. 5986 ~ 5994
ISSN: 2302-4046



# Face Detection Algorithm Based on Multi-orientation Gabor Filters and Feature Fusion

**Chuan Lin*[1], Xi Qin[2], Guo-liang Zhu[1], Jiang-hua Wei[1], Cong Lin[1]**
[1]College of Electric and Information Engineering, Guangxi University of Science and Technology
[2]Department of Electronic Information and Control Engineering, Lushan College of Guangxi University of Technology
*Corresponding author, e-mail: gxustlc@126.com, gxustceielc@126.com

***Abstract***

*In order to enhance the accuracy of multi-pose and multi-expression face detection, this paper proposes an algorithm based on multi-orientation Gabor feature fusion of mean and variance of sub-images. Firstly, to remove the huge background regions, we segmented images based on YCbCr space and then used two-eye templates to locate faces in skin-color regions by template matching. Secondly, the features were extracted from the face regions which had been removed the nose parts. After filtering by four-direction Gabor filters, the images were divided and then the corresponding constituent parts were formed. Thirdly, we calculated the mean and variance of the constituent parts block by block and processed them with features fusion. At last Support Vector Machine (SVM) and Back Propagation net (BP) were used for classification detection. The experimental results show that the algorithm has higher detection accuracy than other similar algorithms in multi-pose and multi-expression facial detections.*

***Keywords***: *Face Detection, Gabor Filters, Feature Fusion, Sub-images*



## 1. Introduction

Face detection has become a hot area of research in recent years. Face detection is the key technology and necessary prerequisite for face recognition, expression recognition, digital video processing and video detection, which aroused widespread attention in the place of pattern recognition, man-machine interaction, intelligent surveillance and video retrieval [1]. Essentially, face detection is called as the separation of face and background, which also means that determine its specific locations, sizes and quantities if there are human faces in the background of environment.

At present both domestic and international scholars have put forward many effective algorithms of face detection, which can be classified into four categories: Template matching algorithm [2, 3] achieves face detection through calculating the correlation between face templates and detecting images, using sequential module and deformable module could obtain a good result for different sizes, angles and poses, but the detection time is relatively longer; Knowledge-based face detection algorithm defines the rule through the prior knowledge of human beings [4, 5], but it is difficult to generalize face knowledge into human face rule which apply to computer, thus it is usually used to reduce the detection error rate; Feature invariant algorithm [6, 7] uses the geometry features, colour features and texture features of human faces and other parts for detection, but it requires a more demanding image quality, the noise, illumination and shadow are likely to damage the edge of human faces and thus affect the efficiency of the algorithm; Statistical theory-based algorithm [8, 9] uses the principal component analysis, SVM and neural network approach to map the detecting images to low-dimensional space and then separates the face and non-face. The algorithm [10, 11], which uses the grey information of human faces as features and detects them by BP and SVM after reducing the dimension by Principal Components Analysis (PCA), has a good instantaneity and detection accuracy. To enhance higher accuracy, feature selection and extraction are became the main research object. Reference [12] research makes clear that Gabor wavelet is very similar to the response to visual stimulation of the simple cell which is in Human Visual System (HVS). It has

*Received April 15, 2013; Revised July 10, 2013; Accepted July 20, 2013*

been successfully applied in many face detections and recognitions because of the good spatial locality and directional selectivity. The algorithm, which extracts feature by polydirectional Gabor filters and then uses BP and SVM for detection after dimension reduction [13, 14, 15, 16], had obtained a good effect, but the detection accuracy is easily influenced by dimension reduction algorithm due to the large feature dimension, the contributions of facial components also can't be embodied, hence, it is relatively difficult to deal with the face detection which has multi-pose face. Based on that, this paper focuses on feature extraction and selection that segment the skin color and locate two eyes at first, then features of the mean and variance of the four-direction 2D Gabor are fused, the SVM and BP are used for final classification detection at last. The experimental results show that the algorithm has higher detection accuracy than other similar algorithms [10, 11, 13, 14, 15, 16] in multi-pose and multi-expression face detection.

## 2. Skin-Color Segmentation of Human Face

### 2.1. Color-Space Selection

It's very effective to use skin-color feature to segment face in the drab surrounding and the situation which is significant different between skin color and surrounding. The calculation speed is also fast. Present color-space applications have several main types such as RGB, YCbCr, HSI, YIQ and YUV. YCbCr can be obtained by linear transformation of RGB. It can meet the needs of effectiveness and rapidity of the detection since the skin color has good cohesion in YCbCr [17]. As a result, YCbCr is used for color space of skin-color detection in this paper.

YCbCr derives from YUV. In YCbCr space the component *Y* denotes the brightness information of color, *Cb* and *Cr* denote the chrominance information and they clearly separate from each other. *Cb* denotes the difference between blue component and a reference value, *Cr* denotes the difference between red component and a reference value. After normalizing the components of RGB, the conversion formula from RGB to YCbCr has various kinds of forms. Through the experimental analyses, the formula is chosen as follows:

$$\begin{bmatrix} \mathrm{Y} \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ -0.1687 & -0.3313 & 0.5000 \\ 0.5000 & -0.4187 & -0.0813 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \tag{1}$$

### 2.2. Establishment of Skin-Color Model

Skin-color model is used as a form of algebraic to characterize the level of similarity between skin color and the hue of the detecting image. Its function is to construct the likelihood score of skin color through calculating the similarity probability, and then extract skin-color regions from images. Existing skin-color models include: Single Gaussian Model, Gaussian Mixture Model, histogram model and so on [18]. Because the statistical distribution of skin color is similar to Gaussian distribution in YCbCr space, Gaussian Model which has a simple mathematical expression is chosen in this paper. The model is also conducive to the detection efficiency. When establishing skin-color models, the selected images are all base on typical samples so that skin-color Gaussian Model can be applied to different skin colors, ages and sexes of individuals. Gaussian distribution model is defined as *G*(*m*,*c*), where *m* denotes the mean, *c* denotes the covariance. The expression is as follows:

$$m = E\{X\}, \quad X = (Cb, Cr) \tag{2}$$

$$C = E\left\{(X - m)(X - m)^T\right\} \tag{3}$$

$$P(Cb, Cr) = \exp[-0.5(X - m)^T C^{-1}(X - m)] \tag{4}$$

where EQ. (4) is the fit skin-color Gaussian Model function. The values of *m* and *c* are confirmed through calculating above samples. In this paper, the values of *m* and *c* are given as:

$$m = [108.4870, 150.2755], C = \begin{bmatrix} 77.8753 & -50.5381 \\ -50.5381 & 87.4238 \end{bmatrix}$$

After obtaining the values of *m* and *c*, the obtained skin-color likelihood images are calculated through all of the pixels of detecting samples to lay the foundation of skin-color segmentations.

### 2.3. Skin-Color Segmentation

The procedure of skin-color segmentation is shown in Figure 1. The color excursions of the collected color-images often occur because face images are easily influenced by complex environmental factors such as light and the color of light. Thus, the method of white balance is used for light compensation firstly. Secondly convert the detecting color-images to YCbCr color-space through EQ. (1) in order to obtain and count the corresponding chrominance values, and then calculate the skin-color likelihood images through EQ.(4). After mean filtering and gray stretch, use OTSU for threshold segmentation in order to convert the likelihood images to binary images and separate the color-images into skin and non-skin regions. At last, remove small areas of skin regions, find the bounding rectangles of identified skin regions and calculate the aspect ratios and then reserve the regions which have the height-width ratio from 0.6~1.5 as following detecting images. The process of the skin-color segmentation is shown in Figure 2.

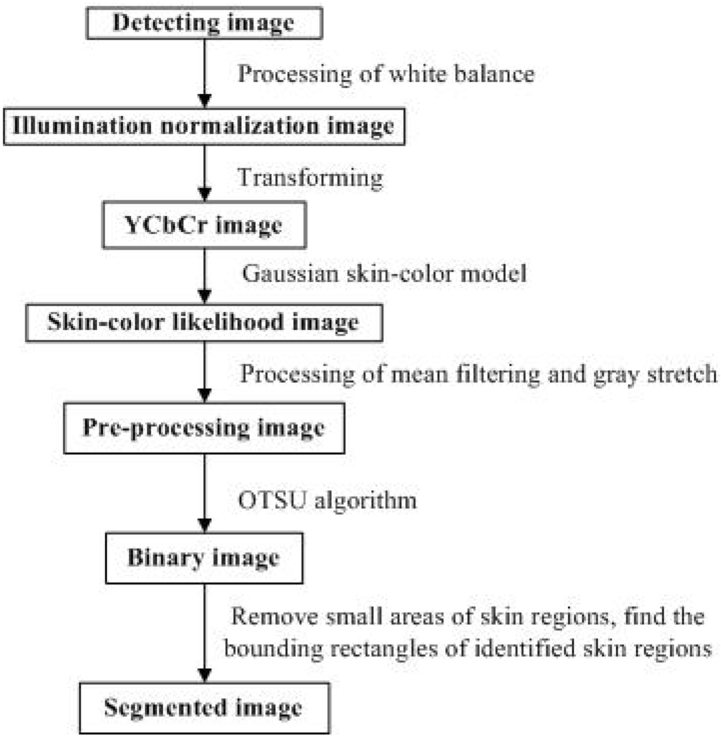
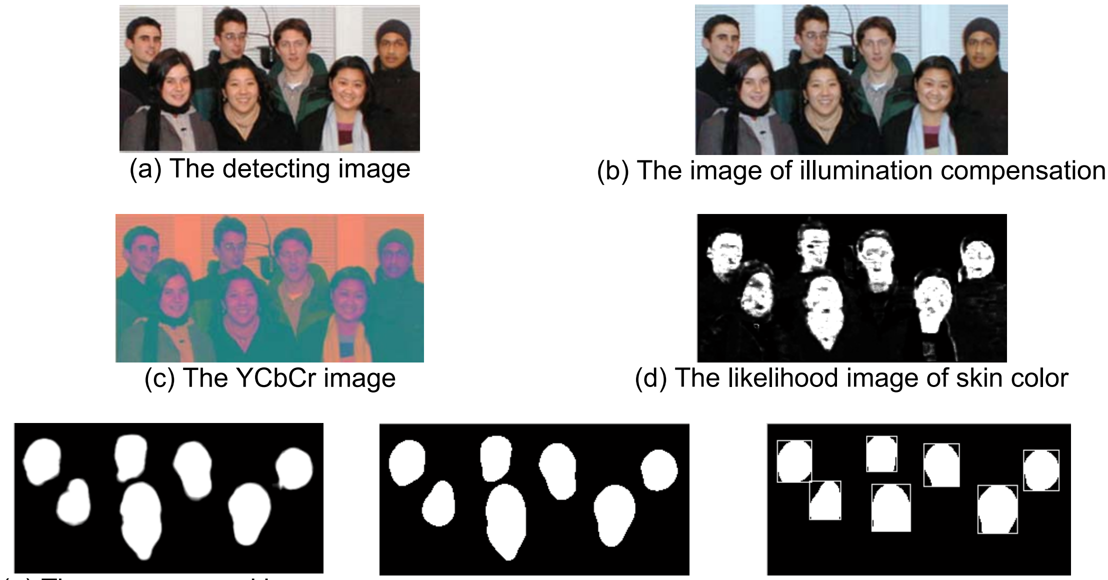



Figure 1. The procedure of the skin color segmentation

## 3. Facial Region Localization

### 3.1. Two-Eye Template Design

It has greater gray difference between eyes and skin color of human face, hence it is ideal to use eye templates for matching and face region localization. To solve the facial problems of different scales and adapt different aspect ratios of different eyes, five two-eye templates which have different aspect ratios are used for first screening in face detection.



(a) The detecting image (b) The image of illumination compensation

(c) The YCbCr image (d) The likelihood image of skin color

(e) The preprocessed image (f) The binary image (g) The image after segmenting

Figure 2. The skin-color segmentation process

Suppose the width and height of the two-eye templates are *W* and *H*, now the width-height ratios are set as: *W/H*=1.5, 2.0, 2.5, 3.0, 3.5.

Select different races and ratios of sample images and then crop 200 two-eye areas as two-eye samples. Because each sample has different scales and grayscale distributions, the samples should be processed with scale normalization and grayscale distribution normalization in order to remove the impacts of illumination. It means that grayscale mean and variance of samples should be adjusted as given values.

Suppose the sample image as*I(i. j)*, the computational formula of grayscale mean is:

$$\mu = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} I(i, j) \tag{5}$$

The computational formula of gray variance is:

$$\sigma^2 = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} \left[ I(i, j) - \mu \right]^2 \tag{6}$$

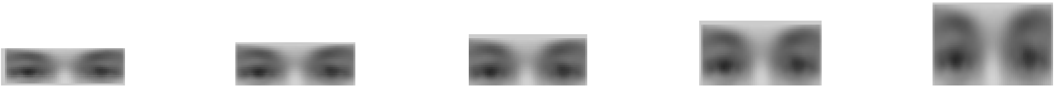
Calculate the mean $\mu$ and variance $\sigma$ of the samples through EQ.(5) and EQ.(6) and then set the standardized mean $\mu_0$ and variance $\sigma_0$(in this paper $\mu_0$=130, $\sigma_0$=60). The gray value on each pixel of the samples *I(i,j)* is transformed to obtain the samples $I_S(i, j)$ which have been processed with grayscale distribution normalization. The computational formula is:

$$I_S(i, j) = \frac{\sigma_0}{\sigma} \left[ I(i, j) - \mu \right] + \mu_0 \tag{7}$$

Then all of samples are taken grayscale mean and compressed into a scale which we need, after that the original two-eye templates are obtained, as shown in Figure 3.

Figure 3. The original two-eye template

At last different scales of templates are generated by original templates with height-width ratios, as shown in Figure 4.



(a) *W/H*=1.5 (b) *W/H*=2.0 (c) *W/H*=2.5 (d) *W/H*=3.0 (e) *W/H*=3.5

Figure 4. Different scales of templates

### 3.2. Two-Eye Template Location

Through using different scales of two-eye templates, retrieve the top half of bounding rectangles of skin-color regions which have been segmented, then zoom out the scales of detection regions as same as the templates and use similarity distance as matching method. Suppose the two-eye template as *P(i,j)*, its grayscale mean is $\mu_p$, mean square error is $\sigma_P$; *G(i,j)* denotes the input image, its grayscale mean is $\mu_G$, mean square error is $\sigma_G$,, the computational formula of similarity distance is given by :

$$d = \frac{\sum_{i=1}^{H}\sum_{j=1}^{W}\left[P(i,j)-\mu_P\right]\times\left[G(i,j)-\mu_G\right]}{H\times W\times\sigma_P\times\sigma_G} \tag{8}$$

After matching calculations, set the threshold in order to remove the regions which have the threshold less than 0.5, then array the rest of regions in descending order by similarity distance and select 15 regions which have the minimum distance, as shown in Figure 5. There may be some overlapping parts in the regions, thus combine the close regions in order to reduce redundant computations of the consequent face location.

The result after combination is shown in Figure 6.

Figure 5. The result image of template matching

Figure 6. The result after combination

## 4. Gabor Feature Extraction and Selection

### 4.1 Gabor Feature Extraction

Gabor wavelet is widely applied to various problems of extracting feature in face recognitions because wavelet transform has the time-frequency localization properties. Gabor wavelet transform is used for feature extraction and selection in this paper. 2D Gabor wavelet is consisted of a group of filters which have different scales and directions. It can extract the local feature of images accurately due to its nice locality, direction selectivity, bandpass and virtues of strong anti-jamming capability. 2D Gabor function is given by:

$$g(x,y) = \frac{1}{2\pi\sigma_{xy}^{2}}\exp\left[-\left(\frac{x'^{2}+y'^{2}}{2\sigma^{2}_{xy}}\right)\right]\times\left\{\exp(2\pi i r_0 x') - \exp(-\frac{r_0^{2}}{2\sigma_{uv}^{2}})\right\} \tag{9}$$

$$x' = x\cos\theta + y\sin\theta \tag{10}$$

$$y' = -x\sin\theta + y\cos\theta \tag{11}$$

where $r_0$ is the frequency of Gabor filters and $\sigma_{xy}$ is the standard deviation of the Gaussian envelop in the direction of $x$ and $y$, then we have: $\sigma_{xy}$=π, $1/\sigma_{uv}=2\pi\sigma_{x,}$, $r_0$=0.1681.

$\theta$ is denoted as the direction of 2D Gabor filters. Therefore, the four-direction Gabor filters are considered:

$$\theta_k = \frac{\pi(k-1)}{4}, k = 1,2,3,4 \tag{12}$$

The corresponding Gabor feature can be obtained after filtering the images. In the existing literature, most algorithms use the face regions which include eyes, noses and mouths as the samples for Gabor feature extraction [13,14,15,16], as shown in Figure 7. The methods have great limitation in multi-pose face, especially tilted face. Thus extract only eyes and noses (hereafter referred to as half face). Suppose detecting gray image as $I(x,y)$, the convolution of image $I(x,y)$ and a 2D Gabor filter is denoted as:

$$F(x,y,\theta_k) = I(x,y) * g(x,y,\theta_k) \;\; k = 1,2,3,4 \tag{13}$$

where $F(x, y, \theta_k)$ is called as the Gabor representation of image $I(x,y)$, the extraction effects are shown in Figure 8.

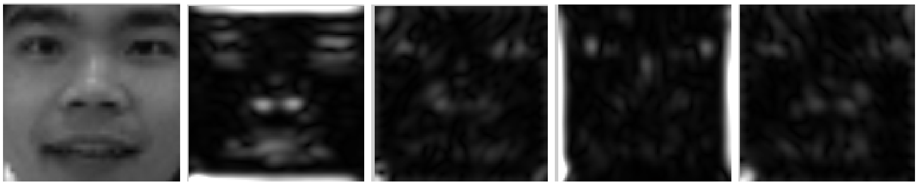


Figure 7. The Gabor feature extraction effects of full face

Figure 8. The Gabor feature extraction effects in this paper

### 4.2. Feature Fusion between Mean and Variance of Multi-Orientation Gabor

After extracting by four-directional Gabor filters, the feature dimension is four times the image size. In existing algorithms, the features which have better classifier are selected by dimension reduction to achieve final classifications [13, 14, 15, 16]. Hence, the classified accuracy is easily influenced by dimensionality reduction algorithms, the contributions of face and other parts also can't be incarnated in Gabor features. The algorithm of feature fusion between mean and variance of multi-orientationGabor is presented for feature extraction in this paper. On the basis of facial structure, firstly the face samples which only have the eyes and nose regions are processed with scale normalization: the height is 46 and the width is 60. As shown in Figure 9, divide the face into six parts: the left-eye region (A), the right-eye region (B), the glabellas region (C), the left cheek region (D), the right cheek region (E), the nose region (F). Among of them, the height of eyes region is 22 and the width is 25; the height of glabellas region is 30 and the width is 10; the height of cheek region is 24 and the width is 10; the height of nose region is 20 and the width is 40. Compose above regions into 13 different kinds of constituent parts according to Table1 and then calculate their mean $\mu$ and variance $\sigma$ to constitute featurematrix $G_i$:

$$G_i = [\mu_{i1}, \sigma_{i1}, \mu_{i2}, \sigma_{i2}, \mu_{i3}, \sigma_{i3}, \mu_{i4}, \sigma_{i4}] \qquad i = 1,2\cdots 13$$

where the subscript *i* denotes the number of constituent parts, the subscript (1, 2, 3, 4) denote four directions of the 2D Gabor filters. The feature vector *X* can be obtained through the ordered fusions of all constituent parts' feature matrixes:

$$X = \left[G_1, G_2, \cdots, G_i, \cdots, G_{13}\right]$$

Then the feature vector which has 104×1 column vectors is obtained:

$$X = \left[X_1, X_2, \cdots, X_i, \cdots, X_{104}\right]$$

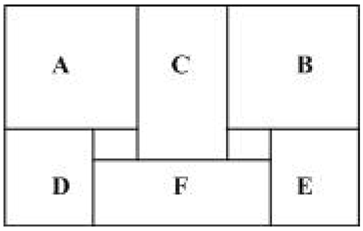



Figure 9. The block sketch of half face regions

Table 1. The table of constituent parts

| NO. of constituent parts | NO. of regions |
|---|---|
| 1 | A |
| 2 | B |
| 3 | C |
| 4 | D |
| 5 | E |
| 6 | F |
| 7 | AB |
| 8 | DE |
| 9 | ABC |
| 10 | DEF |
| 11 | AD |
| 12 | BE |
| 13 | ABCDEF |

## 5. Experimental Results and Analyses

To verify the effectiveness of the algorithm, the algorithm and previous algorithms are both processed with BP and SVM classification experiments. The two-eye templates are obtained through calculating 300 hand-cut samples, training samples and detecting samples are processed with normalization. The number of detecting samples is 2000, which include the front face, profile face and tilted face. Furthermore, these types have various situations, such as multi-pose (glasses and beard) and multi-expression, some face samples are shown in Figure 10. 1750 non-face samples are selected randomly from skin-like or skin regions, which contain no complete face regions. As shown in Figure 11.

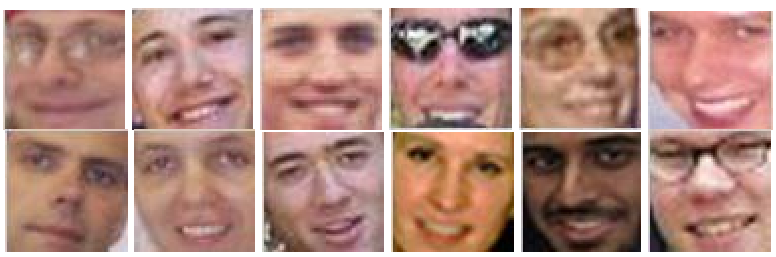


Figure 10. Face samples

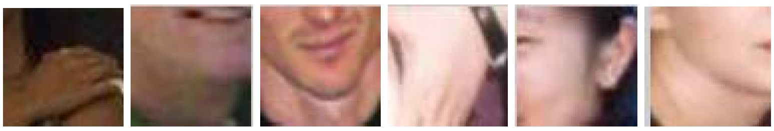


Figure 11. Non-face samples

In the references [10, 11], face gray is selected as the feature and extracted through reducing the dimensionality by PCA (hereafter referred to as the algorithm 1). In the references [13, 14, 15, 16], the four directional filters are used for filtering and extracted through reducing the dimensionality by PCA (hereafter referred to as the algorithm 2). In these two algorithms, the index which contributed more than 90% is selected as feature. Extract and select features from above training samples according to the algorithm 1, algorithm 2 and the algorithm in this paper, and then train them by BP and SVM.

Select 330 samples which include the front faces and tilted faces, then detect them by BP and PCA. The test results of the algorithm in this paper are shown in Figure 12 and Figure 13. The detection accuracies of each algorithm are shown in Table 2. It shows that the algorithm in this paper has higher detection accuracy than other algorithms. It's important to note that the algorithm 1 and 2 select different contributions for experiment when reducing the dimensionality by PCA. The accuracies listed in Table 2 are the highest among all contributions. It means that without dimension reduction the algorithm in this paper has the advantage of feature extraction and selection.

Table 2. Detection accuracies of algorithms

| Types of faces | Algorithm 1 + BP | Algorithm 1 +SVM | Algorithm2 + BP | Algorithm2 +SVM | Algorithm inthispaper +BP | Algorithm inthispaper +SVM |
|---|---|---|---|---|---|---|
| Front | 86.36% | 85.45% | 96.36% | 97.58% | 99.70% | 99.09% |
| faceTilted face | 72.73% | 76.36% | 90.91% | 91.81% | 96.97% | 97.58% |

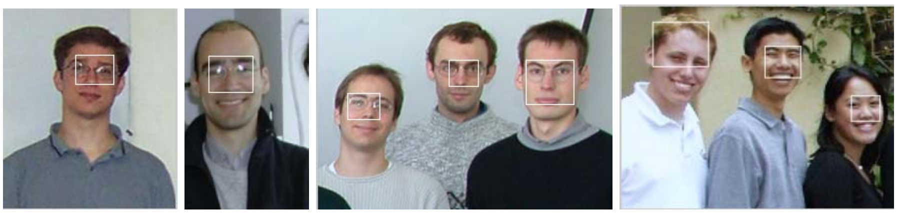


Figure 12. Detection results of multi-pose and multi-expression front faces

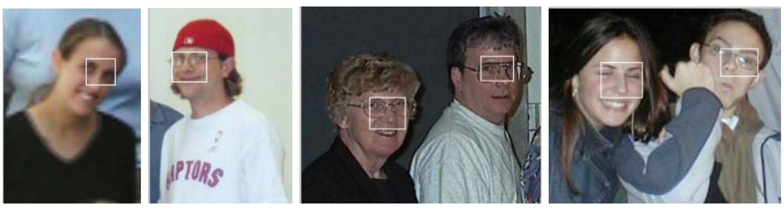


Figure 13. Detection results of multi-pose and multi-expression tilted faces

## 6. Conclusion

This paper researches human eye localization and skin-color segmentation of human face. Features are extracted and selected by using 2D Gabor filters and finally detected by SVM and BP. Furthermore, this paper proposes a feature selection algorithm that the face images which only include eyes and nose regions are processed with 2D Gabor wavelet transformation, block calculations of mean value and variance and fusion. The algorithm uses the Gabor wavelet transform which has the time-frequency localization properties and the invariance in translation, size and rotation. The experimental results show that the algorithm has higher detection accuracy in the detection of multi-pose and multi-expression face.

**Acknowledgements**

We sincerely thank the guest editors and the anonymous referees for their valuable suggestionsand comments. Their comments helped to improve the quality of the paper immensely. This research issupported by the project of Guangxi University Students' innovative undertaking training program in 2012 (No.31).

**References**

[1] Hjelmas E, Low BK. Face detection: A survey. *Computer Vision and Image Understanding*. 2001; 83(3): 236-274.
[2] Cootes TF, Taylor CJ. *Locating face using statistical feature detectors*. Proc. Second Int'1 Conf. Automatic Face and Gesture Recognition. 1996; 204-209.
[3] Lanitis A, Taylor CJ, Cootes TF. Automatic tracking coding and reconstruction of human faces, using flexible appearance models. *IEEE Electron. Letters*, 1994; 30: 1578-1579.
[4] Yang MH, Driegman DJ, Ahuja N. Detecting f aces in images: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002; 24(1): 34-58.
[5] Yang GZ, Huang TS. Human face detection in a complex background. *Pattern Recognition.* 1994; 27(1): 53-63.
[6] Kong SG, Heo J, Abidi BR, et al. Recent advances in visual and infrared face recognition a review. *Journal of Computer Vision and Image Understanding*. 2005; 97(1): 103-135.
[7] Dai Y, Nakano Y. Face-texture model based on SGLD and its application in face detection in a Color Sence. *Pattern Recognition*. 1996; 29(6): 1007-1017.
[8] HA Rowley, S Baluja, T Kanade. Neural network-based face detection. *IEEE Transform Pattern Analysis and Machine Intelligence*. 1998; 23-38.
[9] HA Rowley, S Baluja, T Kanade. Rotation invariant neural network-based face detection. *IEEE Conf. Computer Vision and Pattern Recognition*. 1998; 38-44.
[10] Wang Hongman, Ou Zongying. Face Recognition Based on Features by PCA/ICA and Classification with SVM. *Journal of Computer-Aided Design and Computer Graphics*. 2003; 4(15): 416-420.
[11] Zhou Jingli, Wu Guilin, Yu Shengsheng. Algorithm of BP Neural Network-based Face Detection. *Computer Engineering*. 2004; 11(30): 34-36.
[12] Daugman J. *Complete Discrete 2-D Gabor Transform by Neural Networks for Image Analysis and Compression*. IEEE Trans. Acoustics. Speech and Signal Processing. 1998; 36(7): 1169-1179.
[13] Huang Lin Lin, Shimizu A, Kobatake H, *Robust face detection using Gabor filter features*. Pattern Recognition Letters. 2005; 1641-1649.
[14] Lin Ke-zheng, Xu Ying, Li Shu, Face. Recognition Using 2D Gabor Mean Values in Subspace. *Journal of HARBIN University of Science and Technology*. 2012; 17(5): 65-68.
[15] Lin Yu-sheng, Yang Jing-yu. Face detection based- on gabor filter and support vector machines. *Computer Engineering and Applications*. 2007; 43(1): 33-34.
[16] Yan Wen-xiu, Pei Jian-gang, Sun Ying. A Face Detection Method Based on Gabor Filters and Improved BP Neural Network. *Journal of Chongqing Institute of Technology (Natural Science).* 2009; 3(4): 98-102.
[17] Hsu RL, Mohamed AM, Anil KJ. Face detection in color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002; 24(5): 696-706.
[18] Ishii, Hideaki, Minoru Fukumi, and Norio Akamatsu. *Face detection based on skin color information in visual scenes by neural networks.* Systems, Man, and Cybernetics, 1999. *IEEE SMC'99 Conference Proceedings*. 1999; 350-355.