http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 10 (12): 2308-2314, 2011
ISSN 1812-5638 / DOI: 10.3923/itj.2011.2308.2314

Face Recognition using Skin Color Segmentation and Template Matching Algorithms

[1,2,3]Zhiwen Wang and [2,3]Shaozi Li
[1]Department of Computer and Engineering, Guangxi University of Technology, Liuzhou, China
[2]Department of Cognitive Science, Xiamen University, Xiamen, China
[3]Key Laboratory of the Brain-like Intelligent Systems, Xiamen University, Xiamen, China

Abstract: Currently, relatively popular and representative face recognition algorithms are algorithm based on template matching and algorithms based on skin-color segmentation. The computation of recognition algorithm based on template matching is very high and the recognition rate of recognition algorithms based on skin color segmentation is low and is vulnerable to the impact of background which is similar with skin color, In order to overcome these deficiencies, face recognition using skin color segmentation and template matching algorithm is presented in this study. According to the clustering properties that the skin-color of human faces emerge in the YCbCr color space, the regions closing to facial skin color are separated from the image by using Gaussian mixture model in order to achieve the purpose of rapidly detecting the external face of human face. Adaptive template matching is used to overcome the affect of the backgrounds which are similar with skin color on face detection and recognition. Computation in the matching process is reduced by using the second matching algorithm. Extraction of face images by using singular value features is used to identify faces and to reduce the dimensions of the eigenvalue matrix in the process of facial feature extraction. Experimental results show that proposed method can rapidly detect and recongnise human faces and improve the accuracy of face detection and recognition.

Key words: Human face recognition, singular value feature, mixture Gaussian model, secondary matching algorithm, template matching

INTRODUCTION

Research on recognition technology is currently the research hot spot in the field of pattern recognition and artificial intelligence. There are wide ranges of applications, such as identification of criminals in public security systems, security verification systems, credit card verification, medicine, file management, video conference, HMI systems, verification of actual holders of driver's license or passport, monitoring systems, automatic entry systems and so on. Up to now, the mainly algorithms of face recognition can be divided into following categories:

- **Geometric approach:** Such methods mainly employ the geometric shape of human face and the proportional relation of facial organs to detect and recognize a face. There are two of them, one is bottom-up approach and the other is top-down method. The former firstly detects each feature sites (e.g., eyes, nose and lip) of a face and then reconstructs the face according to the sites. While the later locates the potential locations of faces at first and then verifies them according to the facial feature sites. The location of eyes is very important in these methods. Due to this, the high quality image around the eyes area is required which will to some extent restrict their applications

- **Template matching approach:** A pixel-by-pixel scanning is conducted in an image to be detected by using a template and the matching degree is calculated between pixels in the image and the template, according to which faces can be detected and recognized. The speed and efficiency of the detection are relatively low because of the large amount of calculation

- **Approach based on classification:** In this approach, face detection and recognition by using facial global features (e.g., skin color, or gray distribution) can avoid the detection of each facial organ. Face detection methods based on color information are receiving more and more concern in recent years and have become the research hot spot because of their small amount of calculation, good stability, simplicity of description and contribution to achieve real-time processing. However, face detection by using color information suffers a lot from the impact of complex

Corresponding Author: Zhiwen Wang, Department of Computer and Engineering, Guangxi University of Technology, Liuzhou, China

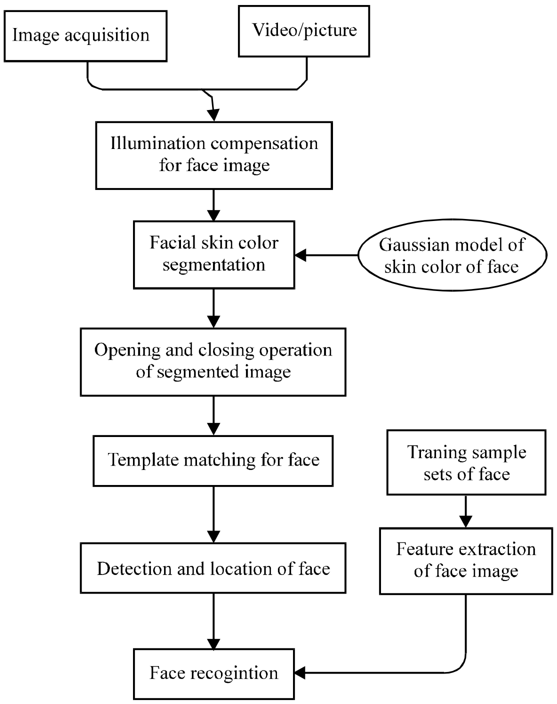

Fig. 1: The block diagram of face recognition

background, especially the impact of skin-like background and luminance. What's more, it's rather difficult to describe the facial features explicitly (Zong *et al*., 2001; Greenspan *et al*., 2001; Wang *et al*., 2010)

In order to overcome these deficiencies that computation of face recognition algorithm based on template matching is very high and the recognition rate of face detection and recognition algorithms based on skin-color segmentation is low and is vulnerable to the impact of background which is similar with skin-color, a face recognition algorithm based on skin color segmentation and template matching is presented in this study. The block diagram of face detection and recognition of human face image is shown in Fig. 1. According to the characteristic that human skin-color clusters in the YCbCr color space, regions closed to facial skin-color can be separated from the image by using mixture Gaussian model for human skin color to conduct segmentation and then achieve the purpose of rapid outer-face detection. A self-adapted template has been used to modify the impact of skin-like background. Furthermore, second matching algorithm has been applied to reduce the amount of calculation. Satisfied result has been obtained when singular values extracted from the image were used for face recognition, meanwhile, the dimensions of the singular value matrix has been deduced during the process of extracting facial features (Toews and Arbel, 2009; Raghuraman and David, 2010).

SKIN COLOR SEGMENTATION ALGORITHM BASED ON GAUSSIAN MIXTURE MODEL FOR HUMAN SKIN COLOR

Usually, skin color segmentation algorithm model by using Gaussian mixture model, because facial color region can be described by Gaussian distribution.

Gaussian mixture model for human skin color: Skin color is an important feature of human face. Despite that the facial color of people from different racials, ages looks different but the main differences are caused by luminance (Zong *et al*., 2001). After eliminating the luminance, the distribution of facial color from different people has good clustering property (Chao-Kuei *et al*., 2009). And its statistical distribution in the YCbCr skin-color space satisfies (1):

$$98 \leq Cb \leq 127,\ 133 \leq Cr \leq 170 \tag{1}$$

Hence, this skin-color clustering property can be applied to detect face. In the two-dimensional color space, the facial color region can be described by Gaussian distribution (Toews and Arbel, 2009; Raghuraman and David, 2010). Facial color Gaussian distribution of three main racials from different genders and various ages can be illustrated by Fig. 2.

According to the Gaussian distribution of face color in YCbCr space, the values of Cb and Cr of the color samples of face are more concentrated, so they are suitable to be used to establish the Gaussian model. For each pixel in a color image, the probability of whether it belongs to facial color region can be calculated, as long as it is converted from RGB color space to YCbCr color space by using formula (2). Put it another way, the similarity between a pixel and facial color can be calculated on the basis of the distance between the pixel and the Gaussian distribution center. And then, convert the color image into a gray image whose intensity of each pixel is associated with this similarity. Figure 3 illustrates the 2D-Gaussian model of facial color which is established in accordance with Eq. 3:

$$\begin{cases} Y = 0.257\times R + 0.504\times G + 0.098\times B + 16 \\ Cb = -0.148\times R - 0.291\times G + 0.439\times B + 128 \\ Cr = 0.439\times R - 0.368\times G - 0.071\times B + 128 \end{cases} \tag{2}$$

$$m = (\overline{Cb},\overline{Cr}),\ \overline{Cr} = \frac{1}{N}\sum_{i=1}^{N} Cr_i,\ \overline{Cb} = \frac{1}{N}\sum_{i=1}^{N} Cb_i,\ V = \begin{bmatrix} \sigma_{CrCr} & \sigma_{CrCb} \\ \sigma_{CbCr} & \sigma_{CbCb} \end{bmatrix} \tag{3}$$

where, $\overline{Cb}, \overline{Cr}$ are means of Cb, Cr, respectively; V is covariance matrix and N is the total number of pixels of the face (Ksantini *et al*., 2010; Chunhua *et al*., 2007).

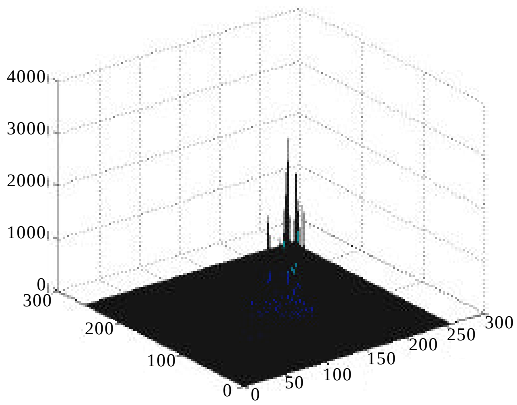

Fig. 2: Gaussian distribution of face color in YCbCr space

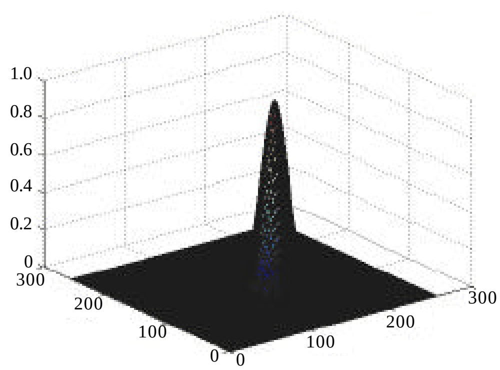

Fig. 3: 2D-Gaussian model of facial color

Gaussian density function for the facial color of the left face can be calculated by Eq. 4:

$$p_L(CbCr) = k_L \exp[-0.5(x_L - m)^T \textstyle\sum_L^{-1}(x_L - m)] \tag{4}$$

Gaussian density function for the facial color of the right face can be calculated by Eq. 5:

$$p_R(CbCr) = k_R \exp[-0.5(x_R - m)^T \textstyle\sum_R^{-1}(x_R - m)] \tag{5}$$

where, $\sum_L^{-1}$ and $\sum_R^{-1}$ are covariance of the left and right face, respectively. k_L and k_n are Gaussian model constants for the left and right face and:

$$k_L = (2\pi)^{-1}\left|\textstyle\sum_L\right|^{-\frac{1}{2}},\ k_R = (2\pi)^{-1}\left|\textstyle\sum_R\right|^{-\frac{1}{2}}$$

When color image has been converted into similarity gray image by using Gaussian skin color model and then the skin-color regions can be divided from non-skin-color regions by choosing a proper threshold . This model is characterized by skin color model based on statistical which needs to calculate the similarity for every pixel. Therefore, its computational speed is not fast enough. In the practice of skin color detection, the items:

$$[-0.5(x_L - m)^T \textstyle\sum_L^{-1}(x_L - m)]$$

and:

$$[-0.5(x_R - m)^T \textstyle\sum_R^{-1}(x_R - m)]$$

from Eq. 3 and 4 can be used for decision directly in order to improve the detection speed (Gan and Wen-Jun, 2006).

Facial skin color segmentation: After establishing the Gaussian model of face, facial skin color segmentation can be achieved according to the following steps.

Step 1: Design a skin classifier based on color core and fuzzy segmentation and then conduct skin segmentation on a color image by using this classifier, whose input is the original color image and output is a binary image which demonstrates the skin region and the non-skin region.

Step 2: When doing skin segmentation, it is impractical to reliably measure the variance of luminance around the face region which is caused by the ambient light. However, luminance can be removed from the color representation in YCbCr color space. After normalization processing, chromic color can be defined by Eq. 6 according to the formula f (R, G, B) = g (r, b) which can also be regard as solid color without the luminance component:

$$r = R/(R + B + G),\ b = B/(R + B + G) \tag{6}$$

Even though, the skin-color variance due to different people expends into a wide range but this kind of variance is far less than the variance caused by luminance. According to this, a well-performing skin color model can be established in the color space. After much calculation, we get the value of m (117.4316 148.5599).

Step 3: Seen from the skin-probability image, skin regions (e.g., face, hands) have higher brightness than non-skin regions. Hence, skin regions can be divided from non-skin regions by setting threshold. For various images of different skin-color and luminance, it is impossible to

Fig. 4: Skin image segmentation process. From left to right: the original, the YCbCr color space, the skin probability, the segmented image

determine a fixed threshold that will conduct accurate segmentation in all cases. Considering this, a fuzzy classification method is required to distinguish between skin and non-skin. First, do fuzzy classification by clustering criteria described with formula (7). Second, apply a self-adapted threshold based on region growing algorithm. Optimal threshold is adopted each time when doing segmentation. The reason is that, self-adapted threshold based on region growing algorithm is obtained according to the step-by-step result (Jun *et al*., 2003; Yunyang *et al*., 2008; Qiao *et al*., 2010). More specifically, progressively decreasing the threshold will result in the enlargement of the segmented region but the change of enlargement is gradually reduced each time, so that the threshold which cause the minimum regional growth is the optimal one:

$$\min f_k(X, \mu, v) - \sum_{i=1}^{n}\sum_{j=1}^{n}\mu_{ij}v_i - X_i^2 = \sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^2 d_{ij}^2 \tag{7}$$

where, $k \in |0, \infty|$ are weights. $f_k(X, \mu, v)$ is the objective function which is the weighted sum of squares of the within-class error. μ_{ij} is fuzzy partition matrix and v_i is fuzzy cluster centers which can be calculated by formula (8). d_{ij} is within-class distance and c is the number of cluster categories:

$$\mu_{ij} = 1/\sum_{1}^{c}\left(\frac{d_{ij}}{d_{ij}}\right)^{\frac{2}{k-1}}, \quad 1 \le i \le c, 1 \le j \le c, \quad v_i = \sum_{j=1}^{n}\mu_{ij}X_j / \sum_{j=1}^{n}\mu_{ij} \tag{8}$$

And μ_{ij} needs to meet the following three constraints:

$$\begin{aligned} &\mu_{ij} \in [0,1], 1 \le i \le c, 1 \le j \le n, \\ &\sum_{i=1}^{c}\mu_{ij} = 1, 1 \le j \le n, \\ &n > \sum_{j=1}^{n}\mu_{ij} > 0, 1 \le i \le c \end{aligned} \tag{9}$$

Opening and closing operations in image segmentation: After skin-color modeling, we can get some connected regions which contain not only the face regions but also some other skin-color regions (e.g., arms, necks). During the process of binarization of a noise image, the obtained boundaries are always not smooth, some object regions are often wrongly determined, small noise is scattered on the background area. In order to further improve the performance of face skin-color segmentation, opening and closing operations are adopted to segment the image and significant improvements have been achieved to the processed images (Wang, 2010; Hsu *et al*., 2002). Opening operation means conducting corrosion to the image at first and then conducting expansion to the corrosion structure which can be defined by formula (10):

$$A \circ B = (A \odot B) \oplus B \tag{10}$$

where, A and B are the sets in Z, is the opening operator, is the expansion operator. Generally, opening operation can smooth contours of the image, weaken the narrow parts and remove the tiny protrusions.

On the contrary, closing operation is conducting expansion to the image at first and then conducting corrosion to the expanded result. Closing operation is defined by formula (11):

$$A \bullet B = (A \oplus B) \odot B \tag{11}$$

where, • denotes closing operator. Closing operation can also smooth contours of the image. But contrast with the open operation, closing operation can generally merge the narrow gaps and curved slit, remove the small holes and fill the gap on the contours. Some tiny glitches in the image can be removed after opening and closing operations and achieve the purpose of denoise. Next, glitches can be further removed by applying fill-hole technique (Wang and Wu, 2010).

Facial skin color segmentation is shown in Fig. 4.

FACE RECOGNITION BASED ON TEMPLATE MATCHING AND SKIN COLOR SEGMENTATION

Facial feature extraction: For the normalized face images, wavelet transformation combined with DCT is adopted to extract the facial features. Firstly, conduct three-level

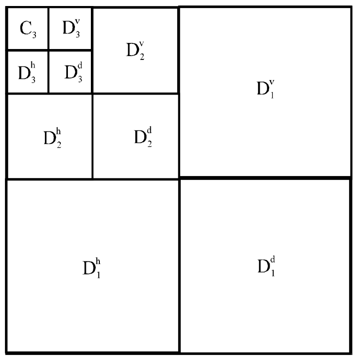

Fig. 5: Three-level wavelet decomposition of an image

wavelet decomposition to the face image (Fig. 5). Secondly, select the low-frequency sub-image as the object of feature extraction. Thereby, obtain the low-frequency sub-images of the training images or the test images (Heusch and Marcel, 2010; Zhi-Wen *et al*., 2010; Queirolo *et al*., 2010). Suppose that x_i stands for the 1D vector of the low-frequency sub-image of ith face image, then the covariance matrix can be represented by formula (12):

$$C = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(x_i - \bar{x})^T \tag{12}$$

where, N means the total number of training samples:

$$\bar{x} = \sum_{i=1}^{N} x_i$$

is the average vector of the training sample set. Because C is a symmetric matrix, so it can be diagonalized as follows:

$$C = U\Lambda U^T = \sum_{i=1}^{R}\lambda_i u_i (u_i)^T \tag{13}$$

where, λ_i are the eigenvalues of C. U is the corresponding eigenvectors; $\{u_1, u_2, ..., u_{i-1}, u_1, ..., u_{R-1}, u_R\}$ is standard orthogonal basis. R is the rank of C. Λ is a diagonal matrix whose diagonal elements are the eigenvalues of C.

Because the dimension of the covariance matrix C is high, the calculation amount is huge if the eigenvalues and the orthonormalized eigenvectors of C are directly calculated. So, SVD can be adopted for further dimension reduction.

Theorem (SVD theorem): Suppose $A_{m\times n}$, without loss of generality, assume $n \geq m$ and Rank (A), then there exists two orthogonal matrices $U_{m\times m}$, $V_{n\times n}$ and a diagonal matrix A which satisfies $A = AAV^T$, where, $A = \text{diag}(\lambda_1, \lambda_2 ... \lambda_{k-1}, \lambda_k, 0, ..., 0)$ and $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k . \lambda^2_i$ (I = 1, 2, ..., k) are eigenvalues of AA_T and A^TA while U and V are both orthogonal matrices.

Regard a face image as a matrix A, then k non-zero eigenvalues of A combined with n-k 0 form a n dimensional column vector. Assume Y is a singular value feature vector of A (Chunhua *et al*., 2007; Gan and Wen-Jun, 2006; Jun *et al*., 2003). For an arbitrary real matrix A, there exists only one singular value feature vector corresponding to the original face image when $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_{k-1} \geq \lambda_k$. The so called eigenfaces which represent the algebraic features of face (Yunyang *et al*., 2008; Hsu *et al*., 2002; Wang and Wu, 2010). Singular value features are insensitive to the gray variation caused by noise or illumination variation and can overcome the impact from illumination, image size, image rotation and image deflection, pose variation and some other factors during the recognition process. Hence, singular value features can be used more effectively to recognize face images.

Face recognition: During the recognition process, A (video A_j, j is the frame subscript of the video) is an image to be recognized, from which a singular value feature vector Y is extracted. The identification feature vector (Id) of this image can be calculated by:

$$Id = H^TV^TY.$$

where, $H = (h_1, h_2, ..., h_{n-c})$ are eigenvectors corresponding to the n-c biggest eigenvalues of the total distributed matrix which is made up by Y from all training images. $V = (v_1, v_2, ...v_d)$ is the projection vector of the eigenvectors corresponding to the first d largest eigenvalues. There is a unique Id belonging to each face image A. In order to let the test samples and the training samples be comparable, feature vectors of all training samples have been extracted. The average feature vector of the training samples is calculated by formula (14):

$$m_k = \frac{1}{N}\sum_{i=1}^{N} V_{k,i} \tag{14}$$

where, N is the number of training samples, $V_{k,I}$ is the ith feature of the ith sample, m_k is the k th eigenvector of average identification feature vector of the training face images. Then a minimum distance classifier is adopted to classify the identification feature vector (Id). For a test sample A, the between-class distance is calculated by $d(m_k, Id) = \|m_k - Id\|_2$. If $d(m_k, Id) = \min d(m_k, Id)$, then $A \in \omega_k$. Recognition results are shown in Fig. 6.

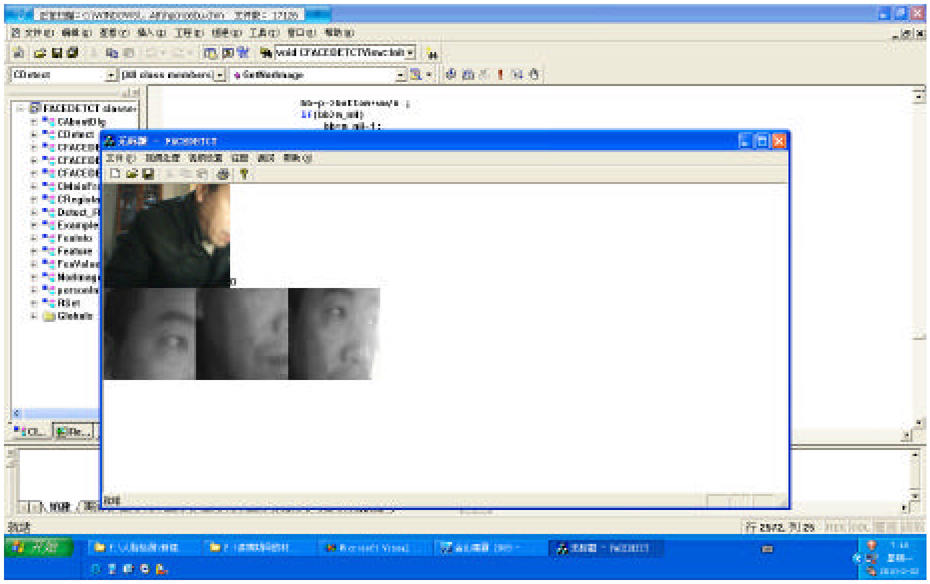

Fig. 6: The results of face recognition for real-time video

RESULTS ANALYSIS OF FACE RECOGNITION BASED SKIN COLOR SEGMENTATION AND TEMPLATE MATCHING

The typical face databases (ORL, UMIST, CVL) as well as the real-time video have been used for tests in the experiments, meanwhile, some face images downloaded from the internet have been used for supplement experiments. The size of face image is range from 5 to 800 k, the face images include photographs taken from different views, photographs that have various degree of change in facial expression and facial details, photographs whose face poses have substantially changes, photographs whose faces have various degree of rotation, photographs that contain faces of different size, illumination, age and skin-color, photographs of high or low collar, etc., have been chosen in this study. Algorithms proposed by this study or the references have been used for face recognition.

Experimental results show that, the proposed algorithms for face recognition could well overcome the impact from illumination variance, collar height difference, face size changes, different facial expressions, views, ages, poses, complex backgrounds, various shooting angles, skin-like backgrounds and some other factors. In other to compare with other face recognition algorithms such as algorithm based on Haar feature, Adaboost approach and algorithm based on Bayesian decision rule proposed by Wang and Wu (2010). Table 1 shows comparison results of face recognition:

Table 1: The comparison results of face recognition

Method	Training samples	Test samples samples	Recognition rate (%)
LDA	1180	1180	94.253
EBGM	1180	1180	95.384
LBP	1180	1180	97.015
Method proposed by Qiao *et al.* (2010)	1180	1180	96.185
Method proposed by Wang (2010)	1180	1180	97.156
Method proposed in this study	1180	1180	99.018

$$Rr = \frac{\text{No. of detected true face}}{\text{Total No. of face in image}} \times 100\% \qquad (15)$$

CONCLUSIONS

We have proposed a face recognition algorithm based on skin-color segmentation and template matching in this study. According to the clustering property of the human skin color in the YCbCr color space, a Gaussian mixture model for facial color has been adopted to segment the regions which are close to the facial color from the image. In the segmented region, the location of the center point is realized, hence, achieve the purpose of detecting the outer face rapidly and overcome the disadvantages of the traditional morphological processing and edge tracking. The illumination compensation technique has overcome the impact from luminance during the detection and recognition process. Established an eyes-extracting-model to get the eyes' regions. A self-adaptive template has been adopted to overcome the impact from the skin-like background. Secondary matching algorithm has been applied to cut down the calculation amount and improve the detection speed during the matching process. The multi-scale singular-value- features which are extracted from the face images are used in the recognition process by face recognition algorithm based on template matching and achieve the purpose of reduce the dimension of the singular-value-feature matrix. By doing so, the local features extracted from the face image under the multi-scale can better reflect the differences between images, can more fully reflect the characteristics which are important for recognition, meanwhile, the template matching algorithm which has the better classification ability is adopted, so it is more suitable for pattern recognition. The experimental results show that, faces which have been rotated by large angle can also be detected; have strong adaptability under various illumination, poses and the background interference; the detection rate and recognition rate have both been largely promoted. The proposed approach is easy to manipulate and has a relatively faster speed which can satisfy the demand of real-time processing.

ACKNOWLEDGMENTS

This research was supported in part by a National Natural Science Foundation of China (No. 60873179), a Natural Science Foundation of Guangxi (No. 2011GXNSFA018162), a Research Fund for the Doctoral Program of Higher Education of China (No. 20090121110032), a Shenzhen Municipal Science and Technology Planning Program for Basic Research (No. JC200903180630A).

REFERENCES

Chao-Kuei, H., L. Shang-Hong and C. Yung-Chang, 2009. An optical flow-based approach to robust face recognition under expression variations. IEEE Trans. Image Process., 19: 233-240.

Chunhua, H. M. Xudong, D. Xianzhong and Q. Kun, 2007. Method for fast face-detection based on normalization Gaussian mixture model. J. Southeast Univ., 37: 389-394.

Gan, D.U. and Z. Wen-Jun, 2006. Face recognition method based on singular value decomposition and fuzzy decision. J. Image Graphics, 11: 1456-1459.

Greenspan, H., J. Goldberger and I. Eshet, 2001. Mixture model for face-color modeling and segmentation. Pattern Recognit. Lett., 22: 1525-1536.

Heusch, G. and S. Marcel, 2010. A novel statistical generative model dedicated to face recognition. Image Vision Comput., 28: 101-110.

Hsu, R.L., M. Abdel-Mottaleb and A.K. Jain, 2002. Face detection in color images. IEEE Trans. Pattern Anal. Mach. Intell., 24: 696-706.

Jun, L.X., W.D. Feng, Z.L. Fei, S.Y. Gang and Z.M. Yan, 2003. An approach for face recognition based on singular value decomposition and hidden Markov model. Chin. J. Comput., 26: 340-344.

Ksantini, R., B. Boufama, D. Ziou and B. Colin, 2010. A novel Bayesian logistic discriminant model: An application to face recognition. Pattern Recognit., 43: 1421-1430.

Qiao, L., S. Chen and X. Tan, 2010. Sparsity preserving projections with applications to face recognition. Pattern Recogn., 43: 331-341.

Queirolo, C.C., L. Silva, O.R.P. Bellon and M.P. Segundo, 2010. 3D face recognition using simulated annealing and the surface interpenetration measure. IEEE Trans. Pattern Anal. Mach. Intell., 32: 206-219.

Raghuraman, G. and J. David, 2010. Comparing and combining lighting insensitive approaches for face recognition. Comput. Vision Image Understanding, 14: 135-145.

Toews, M. and T. Arbel, 2009. Detection, localization and sex classification of faces from arbitrary viewpoints and under occlusion. IEEE Trans. Pattern Anal. Mach. Intell., 31: 1567-1581.

Wang, W., 2010. Research on face recognition technology based on average gray scale. Lect. Notes Electr. Eng., 56: 158-165.

Wang, Y. and Y. Wu, 2010. Complete neighborhood preserving embedding for face recognition. Pattern Recogn., 43: 1008-1015.

Wang, Z., S. Li, Y. Lv and K. Yang, 2010. Remote sensing image enhancement based on orthogonal wavelet transformation analysis and pseudo-color processing. Int. J. Comput. Intell. Syst., 3: 745-753.

Yunyang, Y., G. Zhibo and Y. Jingyu, 2008. Face recognition based on multi-scale singular value features. J. Tsinghua Univ., 3: 745-753.

Zhi-Wen, W., L. Shao-Zi, S. Song-Zhi, L. Mei-Zhen and C. Qi-Xian, 2010. A quickly-speed running watermarking algorithm based on quantum evolutionary algorithm. J. Optoelectron. Laser, 21: 737-742.

Zong, J.I.N., Y. Jing-yu, H.U. Zhong-shan and Z. Lou, 2001. Face recognition based on uncorrelated discriminant transformation. Pattern Recogn., 34: 1405-1416.