

Robust Face Recognition against Eyeglasses Interference by Integrating Local and Global Facial Features

Hansheng Fang¹, Jie Wen¹, Yong Xu^{1,2,3*}

¹ Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology,
518055 Shenzhen, China

² Key Laboratory of Network Oriented Intelligent Computation, 518055 Shenzhen, China

³Medical Biometrics Perception and Analysis Engineering Laboratory, 518055 Shenzhen,
China.

fanghansheng@163.com
ckghostwj@126.com
yongxu@ymail.com;

Abstract. In this paper, we proposed a feature extraction method to solve a challenge problem of face recognition, i.e., recognition of faces with eyeglasses. By fusing the local and global facial features, the proposed method can extract robust facial features that can greatly reduce the negative influence of eyeglasses on face recognition. Firstly, we use the Ununiformed Local Gabor Binary Pattern Histogram Sequence (ULGBPHS) method to extract local facial features. Secondly, we apply 2D-Discrete Fourier Transform (2D-DFT) method to obtain global facial features. Finally, we use a weighted fusion strategy to combine the two kinds of facial features for face recognition. Extensive experimental results on the well-known public GT and CMU_PIE face datasets, and real scene dataset which is built by our group show that the proposed feature extraction method obtains the best performance among some state-of-the-art methods. The relevant code and data will be available at <http://www.yongxu.org/lunwen.html>.

Keywords: face recognition, eyeglasses, local and global facial features

1 Introduction

Computer vision (CV) has become a hot research direction in recent years because of the advance of both related theories and computer hardware. There are many research areas in CV, such as the biometrics technology including fingerprint recognition, palm print recognition, face recognition etc. [1-3]. Among them, face recognition has received much attention in both academia and industry [4-5]. As we know, many external factors affect the actual recognition rate [6-7]. The eyeglass is a common problem in face recognition [8]. Suppose the user wears a pair of eyeglasses when he registers in a face recognition system, he may fail to pass the system when he doesn't wear eyeglasses in the recognition stage. In order to reduce the negative impact caused by eyeglasses, some algorithms have been proposed in the past.

Extracting more robust facial features is one of the most effective approaches to reduce the negative influence of eyeglasses. Martinez [9] proposed to extract the local facial features to solve the occlusion problem of face recognition. He divided a face into k blocks and calculated the feature for each block. After comparing the training set and testing set, he obtained the probability of occlusion for each block. These probability values were used to adjust the weight of each block when they use Mahalanobis distances in the recognition stage. Zhang et al. [10] proposed a Local Gabor Binary Pattern Histogram Sequence (LGBPHS) method to address the occlusion problems of face recognition. They applied Gabor filters and local binary pattern (LBP) operators [11-13] to a face image to extract the local facial features for face recognition. Yi et al. [14] performed the sparse representation based classification method on the extracted local features for face recognition. Liu et.al [15] proposed a novel ununiformed division strategy based on the LGBPHS method [10]. Experimental results in those papers all prove the effectiveness of the local features for eyeglass-face recognition.

In this paper, we try to solve the eyeglass-face recognition problem by extracting robust facial features [16-17]. The new feature extraction method can show outstanding performance in resisting the trouble caused by eyeglasses in a face. Firstly, we normalize the face image and extract its local facial features. We use non-uniform division strategy proposed in [15] to segment a face into several non-overlapping blocks with different sizes. For each block, Gabor filters with different scales and orientations are utilized to obtain multiple Gabor Magnitude Pictures (GMPs) [10, 18]. Then we compute the histogram for each GMP and concatenate these histograms as the local facial features for each facial block. In this way, we can obtain the local facial features of each face image by integrating all histograms of these blocks. Secondly, we use the 2D-DFT method to extract the global facial features. After transforming the face image into the frequency domain via the 2D-DFT transformation, we can obtain the real component and imaginary component of a face image. In this work, we only exploit the low frequency coefficients as the global facial features since they are the intrinsic global image information [19]. Finally, the extracted local and global facial features are combined via an adaptive weighted fusion approach for face recognition.

The remainder of this paper is organized as follows: Section 2 introduce a brief review of the local facial feature extraction method. The details of our proposed method are described in section 3. Section 4 presents the experimental evaluations. The conclusion and discussion of this paper are offered in section 5.

2 A Brief Review of Local Facial Feature Extraction Method

2.1 Facial Feature Extraction Using the LGBPHS Method

The eyeglasses can be viewed as occlusion of the face. Many excellent algorithms have been proposed and achieved a good performance in solving occlusion problems of face recognition. LGBPHS is one of the representative methods. It simultaneously combines the advantages of Gabor filters and LBP operators [20]. Gabor filters have powerful ability in obtaining robust and discriminative local features. LBP is a typical

visual descriptor and has been widely used in computer vision owing to its effectiveness in extracting the texture information of an image. The LGBPMS method uses Gabor filters to perform convolution operation on a normalized face image and gets plenty of GMPs in the first step. Then it uses the LBP operator to extract the LBP feature maps base on the obtained GMPs. In the second step, it divides these LBP feature maps uniformly and obtains the statistical histogram information of all blocks. Finally, it concatenates these feature histograms of all blocks into a feature histogram sequence as the final extracted local features. Experimental results show that this method is efficient for general occlusion problems.

2.2 Facial Feature Extraction Using the Ununiformed Division Strategy

Although the LGBPMS method shows good performance in general occlusion face recognition problems, it was proved to be inefficient in eyeglass-face recognition. As we know, the eyeglass-face problem is different from other general occlusion problems since eyeglasses always exists around our eyes. Human eyes are very important features in face recognition. The LGBPMS method uses a uniform way to divide a face into several blocks of the same size. This may break the integrity of eyes when we extract the facial features. In order to maintain the integrity of these facial keypoints, an improved non-uniform partition strategy called Ununiformed Local Gabor Binary Pattern Histogram Sequence (ULGBPMS) was proposed in [15]. This method can be viewed as extension of the LGBPMS method which uses a non-uniform partition strategy to extract more robust local features. The main idea of this method comes from the reality that the facial keypoints in the face should keep their own completeness when we divide the face feature map. Therefore, this method proposes to partition a normalized face image into different blocks non-uniformly. Then Gabor filters are applied to each block with different sizes to obtain their GMPs. Finally, it performs LBP on each group of GMPs to obtain the LBP feature maps. It gets the histograms of the local LBP feature maps for each group and concatenated them into a final feature histogram sequence by using a weighted strategy as the final features for face recognition [21-22]. Experimental results prove that the ULGBPMS method is effective to improve the accuracy in this problem. Although the non-uniform strategy is effective in handling eyeglass-face problems in comparison with the LGBPMS method, it doesn't always work well since it still focuses on extraction of facial local features.

3 The Proposed Feature Fusion Method

3.1 Method Analysis

Actually each face has its own overall appearance. Rather than focusing on local feature, we care about both local and global facial features in the eyeglass-face problems. In other words, we should take overall appearance into our consideration when extracting facial features. Usually humans can recognize a person correctly no matter whether he/she wears eyeglasses or not. This is because humans can recognize a face image as a whole at first glance and confirm their appearance by recognizing the local

facial keypoints. In this paper, we use the combination of local and global facial features instead of the local parts only. Specially we apply 2D-DFT method to extract the global facial feature of face image with and without eyeglasses. As a matter of fact, we can get two similar face fuzzy contours by using low frequency coefficients to reconstruct face images with and without eyeglasses. This proved the correctness of our idea.

3.2 The Procedure of the Proposed method

Firstly, we use the ULGBPMS method to extract local facial features. We use Gabor filters of different scales and orientations to perform convolution with the facial blocks. The convolution can be expressed by the following formulation:

$$G_{\mu,\nu}(x, y) = f(x, y) * \Psi_{\mu,\nu} \quad (1)$$

where $f(x, y)$ represents the pixel value of segmented face image block f and operator $*$ denotes the convolution operation. We call the $G_{\mu,\nu}(x, y)$ GMP after the convolution operation of face image block and Gabor filters. μ and ν represent the scales and orientations respectively. We use these GMPs in the next step to calculate their LBP feature maps by:

$$LGBP_{\mu,\nu}(x, y) = \sum_{p=0}^7 S(G_{\mu,\nu}(x_p, y_p) - G_{\mu,\nu}(x, y)) 2^p \quad (2)$$

where $LGBP_{\mu,\nu}(x, y)$ denotes each feature map after using LBP operator in each obtained GMP. (x, y) , (x_p, y_p) represent each pixel in a GMP and its neighbor pixels in this GMP respectively. S represents the binary pattern operation and it can be represented by the following equation:

$$S(x, y) = \begin{cases} 1, & G_{\mu,\nu}(x, y) > G_{\mu,\nu}(x_p, y_p) \\ 0, & G_{\mu,\nu}(x, y) \leq G_{\mu,\nu}(x_p, y_p) \end{cases} \quad (3)$$

The histogram of a feature map block for a face image f can be represented by histogram ranging in $[0..L-1]$ in the following format:

$$h = \sum_{x,y} \text{binary}\{LGBP_{split}(x, y) = i\}, i = 0, 1, \dots, L-1 \quad (4)$$

where h denotes the histogram for a feature map of each GMP. i refers to the gray level in this feature map and $LGBP_{split}(x, y)$ refers to the Local Gabor Binary Pattern (LGBP) feature map for each corresponding segmented face image block. The calculating procedure can be expressed by:

$$\text{binary}(A) = \begin{cases} 1, & A \text{ is true} \\ 0, & A \text{ is false} \end{cases} \quad (5)$$

We assume that we divide a face into m blocks, therefore the r -th block histogram is:

$$H_{\mu,\nu,r} = (h_{\mu,\nu,r,0}, \dots, h_{\mu,\nu,r,L-1}) \quad (6)$$

In this paper, we set μ equals to 5 and ν equals to 8 respectively. Therefore, we can obtain the local facial features by the following equation:

$$H_{local} = (H_{0,0,0}, \dots, H_{0,0,m-1}, H_{0,1,1}, \dots, H_{0,1,m-1}, \dots, H_{7,4,m-1}) \quad (7)$$

Secondly, we use 2D-DFT to extract global facial features. We can clearly find that the current algorithms only focus on extraction of the local facial feature. Usually we can recognize a person approximately in the first sight whether he wears eyeglasses or not. We just know the faces from the overall prospective. Therefore, we proposed our method by combining global and local facial features to recognize a person. For local facial features, we use ULGBP method. For global facial features, we use DFT method. This process of extracting global facial features can be shown in the following equation:

$$F(u', v') = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} g(x, y) e^{-j2\pi(\frac{u'x}{M} + \frac{v'y}{N})} \quad (8)$$

where g represents a $M \times N$ size face image, μ' and ν' refers to frequency variables. The output of the above formulation can be shown in the next equation:

$$F(u', v') = R(u', v') + jI(u', v') \quad (9)$$

$R(\mu', \nu')$ and $I(\mu', \nu')$ refers to the real part and imaginary part of $F(u', v')$. Via the DFT operation, each face can be converted into a real component and imaginary component in frequency domain. And we extract the low frequency part for global information of a face (including real part and imaginary part.).

We use $H_{global} = (H_R, H_I)$ to extract global facial information. H_R and H_I are feature vectors which can be calculated by:

$$H_R = \sum_{x,y} \text{binary}\{g_R(x, y) = i\}, i = 0, 1, \dots, L-1 \quad (10)$$

$$H_I = \sum_{x,y} \text{binary}\{g_I(x, y) = i\}, i = 0, 1, \dots, L-1 \quad (11)$$

$g_R(x, y)$ and $g_I(x, y)$ represent the real and imaginary magnitude pictures of a face image calculated by $R(\mu', \nu')$ and $I(\mu', \nu')$ respectively.

Finally, we use a weighted fusion approach to combine the two different kinds of facial features. We use D_{local}^k and D_{global}^k represent the Euclidean distance between the testing sample and training samples for local and global facial features respectively, where c refers to face classes and k refers to the training number of the face images in each class. We assume $c = 1, 2, \dots, C$, $k = 1, 2, \dots, K$ and the final distance between the testing sample and the training samples can be calculated by the following equation:

$$D_c^k = \omega_{local} * D_{local}^k + \omega_{global} * D_{global}^k \quad (12)$$

where ω_{local} and ω_{global} refer to the distance weight coefficients and $\omega_{local} + \omega_{global} = 1$. D_c^k refers to score level fusion of the distances calculated by two different feature vectors. The final classification result can be represented by:

$$t = \arg \min_c D_c^k \quad (13)$$

And then the testing sample is assigned to the t -th class, $t \in \{1, 2, \dots, C\}$. The Fig.1 shows the whole procedure of feature extraction.

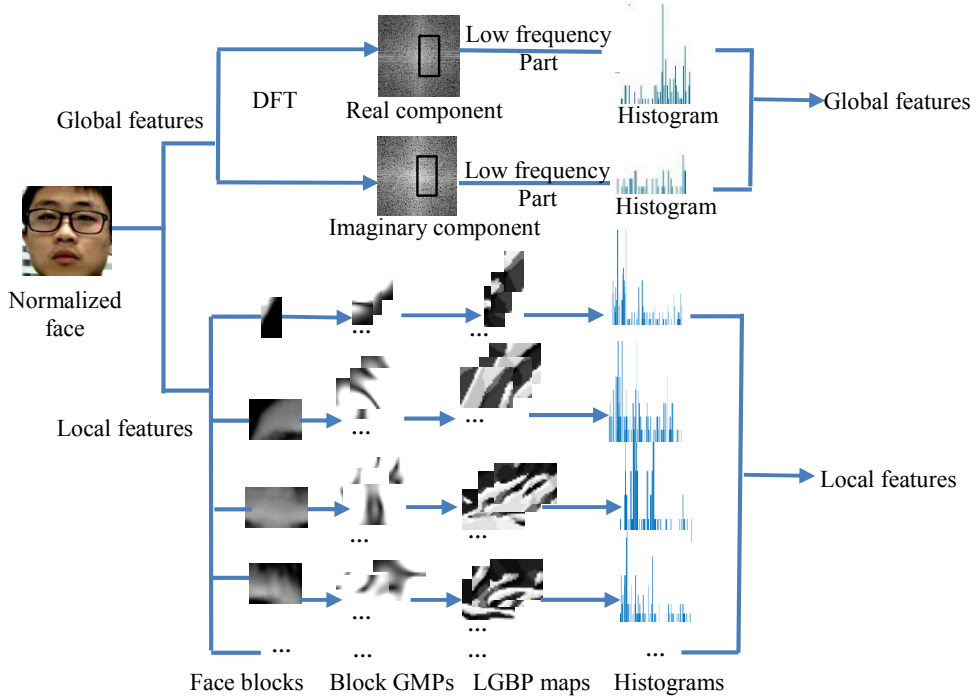


Fig. 1. Procedure of feature extraction using our proposed method

4 Experimental Evaluations

The datasets in our experiments come from two sources. Firstly, we select some typical face datasets and pick out the eyeglasses faces and non-eyeglasses faces from these original datasets. We choose the GT [23] and CMU_PIE [24] datasets for our experiments. In order to reduce the negative influence caused by some irrelevant factors, such as illumination, pose, facial expressions, we should keep these factors unchanged in the face images of each individual since they might affect the final recognition rate. We discard these unqualified face images of the original datasets before we do the experiments. After that we rename these remaining images in order to put images with eyeglasses in the beginning half part while these without eyeglasses are set in the ending half part. Since the individuals in GT dataset do not always wear eyeglasses, we select some of individuals which contain samples with and without eyeglasses. The GT dataset we selected consists of a total 330 images of 22 individuals. Due to the different poses in the CMU_PIE dataset, we divide them into five sub-datasets, which named CMU_PIE_pose05, CMU_PIE_pose07, CMU_PIE_pose09, CMU_PIE_pose27 and CMU_PIE_pose29. The CMU_PIE_pose05 consists of a total 1,120 images of 28 individuals. And 540 face images of 27 individuals from the CMU_PIE_pose07 are used for the experiments. 25 peoples with 20 images per person from the CMU_PIE_pose09 are used for our experiments. The CMU_PIE_pose27 we use contains a total 1,120 images of 28 individuals. In the CMU_PIE_pose29, we use a total 540 images of 27 individuals.

Secondly, we built a face dataset in real scene which is named BCC_Lab_Face. Our dataset consists of 50 peoples with 40 images per person. For each person, we collect four groups of face images. Two groups of face images are with eyeglasses while the other two are not. In the procedure of collection, we control the range of illumination, facial pose and facial expressions. Besides, we prepare several pairs of eyeglasses and each of them has different size and color. For a volunteer, if he wears eyeglasses, we collect the first group of face images and register another group of the face images by randomly selecting a pair of eyeglasses from our given eyeglasses. We also collect the other two groups of face images without eyeglasses which are controlled in the same outside scene. Fig. 2, Fig.3 and Fig.4 are some samples from our experimental datasets.



Fig. 2. One face sample from CMU_PIE selected dataset.



Fig.3. One face sample from GT selected dataset



Fig. 4. One face sample from BCC_Lab_Face.

Actually, we carry out our experiments in two cases. In the first case, we take the first several face images with eyeglasses as training samples, and the rest of them are treated as testing samples. In the second case, we take the last several face images without eyeglasses as training samples whereas the rest of face images are set as testing samples. In this way, we can validate that the features extracted by using our proposed method is robust for the eyeglass-face problem. In other words, we use a kind of facial features which can resist eyeglasses problems in face recognition. The details in setting training and testing sets will be discussed in Table 1. According to the setting in Table 1, we conduct two kinds of comparative experiments based on two different experimental cases. We compare our proposed method with other two previous representative methods called LGBPMS and ULGBPMS to show the superiority of our method. Moreover, we also use single features involved in our method as facial features to make comparison with our proposed method. Here, we use DFT, Gabor and LBP features as the single facial feature in comparative experiments. Fig.5 and Fig.6 show our experimental results on different datasets based on two different cases respectively. The different ID numbers in X axis show different datasets in our experiments and they represent CMU_PIE_pose05, CMU_PIE_pose07, CMU_PIE_pose09, CMU_PIE_pose27, CMU_PIE_pose29, GT, BCC_Lab_Face correspondingly. The Y axis shows different recognition rates of different methods. Since we mainly focus on feature extraction, we use the Euclidean distance as a common measure of standards in final recognition stage for these different methods. In this way, we can compare these different methods with our proposed methods. Fig.5 and Fig.6 indicate that our method is the best one among these different methods in both two different testing cases. The light blue line with diamond shape shows the accuracy of our method on different datasets. We can clearly find that our method has a higher recognition rate on different datasets and it also shows stronger stability on different testing datasets.

Table 1. Training and testing sets division.

Experimental Cases	Case1: First several faces with eyeglasses as training sets		Case2: Last several face images without eyeglasses as training sets	
	Training sets	Testing sets	Training sets	Testing sets
CMU_PIE_pose05	10	30	10	30
CMU_PIE_pose07	10	10	10	10
CMU_PIE_pose09	10	10	10	10
CMU_PIE_pose27	10	30	10	30
CMU_PIE_pose29	10	10	10	10
GT	7	8	7	8
BCC_Lab_Face	10	30	10	30

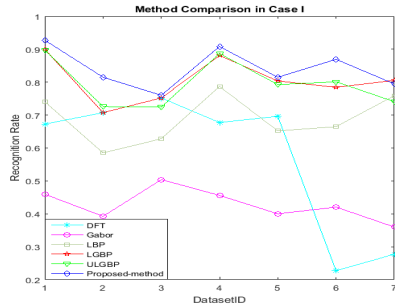


Fig. 5. Experimental results in Case I

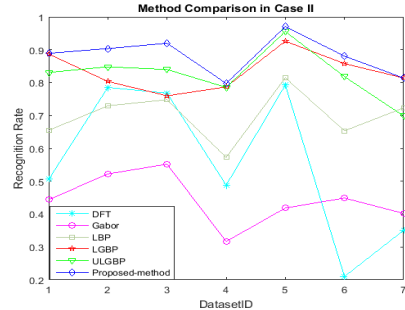


Fig. 6. Experimental results in Case II.

5 Conclusion and Discussion

This paper proposed a novel method to solve the challenge problem of recognition of eyeglass-faces. Compared with other methods, the proposed method simultaneously takes into account the local and global facial features. In other words, the proposed method can extract more discriminant features from the face, which enable it to obtain the best performance. Experimental results on the GT, CMU_PIE, and our collected BCC_Lab_Face datasets powerfully prove the effectiveness of the proposed method.

In the future, we will try to employ better classification strategies in recognition stage to improve the overall performance of the proposed method in eyeglass-face recognition problem.

Acknowledgements. This paper is partially supported by Guangdong Province high-level personnel of special support program (No. 2016TX03X164).

References

1. Coetzee L. Fingerprint Recognition. In: International Conference on Internet Monitoring and Protection. IEEE, 1996:32-32.
2. Kong A, Zhang D, Kamel M. A survey of palmprint recognition. J. Pattern Recognition, 2009, 42(7):1408-1418.
3. Zhao W, Chellappa R, Phillips P J, et al. Face recognition:A literature survey. J. Acm Computing Surveys, 2003, 35(4):399-458.
4. Arya S, Pratap N, Bhatia K. Future of Face Recognition: A Review. J. Procedia Computer Science, 2015, 58(2):578-585.
5. Solanki K, Pittalia P. Review of Face Recognition Techniques. J. International Journal of Computer Applications, 2016, 133.
6. Xu Y, Fang X, Li X, et al. Data Uncertainty in Face Recognition. J. IEEE Transactions on Cybernetics, 2014, 44(10):1950-1961.

7. Givens G, Beveridge J R, Draper B A, et al. How Features of the Human Face Affect Recognition: A Statistical Comparison of Three Face Recognition Algorithms. In: Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on. IEEE, 2004:II-381-II-388 Vol.2.
8. Yi D, Li S Z. Learning sparse feature for eyeglasses problem in face recognition. In: IEEE International Conference on Automatic Face & Gesture Recognition and Workshops. IEEE, 2011:430-435.
9. Martínez A M. Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class. J. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 24(6):748-763.
10. Zhang W, Shan S, Gao W, et al. Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A Novel Non-Statistical Model for Face Representation and Recognition. In: Tenth IEEE International Conference on Computer Vision. IEEE Computer Society, 2005:786-791.
11. Ojala T, Harwood I. A Comparative Study of Texture Measures with Classification Based on Feature Distributions. J. Pattern Recognition, 1996, 29(1):51-59.
12. Ojala T, Pietikäinen M, Mäenpää T. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. J. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 24(7):971-987.
13. Ahonen T, Hadid A, Pietikäinen M. Face Recognition with Local Binary Patterns. In: Computer Vision - ECCV 2004, European Conference on Computer Vision, Prague, Czech Republic, May 11-14, 2004. Proceedings. 2004:469-481.
14. Yi D, Li S Z. Learning sparse feature for eyeglasses problem in face recognition. In: IEEE International Conference on Automatic Face and Gesture Recognition. 2011:430-435.
15. Liu L, Sun Y, Yin B, et al. A Novel Nonuniform Division Strategy for Wearing Eyeglasses Face Recognition. In: International Conference on Image and Graphics. IEEE, 2009:907-911.
16. Deng Y, Guo Z, Chen Y. Fusing Local Patterns of Gabor and Non-subsampled Contourlet Transform for Face Recognition. In: 2013 Second IAPR Asian Conference on Pattern Recognition, 2013:482-485
17. Zanchettin C. Face recognition based on global and local features. In: ACM Symposium on Applied Computing. ACM, 2014:55-57.
18. Yang P, Shan S, Gao W, et al. Face Recognition Using Ada-Boosted Gabor Features. In: IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings. IEEE Xplore, 2004:356-361.
19. Yu S U, Shan S G, Chen X L, et al. Integration of Global and Local Feature for Face Recognition. J. Journal of Software, 2010, 21(8):1849-1862.
20. Zhang W, Shan S, Chen X, et al. Local Gabor Binary Patterns Based on Kullback–Leibler Divergence for Partially Occluded Face Recognition. J. IEEE Signal Processing Letters, 2007, 14(11):875-878.
21. Guo K, Wu S, Xu Y. Face recognition using both visible light image and near-infrared image and a deep network. J. Caa Transactions on Intelligence Technology, 2017, 2(1):39–47.
22. Xu Y, Lu Y. Adaptive weighted fusion: A novel fusion approach for image classification. J. Neurocomputing, 2015, 168:566-574.
23. “Georgia Tech Face Database,” http://www.anefian.com/research/face_reco.html.
24. “The CMU Multi-PIE Face Database,” <http://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/Home.html>.