Face Recognition by Fusing Binary Edge Feature and Second-order Mutual Information

Jiatao Song, Beijing Chen, Wei Wang, Xiaobo Ren
School of Electronic and Information Engineering,
Ningbo University of Technology
Ningbo, China

***Abstract*—Illumination change and expression variation are two main factors affecting the performance of some existing face recognition algorithms. Edge feature based methods are robust to illumination change and are easy to implement. But they don't work well for the images with expression variation. In this paper, a novel face recognition method based on the fusion of binary edge feature and grayscale information is proposed to improve both the illumination and expression robustness. The second-order mutual information (MI_2) is introduced as the similarity metric of grayscale face images. Experimental results on AR dataset and Yale dataset, both with illumination and expression changes, show that the overall face recognition rate of the proposed method is better than that of some commonly used approaches, indicating that our method is more effective for practical uses.**

***Keywords*—face recognition, illumination change, expression change, feature fusion, binary edge feature, grayscale feature**

I. INTRODUCTION

Human face recognition is one of the most direct technologies for personal identification and is the hotspot of pattern recognition and computer vision. It has many potential applications, such as bankcard identification, access control, security monitoring and surveillance. For the past two decades, many face recognition approaches have been proposed. Some of them were put into practical use [1-3].

Illumination change is one of the main factors affecting the performance of some existing face recognition algorithms. In order to reduce the influence of lighting variation, many new face recognition algorithms have been proposed in the past several years These methods can be roughly classified into three categories: statistical methods [4-6], model-based methods [7-10] and edge-based methods [11-13]. Statistical approaches, such as Eigenface method (also named PCA method) [4], Fisherface method [5], Bayesian-based methods [6] and so on, are easy to implement, but they need a number of images for each subject as training images to cover different illumination conditions, while this is often impractical. Model-based methods exploited the fact that the set of images of an object in fixed pose is a convex cone in the space of images. By using a small number of training images per subject taken from different lighting condition, a generative model of the face can be reconstructed. For each pose, the corresponding illumination cone was approximated by a low-dimensional linear subspace and the basis vectors can be estimated using the reconstructed model. In order to construct the illumination cone, model-based methods also require a set of training images for each subject. Further more, because of the complexity of light sources and illumination changes, how to compute and obtain physically implemented lower-dimensional subspaces basis images requires deep investigation [7-10]. Edge reflects the discontinuity of intensity distribution in an image. It contains contour, structure and shape information of objects in an image. Cognitive psychological studies indicated that human beings recognize line drawings as quickly and almost as accurately as gray-level pictures. Further more, edge is insensitive to illumination changes. Based on edge features, B.Takacs et al. proposed an algorithm using Doubly Modified Hausdorff Distance (M2HD) [11], Y. Gao et al. used line edge to represent face and defined Line Hausdorff Distance (LHD) as similarity metric to match two line edge images [12]. Recently, J. Song developed a face recognition method based on binary template matching [13]. Experimental results show that the face recognition rates of all these edge-based methods are preferable, especially in varied lighting cases.

Comparing with other two kinds of methods, edge-based face recognition approaches extract face features directly from testing images, without using of training images, hence this kind of methods is easier to implement in practical applications. While researches [12-13] show that although edge-based methods can obtain better illumination robustness, they don't work well on the images with facial expression changes, especially with exaggerated expression changes, such as scream, which limits its application. In this paper, a novel face recognition method based on the fusion of binary edge and grayscale information is proposed to improve both the illumination robustness and expression robustness. Experimental results on Yale face database [14] and AR database [15] show that our method is effective for images with both illumination changes and facial expression variations.

II. THE PROPOSED METHOD

A. Image Similarity Metric Based on Single Feature

1) Edge-based similarity metric—binary edge distance

Let $BTaI$ and $BTeI$ be two binary edge images to be matched, N_{TaI} and N_{TeI} be the number of foreground pixels in $BTaI$ and $BTeI$ respectively, and $N_{overlap}$ be the number of common foreground pixels in $BTaI$ and $BTeI$. The similarity

978-1-4244-1674-5/08 /$25.00 ©2008 IEEE

between these two binary edge images, namely Binary Edge Distance (BED), is defined as [13]:

$$BED(BTaI, BTeI) = \frac{2N_{overlap}}{N_{TaI} + N_{TeI}} \quad (1)$$

Obviously, the value of $BED(BTaI, BTeI)$ is between 0 and 1. If two images are the same, $N_{overlap}=N_{TaI}=N_{TeI}$, so $BED(BTaI, BTeI) = 1$. If there is no overlapped foreground pixels, then $N_{overlap}=0$ and $BED(BTaI, BTeI) = 0$. Thus, the greater the value of $BED(BTaI, BTeI)$, the more similar the two matched images.

2) Grayscale-based similarity metric — second-order mutaual information distance

Mutual information (MI) is a basic concept of Information Theory. It measures the statistical correlation of two random variables. It also reflects the amount of information of one variable carried by another variable. The mutual information of two images A and B, denoted with *MI(A,B)*, can be defined in terms of the first-order entropies of these two images, denoted with *H(A)* and *H(B)*, and their joint entropy *H(A,B)*, i.e

$$MI(A, B) = H(A) + H(B) - H(A, B) \quad (2)$$

The first-order entropy *H(A)* is often calculated using the grayscale probability distribution function $p(a)$ of each pixel in the image, i.e.,

$$H(A) = -\sum_a p(a) \log p(a) \qquad a \in A \quad (3)$$

But this may lead to the overestimate of the true entropy of an image, because it is based on the assumption that the intensity values of each pixels in an image are probability-independent, which is often not true in practice. To address this problem, Rueckert et al.[16] proposed a new similarity metric based on second-order entropy and second-order mutual information (MI_2). The second-order entropy of image *A*, denoted with $H_2(A)$, can be defined as

$$H_2(A) = -\sum_i \sum_j p(i, j) \log p(i, j) \quad (4)$$

where *p(i, j)* denotes the joint probability that a pixel has intensity value of *i* while its neighboring pixel has intensity value of *j*. The value of *p(i, j)* can be estimated from 2D joint histogram. In order to compute MI_2, the second-order joint entropy should be firstly calculated, it can be defined as follows:

$$H_2(A,B) = -\sum_i \sum_j \sum_k \sum_l p(i,j,k,l) \log p(i,j,k,l) \quad (5)$$

where $p(i,j,k,l)$ denotes the joint probability that a pixel and its neighbor in image A have intensity values of i and j, respectively, while the corresponding pixel and its neighboring pixel in image B have intensity values of k and l , respectively. $p(i,j,k,l)$ can be estimated from 4D joint histogram. Like first-order MI, the second-order mutual information, denoted with $MI_2(A,B)$, can be calculated with

$$MI_2(A,B) = H_2(A) + H_2(B) - H_2(A,B) \quad (6)$$

It can be further normalized as

$$NMID_2(A,B) = \frac{H_2(A) + H_2(B)}{H_2(A,B)} \quad (7)$$

In this paper, the normalized second-order mutual information distance $NMID_2(A,B)$ defined in (7) is used as the similarity metric of two images. Obviously, it is based on the grayscale feature of images. The intensity value j of the neighboring pixel in (4) and (5) is estimated using the mean value of four-neighboring pixels [17].

B. Face Recognition Based on the Fusion of Binary Edge and Grayscale Features

There exist three kinds of information fusion methods, i.e. fusion at feature level, fusion at matching score level and fusion at decision level. In this paper, the third kind information fusion method is adopted for our face recognition purpose. The detailed feature fusion scheme is illustrated in Fig 1.

Let TeI be grayscale image to be tested, TaI_i $(i = 1,2,\cdots,n)$ be target grayscale images, n is the number

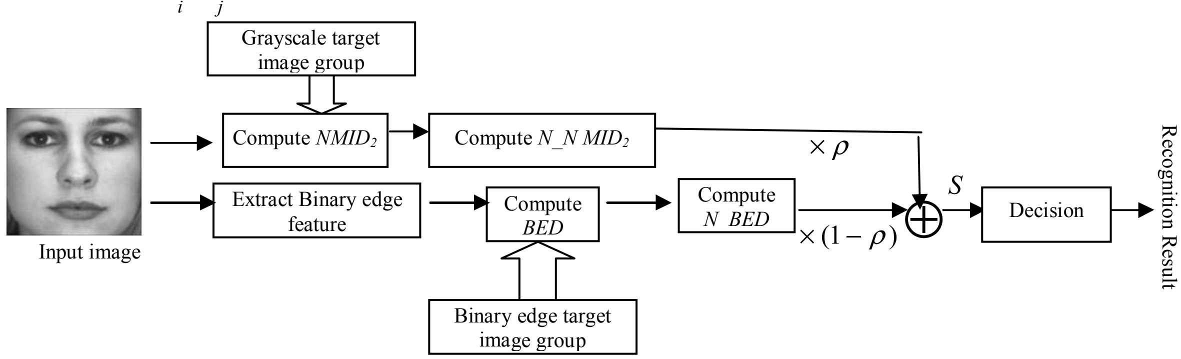

Figure 1. Face recognition based on the fusion of edge and grayscale features

of target images. Let $BTeI$ and $BTaI_i$ be the corresponding binary edge image of TeI and TaI_i, respectively, $NMID_2(TeI,TaI_i)$ be the MI_2 distances between TeI and TaI_i calculated using (7), $BED(BTeI,BTaI_i)$ be the BED distances between $BTeI$ and $BTaI_i$ calculated using (1),. $N\_NMID_2(TeI,TaI_i)$ and $N\_BED(BTeI,BTaI_i)$ be the improved version of $NMID_2(TeI,TaI_i)$ and $BED(BTeI,BTaI_i)$, respectively, with their values normalized to [0,1]. Then, The similarity metric by fusing the binary edge and grayscale features, denoted with $S(TeI,TaI_i)$, is defined as follows:

$$\begin{aligned} S(TeI,TaI_i) &= \rho\times N\_NMID_2(TeI,TaI_i) \\ &\quad +(1-\rho)\times N\_BED(BTeI,BTaI_i) \\ &\quad (i=1,2,\cdots,n) \end{aligned} \tag{8}$$

where ρ is an adaptive weight factor defined as follows:

$$\rho = f(\omega_i) = \frac{1}{1+e^{-5(\omega_i-\omega_0)}} \tag{9}$$

where

$$\omega_i = \omega_{i1} \Big/ (\omega_{i1}+\omega_{i2})$$

$$\begin{aligned} \omega_{i1} &= \frac{|BED(BTeI,BTaI_i)-BED(BTaI_i,BTaI_i)|}{\max\limits_{1\le i\le n}(|BED(BTeI,BTaI_i)-BED(BTaI_i,BTaI_i)|)} \\ &= \frac{|BED(BTeI,BTaI_i)-1|}{\max\limits_{1\le i\le n}(|BED(BTeI,BTaI_i)-1|)} \end{aligned}$$

$$(\because BED(BTaI_i,BTaI_i)=1)$$

$$\begin{aligned} \omega_{i2} &= \frac{|NMID_2(TeI,TaI_i)-NMID_2(TaI_i,TaI_i)|}{\max\limits_{1\le i\le n}(|NMID_2(TeI,TaI_i)-NMID_2(TaI_i,TaI_i)|)} \\ &= \frac{|NMID_2(TeI,TaI_i)-2|}{\max\limits_{1\le i\le n}(|NMID_2(TeI,TaI_i)-2|)} \end{aligned}$$

$$(\because NMID_2(TaI_i,TaI_i)=2)$$

$$\omega_0 = \underset{1\le i\le n}{mean}(\omega_i)$$

where $mean(\,)$ is a mean value function. The matching result of tested image TeI, denoted with TaI_R, is determined by the following rule:

$$R = \arg\max_{1\le i\le n}(S(TeI,TaI_i)) \tag{10}$$

III. EXPERIMENTAL RESULTS AND ANALYSIS

The AR face database [14] and the Yale face database [15] are used for our face recognition experiments. For each subject, only one image is selected as the target image. All the images including those in the target group were manually normalized in geometry and grayscale using the method described in [18]. The distance between two irises is set to 60 pixels and the resolution of the normalized images is 120×120. The binary edge features are extracted using the LAT algorithm proposed in [19]. Some normalized face images and their corresponding binary maps are shown in Fig. 2 and Fig. 3.

The AR image set used includes all 931 images stored on the first four CD- ROMs of the AR database. These images were taken from 133 subjects with three lighting variations (left light on, right light on, and all side lights on) with neutral facial expression, and four facial expression variations (neutral, smile, scream and anger) under normal lighting condition. At testing stage, 133 images with neutral expression and normal illumination condition are selected as target images. The rest 798 images are divided into 6 groups. Each group was tested individually.

The Yale image set used consists of the 135 images of 11 subjects with different illumination and different expression conditions. All these images are divided into 9 groups, corresponding to different lighting conditions and facial expressions. Each group has 15 images. The group with normal illumination and neutral expression is chosen as target group. The other eight groups, corresponding to three variations in lighting condition (center light, left light, right light) and five facial expression variations (happy, sad, sleepy, surprised, wink), are selected as test images.

TABLE Ⅰ and TABLE Ⅱ show the face recognition results of five different methods on the AR images and Yale images, respectively. These approaches are: (1) the BED-based method, (2) the $NMID_2$-based method, (3) the proposed feature fusion based method, (4) the M2HD method and (5) the PCA method.

From TABLE Ⅰ and TABLE Ⅱ, we can conclude that:

a) For images with lighting changes, the BED based method achieves better average face recognition rates than another binary edge feature based M2HD method on both the AR subset and Yale subset used. Also, the $NMID_2$ based method obtain better performance than another grayscale feature based PCA method on these images.

b) For images with expression changes, on both the AR subset and Yale subset, the edge based methods, including M2HD approaches and BED based method, are not as good as the grayscale based methods, such as the $NMID_2$ based method and the traditional PCA method. This is coincide with the results reported in [12] and [13].

c) The proposed feature fusion based method achieves an overall face recognition rate of 87.72% on AR images and that of 84.17% on Yale images. Both are better than the face recognition results obtained using other methods. This indicates that by combing the edge feature and grayscale feature together, the proposed method can achieve both better lighting robustness and better expression robustness, leading it more applicable in practical uses.

IV. CONCLUSION

The illumination change and expression variation are two

Figure 2. Example of normalized face image (top: AR images, bottom: Yale images)

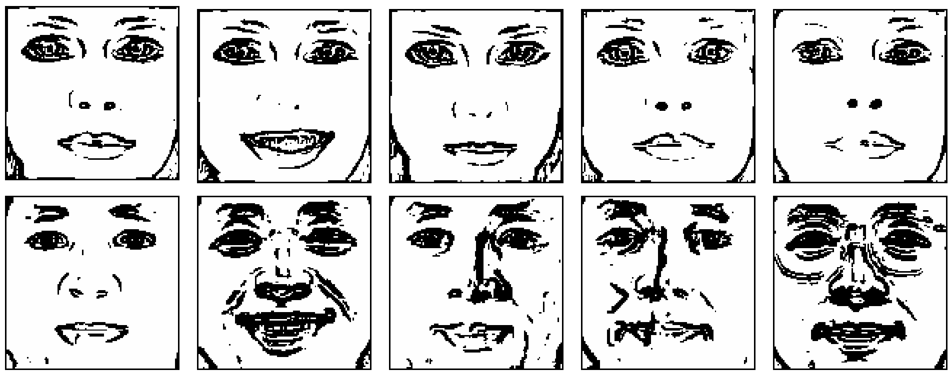

Figure 3. Corresponding binary edge images of Fig.2 (top: AR images, bottom:Yale images)

TABLE I. FACE RECOGNITION RATES OF DIFFERENT METHODS ON AR IMAGES (%)

Methods	**Different lighting**				**Different expression**				**Overall**
	left light on	***right light on***	***all side lights on***	***average***	***Smile***	***Anger***	***Scream***	***average***	
BED	91.73	92.48	87.22	90.48	87.97	72.18	26.32	62.16	76.32
NMID$_2$	95.49	98.50	87.97	93.99	93.98	93.98	50.38	79.45	86.72
Proposed method	**98.50**	**98.50**	**93.98**	**96.99**	**95.49**	**91.73**	**48.12**	**78.45**	**87.72**
M2HD	89.47	90.98	66.17	82.21	73.68	75.94	17.29	55.64	68.93
PCA	66.17	51.88	77.44	65.16	90.98	80.45	42.86	71.43	68.30

TABLE II. FACE RECOGNITION RATES OF DEFFERENT METHODS ON YALE IMAGES (%)

Methods	**Different lighting**				**Different expression**						**Overall**
	center-light	***left-light***	***Right-light***	***average***	***happy***	***sad***	***sleepy***	***surprised***	***wink***	***average***	
BED	80.00	93.33	93.33	88.89	86.67	86.67	60.00	86.67	73.33	78.67	82.50
NMID$_2$	80.00	33.33	66.67	60.00	100.00	93.33	86.67	100.00	80.00	92.00	80.00
Proposed method	**93.33**	**66.67**	**80.00**	**80.00**	**100.00**	**86.67**	**73.33**	**93.33**	**80.00**	**86.67**	**84.17**
M2HD	86.67	86.67	86.67	86.67	73.33	93.33	73.33	73.33	86.67	80.00	82.50
PCA	73.33	26.67	53.33	51.11	100.00	93.33	66.67	86.67	86.67	86.67	73.34

urgent problems to be solved for face recognition nowadays.

Edge feature based methods have good illumination robustness, but they are less robust to expression change than grayscale feature based methods. Considering the limitations of single edge feature or grayscale feature based methods, a novel method by fusing edge feature and grayscale feature is presented in this paper. Experimental results on both AR images and Yale images show that the proposed method achieves better overall face recognition rates than other commonly used edge based or grayscale based method

The future work is to develop more effective feature fusion scheme to further improve the face recognition performance.

ACKNOWLEDGMENT

The work described in this paper was partially supported by a grant from Ningbo Natural Science Foundation (2006A610015 and 2007A610047) and two grants from Zhejiang Provincial Natural Science Foundation of China (Y105239 and 602118). The author also gratefully acknowledges the support of K. C. Wong Education, Hong Kong.

REFERENCES

[1] R. Chellappa, C. L. Wilson, and S. Sirohey, “Human and machine recognition of faces: A survey,” Proceedings of the IEEE, vol. 83, no. 5, pp. 705-740, 1995.

[2] W. Y. Zhao, R. Chellappa, A. Rosenfeld, and J. Phillips, “Face recognition: A literature Survey,” ACM computing Survey, vol. 35, no. 4, pp. 399-458, 2003.

[3] P. J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, M. Bone, “Face recognition vendor test 2002: Evaluation report,”. Technical Report, NISTIR 6965, Gaithersburg. National Institute of Standards and Technology, 2003.

[4] M.Turk, A.Pentland, “Eigenfaces for recognition. Journal of Cognitive,” Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

[5] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, 1997.

[6] B. Moghaddam, T. Jebara, and A. Pentland, “Bayesian face recognition,” Pattern Recognition, vol. 33, no.11, pp. 1771-1782, 2000.

[7] P. Belhumeur, D. Kriegman, “What is the set of images of an object under all possible lighting conditions,” International Journal of Computer Vision, vol. 28, no. 3, pp. 245-260, 1998.

[8] R. Basri, D. Jacobs, “Lambertian reflectance and linear subspaces,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 2, pp. :218-233, 2003.

[9] K. Lee, J. Ho, and D. Kriegman, “Acquiring linear subspaces for face recognition under variable lighting,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 5, pp. 684-698, 2005.

[10] L. Zhang, D. Samaras, “Face recognition from a single training image under arbitray unknown lighting using spherical harmonics,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 3, pp. 351-363, 2006.

[11] B. Takacs, “Comparing face images using the modified Hausdorff Distance,” Pattern Recogni- tion, vol. 31, no. 12, pp. 1873-1881,1998.

[12] Y. Gao, M.K.H.Leung, “Face recognition using line edge map,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 6, pp. 764-779, 2002.

[13] J. Song, B. Chen, and Z. Chi, X. Qiu, and W. Wang, “Face recognition based on binary template matching,” In Proceedings of the Third International Conference on Intelligent Computing(ICIC2007), Springer Publishers, Berlin/ Heidelberg, pp. 1131-1139, August, 2007.

[14] A.M. Martinez and R. Benavente, “The AR Face Database,” CVC Technical Report #24, June 1998.

[15] P. N. Belhumeur, “Yale face database,” http://cvc.yale.edu/projects/yalefaces/yalefaces.html, 2002.

[16] D. Rueckert, M. J. Clarkson, D. Hill, and D. Hawkes, “Non-rigid registration using higher- order mutual information,” In Proceedings of SPIE, Medical Imaging: Image Processing, vol. 3979, pp. 438-447, 2000.

[17] B. Chen, J. Li, and G. Chen, “Study of medical image registration based on second-order mutual information,” In Proceedings of the 1st IEEE International Conference on Bioinformatics and Biomedical Engineering(ICBBE2007),IEEE Xplore, pp. 956-959, 2007.

[18] Z. Bian, X. Zhang, “Pattern Recognition,” 2nd Edn. Tsinghua University Press, Beijing, 2000 (in Chinese).

[19] S. Y. Lee, Y. K. Ham, R.H Park, “Recognition of human front faces using knowledge-based feature extraction and neuro-fuzzy algorithm,” Pattern Recognition, vol. 29, no. 11, pp. 1863-1876,1996.