

RECOGNITION OF EXPRESSION VARIANT FACES FROM ONE SAMPLE IMAGE PER ENROLLED SUBJECT

Hamidreza Rashidy Kanan¹ and Yongsheng Gao²

¹Electrical and Computer Engineering Department, Islamic Azad University, Qazvin Branch, Qazvin, Iran

²School of Engineering, Griffith University, 170 Kessels Road, Nathan, Brisbane, QLD 4111, Australia
rashidykanan@aut.ac.ir, yongsheng.gao@griffith.edu.au

ABSTRACT

Despite remarkable progress on face recognition, little attention has been given to robustly recognize expression variant faces from a single sample image per person. One way to deal with the recognition of faces under above conditions is by using local statistical approaches which appear to be more robust against variations in facial expression. In this paper, we propose a new weighted matching method based on our recent work of AWPPZMA to recognize expression variant faces when only one exemplar image per enrolled subject is available. The proposed weighting method gives more significance to those parts of the face with facial expression variations that change less compared to neutral face image and less significance to those parts that change more. In this contribution, we use the difference between local area in the input face and its corresponding local area in the neutral face image as a measure of observable structure changes. The encouraging experimental results demonstrate that the proposed method provides a new solution to the problem of robustly recognizing expression variant faces in single model databases.

1. INTRODUCTION

Automatic human face recognition has received significant attention for decades due to its numerous potential applications including personal identification, bankcard verification, automated surveillance and intelligent human-computer interaction. Despite remarkable progress so far, general face recognition task is still a challenging problem. The robustness to facial expression variations is an important problem in a face recognition system. This problem can be expressed as follows: “how can we robustly recognize (identify) a person’s face for whom the training and the testing images differ in facial expression?”. Although numerous approaches have been proposed on face recognition, limited studies have been reported on robustly

recognition of expression variant faces from a single sample image per person. Notably, Beymer and Poggio [1] proposed to morphed the testing images to equal in shape the training one. However, as stated by the authors themselves, this is not always possible, because some texture information might be missing in the testing image (e.g. closed eyes). Martinez [2], has proposed an approach that couples the geometric (shape) information of the face with the texture one. In his formulation only a single sample image per class is required. In another paper [3], he has proposed a weighting method based on the optical flow between the training and the testing image as a measure of visible texture changes to robustly recognize expression variant faces from a single sample image per class.

In a recent work [4], we proposed an Adaptively Weighted Patch Pseudo Zernike Moment Array (AWPPZMA) approach for face representation and recognition. Based on this approach, in this paper, we propose a new weighting method to recognize expression variant faces from a single exemplar image per enrolled subject. The proposed weighting method yields more importance to those areas of the face with facial expression variations that change less compared to neutral face image and less importance to those areas that change more. In a scenario of face recognition under uncontrolled condition, different emotions are represented by facial expressions with more emphasis on specific areas of the face than other areas [5]. Therefore, an expression variant face will be easier to recognize using local-based approach and the recognition will be more robust to facial expressions using a weighting method that determines which areas of the face change more and which areas change less (between the gallery and the probe images). In this paper, these changes are detected by a comparison of each local area in the input face image with the corresponding local area at the same position in the neutral face image.

A feasibility and effectiveness investigation for the proposed method is conducted using the standard AR face database [6] which contains images of the frontal view faces of 126 people. The system performance is compared with the performances of three benchmark approaches. Our

experimental results show that the proposed weighting method can improve the performance of single model-based face recognition system when the input face image has facial expression variations.

2. THE PROPOSED APPROACH

As illustrated in Figure 1, there are three main steps in the proposed algorithm: (1) image partitioning and feature extraction, (2) creation of contribution maps and (3) similarity measurement. In this algorithm a face image is partitioned into a set of equal-sized patches in a non-overlapping way. Then Patch Pseudo Zernike Moment (PPZM) is extracted from each patch of the partitioned face image which will be concatenated to form a Patch Pseudo Zernike Moments Array (PPZMA) of the face image. On the other hand, a Contribution Maps (CMs) are obtained from the input face and all PPZMAs extracted from neutral face images (dataset of PPZMA) which will be used to weight the extracted PPZMA in the matching process. Finally, the input face image can be identified using WPPZMA matching score. In the following, we will describe the algorithm in detail.

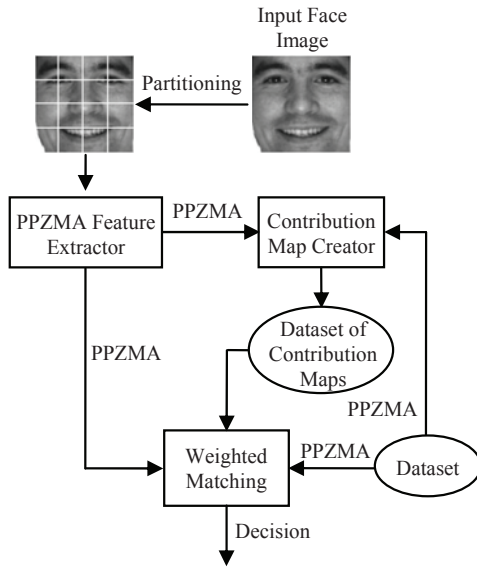


Fig. 1: Schematic illustration of the proposed approach

2.1. Image Partitioning and Feature Extraction

In the proposed algorithm, a face image $f(x, y)$ with a dimension of $N \times N$ is partitioned into a set of equal-sized $W \times W$ patches in a non-overlapping way. These patches can be mathematically represented as:

$$f_{p,q}(a,b) = f(W(p-1)+a, W(q-1)+b) \quad (1)$$

where $f_{p,q}(a,b)$ denotes a patch at the position of (p, q) . a and b are integers ranging from 1 to W while p and q are integers ranging from 1 to $\frac{N}{W}$.

The PZM is a widely used orthogonal moment due to its capability for image representation and robustness to image noise [4]. Basic PZM operation on whole face images to extract the holistic information is not effective under different facial expression changes. In this paper, we use the Patch Pseudo Zernike Moment (PPZM) which performs on a partitioned image to form a PPZM array of the face image.

The discrete form of PPZM of order n with repetition m for mapped digital image intensity $f(x_j, y_i)$ can be written as [4]:

$$PPZM_{n,m}^{p,q}(f(x_j, y_i)) = \frac{2(n+1)}{\pi W^2} \sum_{i=0}^{W-1} \sum_{j=0}^{W-1} V_{n,m}^*(x_j, y_i) f(W(p-1)+x_j, W(q-1)+y_i) \quad (2)$$

As used in [4], the magnitudes of PPZMs of order $n = 1$ up to $n_{\max} = 3$ with $m \geq 0$ will be considered as moment features in this paper. Therefore, the feature vector for patch $f_{p,q}(a,b)$ can be represented as:

$$PPZM [f_{p,q}(a,b)] = \left\{ |PPZM_{u,v}^{p,q}(f(x_j, y_i))| \mid u = 1, \dots, n_{\max}, v = 0, 1, \dots, u \right\} \quad (3)$$

Finally, the extracted PPZMs will be concatenated to form a PPZM Array (PPZMA) of the face image as:

$$PPZMA [f(x, y)] = \left\{ PPZM [f_{p,q}(a,b)] \mid p, q = 1, 2, \dots, \frac{N}{W} \right\} \quad (4)$$

2.2. Creation of Contribution Map (CM)

In this paper, we propose a new local patch-based weighting method to handle the expression variance problem. Our weighting method proposed here yields more significance to those areas of the face with facial expression variations that change less compared to neutral face image and less significance to those areas that change more. These changes can be detected by a comparison of each local area in the input face image with the corresponding local area at the same position in the neutral face image. In other words, the degree of change at each patch in the input face image compared to the corresponding patch in the i^{th} neutral face image in the database can be measured by the difference between the PPZMs of the patch in the input face and the PPZMs of its corresponding patch in the mentioned neutral face as follows:

$$DM_{p,q}^i(a,b) = \left\| \overline{PPZM}^{n_{\max}}[f_{p,q}(a,b)] - \overline{PPZM}^{n_{\max}}[h_{p,q}^i(a,b)] \right\| \quad (5)$$

where $h_{p,q}^i(a,b)$ denotes a patch at the position of (p,q) in the i^{th} neutral face image. Since each patch in the CM should show the degree of similarity instead of dissimilarity between the corresponding patches in the input face and the neutral face, we define the value of each patch in the i^{th} CM as follows:

$$CM_{p,q}^i(a,b) = \max \left\{ DM_{p,q}^i(a,b) \mid i = 1, 2, \dots, M \right\} - DM_{p,q}^i(a,b) \quad (6)$$

Where M is total number of neutral face images in the database. The obtained CM is normalized to the range of [0, 1] for weighting purpose. The advantage of the proposed CM is that we can use the PPZM features in the feature extraction step to create CM. So, there is no additional computational cost required for creating CM.

Figure 2 shows four samples of expression variant face images and their CMs compared to different neutral face images computed in our experiments. For displaying purpose, we rescaled the CM pixels to the range of [0, 255] as shown in Figure 2. Note that, in the CM, a brighter intensity indicates a larger contribution and on the contrary a darker intensity means a smaller contribution.

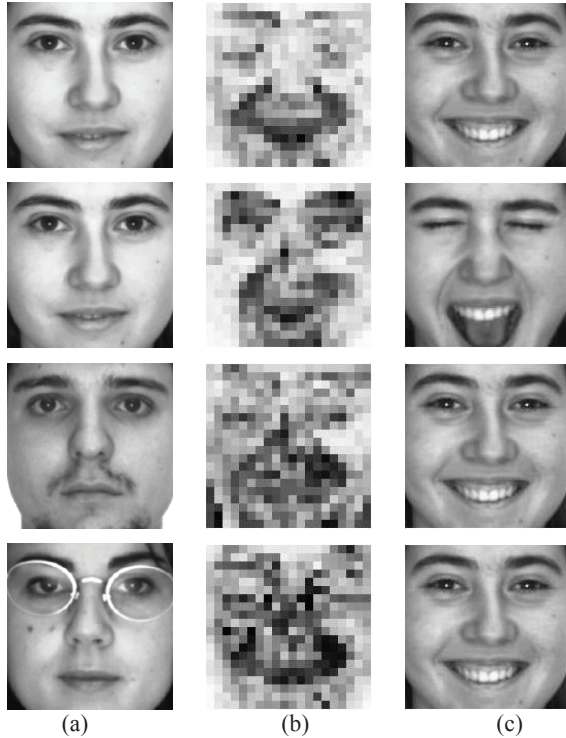


Fig. 2. Neutral face images, expression variant face images and their CMs. (a) Neutral face images; (b) CMs ($W=8$); (c) expression variant face images.

2.3. Similarity Measurement

Using above image representation, a face is described by a PPZMA associated with CMs corresponding to model neutral face images. For a given query face, the face recognition procedure produces the PPZMA descriptor of the query face and calculates the distance between the query PPZMA descriptor and the i^{th} model PPZMA descriptor in the database which is weighted using i^{th} CM of query face as:

$$D_w(f(x,y), h^i(x,y)) = \frac{\left\| \overline{W} \left[\overline{PPZMA}(f(x,y), h^i(x,y)) \right] \bullet \left(\overline{PPZMA}(f(x,y)) - \overline{PPZMA}(h^i(x,y)) \right) \right\|}{\left\| \overline{W} \left[\overline{PPZMA}(f(x,y), h^i(x,y)) \right] \right\|} \quad (7)$$

where $f(x,y)$ and $h^i(x,y)$ are the query and i^{th} model faces respectively, symbol \bullet denotes Hadamard product and $W \left[\overline{PPZMA}(f(x,y), h^i(x,y)) \right]$ is given as:

$$W \left[\overline{PPZMA}(f(x,y), h^i(x,y)) \right] = \left\{ CM_{p,q}^i(a,b) \mid p, q = 1, 2, \dots, \frac{N}{W} \right\} \quad (8)$$

where $CM_{p,q}^i(a,b)$ denotes the value of patch at the position (p,q) in the i^{th} CM. The closest model in the database is considered as the correct return.

$$Re sClass = \arg \min_i \left(D_w(f(x,y), h^i(x,y)) \right) \quad (9)$$

3. EXPERIMENTAL RESULTS

In order to evaluate the effectiveness of the proposed approach, we use the AR face database [6] which contains images of the frontal view faces of 126 people. In order to make a direct comparison of the proposed method against the benchmark system [3], a total of 100 people were randomly selected from AR face database. Each individual consists of eight images with neutral, happy, angry and scream expressions taken during two sessions separated by two-week time interval. These images for one subject in the database are displayed in Figure 3.

All the following experiments employed a single model based identification scheme, in which the neutral face images taken in the first session (e.g. Fig. 3(a)) were used as the gallery images and the other three expressions taken in both sessions (Figs. 3 (b, c and d) and (f, g and h)) were tested as the probe images. The performances of the proposed method are compared with three benchmark approaches, i.e. eigenface which is a widely used baseline approach, the correlation method and the optical flow-based weighted algorithm [3].

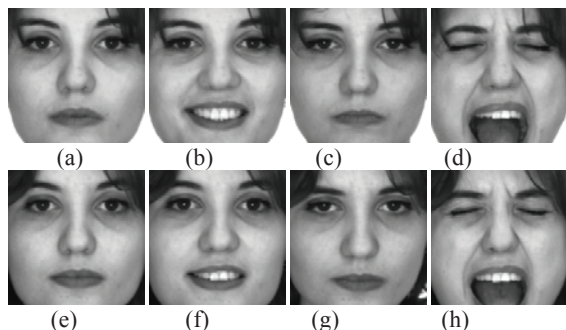


Fig. 3. Samples of AR cropped faces. (a) and (e): neutral faces under controlled/ideal condition taken in the first and second sessions respectively; (b, c and d) and (f, g and h): faces with different facial expressions taken in the first and second sessions respectively.

Experimental results are shown in Figure 4. It is found that the proposed method significantly outperformed the eigenface and correlation methods under different expression variations and optical flow-based weighted algorithm under happy and angry expression variations.

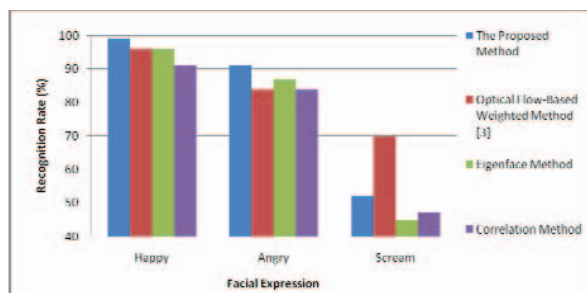


Fig. 4. Performance comparison using single model based identification scheme. The results of Correlation method are from [3].

Figures 5 and 6 also show the recognition rate of the proposed method in cumulative match score curve for images from the first and the second sessions respectively. It can be observed from these figures that the recognition rate of the proposed method under facial expression changes is better than that under duplicated facial expression changes. This is not unexpected because the illumination conditions of images from the first and the second sessions are slightly different.

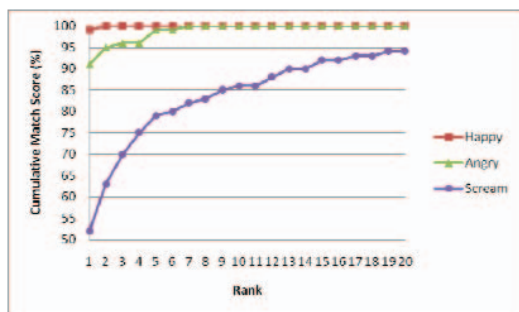


Fig. 5. Recognition rate obtained under different facial expression changes from first session.

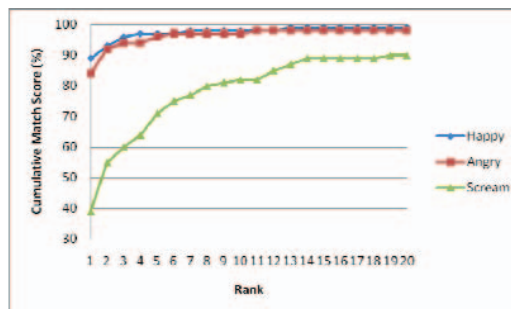


Fig. 6. Recognition rate obtained under different facial expression changes from second session.

4. CONCLUSION

In this paper, we have presented a novel approach to recognize expression variant faces from one exemplar image per enrolled subject. The proposed weighting method assigns proper weights to each PPZM to adjust the contribution of each local area of a face in terms of its similarity compared to the corresponding local area in the neutral face image. It is a very encouraging finding that the proposed approach significantly outperformed the eigenface method, the most widely used baseline for face recognition, and the Correlation method. Performance evaluation also indicates that the proposed algorithm performs superior to the optical flow-based weighted method [3] under happy and angry expression variations.

5. ACKNOWLEDGEMENT

This work was partially supported by the Australian Research Council (ARC) Discovery Grant Scheme.

6. REFERENCES

- [1] D. Beymer and T. Poggio, "Face Recognition from One Example View," *Science* 272(5250), 1996.
- [2] A.M. Mart'inez, "Matching Expression Variant Faces," *Vision Research*, Vol. 43, Issue 9, pp. 1047-1060, 2003.
- [3] A.M. Mart'inez, "Recognizing Expression Variant Faces from a Single Sample Image per Class," *Proceedings of IEEE Computer Vision and Pattern Recognition (CVPR)*, 2003.
- [4] H. R. Kanan, K. Faez and Y. Gao, "Face Recognition Using Adaptively Weighted Patch PZM Array from a Single Exemplar Image per Person", *Pattern Recognition*, Vol. 41, No. 12, pp.3799-3812, 2008.
- [5] R. Campbell, "The Lateralization of Emotion: A Critical Review," *Int'l J. Psychology* 17 (1982) 211-219.
- [6] A.M. Mart'inez, R. Benavente, "The AR Face Database," *CVC Technical Report No. 24*, June 1998, URL: (http://rv11.ecn.purdue.edu/~aleix/aleix_face_DB.html).