

Face Recognition by Using Elongated Local Binary Patterns with Average Maximum Distance Gradient Magnitude

Shu Liao and Albert C.S. Chung

Lo Kwee-Seong Medical Image Analysis Laboratory,
Department of Computer Science and Engineering,
The Hong Kong University of Science and Technology, Hong Kong
liaooshu@cse.ust.hk, achung@cse.ust.hk

Abstract. In this paper, we propose a new face recognition approach based on local binary patterns (LBP). The proposed approach has the following novel contributions. (i) As compared with the conventional LBP, anisotropic structures of the facial images can be captured effectively by the proposed approach using elongated neighborhood distribution, which is called the elongated LBP (ELBP). (ii) A new feature, called Average Maximum Distance Gradient Magnitude (AMDGM), is proposed. AMDGM embeds the gray level difference information between the reference pixel and neighboring pixels in each ELBP pattern. (iii) It is found that the ELBP and AMDGM features are well complement with each other. The proposed method is evaluated by performing facial expression recognition experiments on two databases: ORL and FERET. The proposed method is compared with two widely used face recognition approaches. Furthermore, to test the robustness of the proposed method under the condition that the resolution level of the input images is low, we also conduct additional face recognition experiments on the two databases by reducing the resolution of the input facial images. The experimental results show that the proposed method gives the highest recognition accuracy in both normal environment and low image resolution conditions.

1 Introduction

Automatic facial recognition (AFR) has been the topic of extensive research in the past several years. It plays an important role in many computer vision applications including surveillance development, and biometric image processing. There are still many challenges and difficulties, for example, factors such as pose [8], illumination [9] and facial expression [10]. In this paper, we propose a new approach to face recognition from static images by using new features to effectively represent facial images. Developing a face recognition framework involves two crucial aspects. 1. *Facial image representation*: It is also known as the feature extraction process, in which feature vectors are extracted from the facial images. 2. *Classifier design*: The feature vectors extracted in the first stage

are fed into a specific classifier to obtain the final classification results. In this paper, we focus on the first stage: the feature extraction stage.

Many feature extraction methods for facial image representation have been proposed. Turk and Pentland [4] used the Principle Component Analysis (PCA) to construct eigenfaces and represent face images as projection coefficients along these basis directions. Belhumeur *et al.* proposed the Linear Discriminant Analysis (LDA) method [5]. Wiskott *et al.* used the Gabor wavelet features for face recognition [3].

In recent years, a new feature extraction method is proposed, which is known as the Local Binary Patterns (LBP) [1]. LBP is first applied in the texture classification application [6]. It is a computationally efficient descriptor to capture the micro-structural properties of the facial images. However, there is major limitation of the conventional LBP. The conventional LBP uses circularly symmetric neighborhood definition. The usage of the circularly symmetric neighborhood definition aims to solve the rotation invariant problem in the texture classification application with the cost of eliminating anisotropic structural information. However, for face recognition, such problem does not exist. Anisotropic structural information is an important feature for face recognition as there are many anisotropic structures exist in the face (e.g. eyes, mouths). To this end, we extend the neighborhood distribution in the elongated manner to capture anisotropic properties of facial images called the elongated LBP (ELBP), the conventional LBP is a special case of the ELBP. Also, the conventional LBP does not take the gradient information into consideration. In this paper, we propose a new feature named average maximum distance gradient magnitude (AMDGM) to capture the general gradient information for each ELBP pattern. It is experimentally shown that the ELBP feature and the AMDGM feature are complement with each other and can achieve the highest recognition accuracy among all the other compared methods in both normal environment and low input image resolution conditions.

The paper is organized as follows. In Section 2, the concepts of ELBP and AMDGM are introduced. Section 3 describes the experimental results for various approaches under the conditions of normal environment and low resolution images. Section 4 concludes the whole paper.

2 Face Recognition with ELBP and AMDGM

In this section, the ELBP and AMDGM features are introduced. We will first briefly review the conventional LBP approach and then describe these two features.

2.1 Elongated Local Binary Patterns

In this section, the Elongated Local Binary Patterns (ELBP) are introduced. In the definition of the conventional LBP [6], the neighborhood pixels of the reference pixel are defined based on the circularly symmetric manner. There are

two parameters m and R respectively representing the number of the neighbor pixels and the radius (i.e. the distance from the reference pixel to each neighboring pixel). By varying the values of m and R , the multiresolution analysis can be achieved. Figure 1 provides examples of different values of m and R . Then,

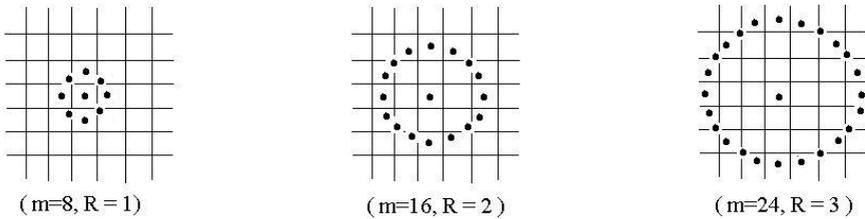


Fig. 1. Circularly symmetric neighbor sets for different values of m and R

the neighboring pixels are thresholded to 0 if their intensity values are lower than the center reference pixel, and 1 otherwise. If the number of transitions between "0" and "1" is less or equal to two, then such patterns are uniform patterns. For example, "00110000" is a uniform pattern, but "01011000" is not a uniform pattern. It is obvious that there are $m + 1$ possible types of uniform patterns. The final feature vector extracted from the conventional LBP is the occurrence of each type of uniform pattern in an input image as the authors in [6] pointed out that the uniform patterns represented the basic image structures such as lighting spots and edges.

As we can see, for the conventional LBP, the neighborhood pixels are all defined on a circle with radius R and reference center pixel. The main reason for defining neighboring pixels in this isotropic manner is aimed to solve the rotation invariant problem in the texture classification application which is the first application of the conventional LBP. Later, the conventional LBP was applied in the face recognition application [1]. However, in this application, the rotation invariant problem does not exist. Instead, anisotropic information are important features for face recognition. To the best of our knowledge, this problem has not been mentioned by any researchers.

Therefore, we are motivated to propose the ELBP approach. In ELBP, the distribution of neighborhood pixels gives an ellipse shape (see Figure 2). There are three parameters related to the ELBP approach: 1. The long axis of the ellipse, denoted by A ; 2. The short axis of the ellipse, denoted by B ; 3. The number of neighboring pixels, denoted by m . Figure 2 shows examples of the ELBP patterns with different values of A , B and m :

The X and Y coordinates, g_{ix} and g_{iy} , of each neighbor pixel g_i ($i = 1, 2, \dots, m$) with respect to the center pixel is defined by Equations 1 and 2 respectively,

$$R_i = \sqrt{\frac{A^2 B^2}{A^2 \sin^2 \theta_i + B^2 \cos^2 \theta_i}}, \tag{1}$$

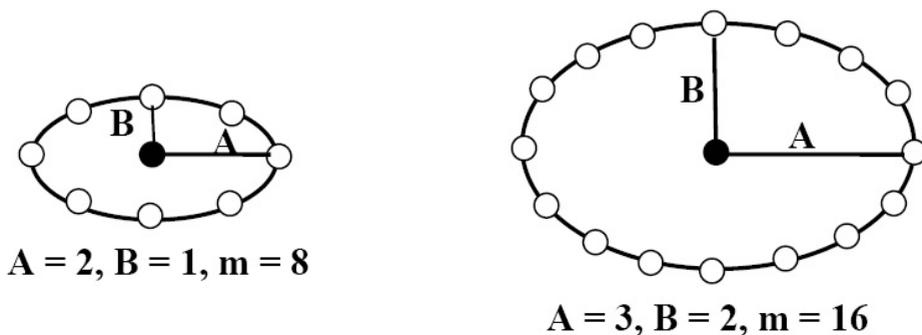


Fig. 2. Examples of ELBP with different values of A , B and m

$$g_{ix} = R_i * \cos \theta_i, \quad \text{and} \quad g_{iy} = R_i * \sin \theta_i, \quad (2)$$

where $\theta_i = (\frac{360}{m} * (i-1))^\circ$. If the coordinates of the neighboring pixels do not fall exactly at the image grid, then the bilinear interpolation technique is applied. The final feature vector of ELBP is also the occurrence histogram of each type of uniform pattern. In this paper, three sets of ELBP are used with different values of A , B and m : $A_1 = 1, B_1 = 1, m = 8$; $A_2 = 3, B_2 = 1, m = 16$; $A_3 = 3, B_2 = 2, m = 16$.

Similar to [1], before processing the input image for face recognition, the input image is divided into six regions in advanced: bows, eyes, nose, mouth, left cheek, and right cheek, which are denoted as R_1, R_2, \dots, R_6 . Each region is assigned a weighting factor according to their importance. The larger the weighting factor, the more important the region. In this paper, the weighting factors for these six regions are set to be: $w_1 = 2, w_2 = 4, w_3 = 2, w_4 = 4, w_5 = 1, w_6 = 1$. The ELBP histograms are estimated from each region. Then, the feature vector is normalized to the range of $[-1, 1]$. Finally, the normalized feature vector is multiplied by its corresponding weighting factor to obtain the region feature vector. As such, the region feature vector encodes the textural information in each local region. By concatenating all the region feature vectors, global information of the entire face image can be obtained.

The ELBP Pattern can also be rotated along the center pixel with a specific angle β to achieve multi-orientation analysis and to characterize elongated structures along different orientations in the facial images. In this paper, four orientations $\beta_1 = 0^\circ, \beta_2 = 45^\circ, \beta_3 = 90^\circ, \beta_4 = 135^\circ$ are used for each ELBP pattern with its own parameters A , B and m . The final ELBP feature vector is an $m+1$ dimension vector F , where each element F_i ($i = 1, 2, 3, \dots, m+1$) denotes the occurrence of a specific type of uniform pattern along all those four orientations $\beta_1, \beta_2, \beta_3$ and β_4 in an input image.

As we can see, the ELBP features are more general than the conventional LBP. More precisely, the conventional LBP can be viewed as a special case of ELBP when setting the values of A and B equal to each other. The ELBP is able to capture anisotropic information from the facial images, which are important

features as there are many important parts in the face such as eyes, mouths are all elongated structures. Therefore, it is expected that ELBP can have more discriminative power than the conventional LBP, which will be further verified in the Experimental Results Section.

2.2 Average Maximum Distance Gradient Magnitude

As mentioned in Section 2.1, ELBP is more general than the conventional LBP and the anisotropic information can be effectively captured. However, both the conventional LBP and the proposed ELBP still do not take the gradient information of each pattern into consideration. Since both the conventional LBP and ELBP are constructed by thresholding the neighboring pixels to 0 and 1 with respect to the reference center pixel, gradient magnitude information is therefore not included. In this paper, a new measure, called the average maximum distance gradient magnitude (AMDGM) is proposed to effectively capture such information. To define AMDGM, we first introduce the concept of distance gradient magnitude (DGM). For each ELBP pattern, there are three parameters: A , B and m denoting the long axis, short axis and the number of neighboring pixels. Then, the distance gradient magnitude for each neighboring pixel g_i , given the center pixel g_c , is defined by Equation 3,

$$|\nabla_d I(g_i, g_c)| = \frac{|I_{g_i} - I_{g_c}|}{|v_i - v_c|^2}, \quad (3)$$

where $v = (x, y)$ denotes the pixel position, I_{g_i} and I_{g_c} are the intensities of the neighbor pixel and the reference pixel respectively. Based on the definition of DGM, the maximum distance gradient magnitude $G(v)$ is defined by Equation 4,

$$G(v) = \max_{g_i} |\nabla_d I(g_i, g_c)|, i = 1, 2, \dots, m. \quad (4)$$

Suppose that, in an input image, for each type of uniform ELBP patterns P_i ($i = 1, 2, \dots, m$), its occurrence is N_i . Then, the average maximum distance gradient magnitude (AMDGM) $A(P_i)$ for each type of uniform patterns is defined by Equation 5,

$$A(P_i) = \frac{\sum_{k=1}^{N_i} G(v_k)}{N_i}, \quad (5)$$

where $v_k \in P_i$, the AMDGM feature has more advantage over the conventional gradient magnitude as it encodes the spatial information (i.e. the distance from the neighbor pixel to the reference center pixel) into consideration. It is essential because the neighborhood distribution is no longer isotropic. The distance from each neighborhood pixel to the reference pixel can be different, unlike the conventional LBP. The AMDGM feature is well complement with ELBP because the ELBP provides pattern type distribution and the AMDGM feature implies the general gradient information with spatial information for each type of uniform patterns.

3 Experimental Results

The proposed approach have been evaluated by performing face recognition experiments on two databases: ORL and FERET [7]. The ORL database contains 40 classes with 10 samples for each class, each sample has resolution of 92×112 pixels. For the FERET database, we have selected 70 subjects from this database with six up-right, frontal-view images of each subject. For each facial image, the centers of the two eyes have already been manually detected. Then, each facial image has been translated, rotated and scaled properly to fit a grid size of 64×64 pixels. The proposed method has been compared with three widely used methods similar to [1]: 1. Principle Component Analysis (PCA), also known as the eigenface approach [4]; 2. Linear Discriminant Analysis (LDA) [5]; 3. The conventional Local Binary Patterns (LBP) [1]. The support vector machine (SVM) [2] with the Gaussian Radius Basis Function (RBF) kernel was used as the classifier in this work. To test the robustness of the proposed method under the condition of low input image resolution, which is a practical problem in real world applications, we have also performed face recognition experiment on the ORL and FERET databases by downsampling the input images.

3.1 Experiment on ORL and FERET Databases with Original Resolution

We have tested all the approaches under the normal environment in both ORL and FERET databases (i.e. all the input images were in their original resolution). Figure 3 and Figure 4 show some sample images of the ORL and FERET databases.

The purpose of this experiment is to test the basic discriminative power of different approaches. In this experiment, half of the facial images for each class were randomly selected as training sets, the remaining images were used as the testing sets. The experiment was repeated for all possible combinations of training and testing sets. The average recognition rates for different approaches in both the ORL and FERET databases are listed in Table 1.



Fig. 3. Some sample images from the ORL Database



Fig. 4. Some sample images from the FERET Database

Table 1. Performance of different approaches under the normal condition in the ORL and FERET Database. Results of the proposed methods are listed in Rows 4 – 5.

Methods	Recognition Rate %	
	ORL	FERET
1. PCA [4]	85.00	72.52
2. LDA [5]	87.50	76.83
3. LBP [1]	94.00	82.62
4. ELBP	97.00	86.73
5. ELBP + AMDGM	98.50	93.16

From Table 1, the proposed method has the highest recognition rate among the compared methods in both databases. Furthermore, the ELBP and AMDGM features are well complement with each other. The discriminative power of the proposed method is obviously implied.

3.2 Experiment on ORL and FERET Databases with Low Resolution Images

In real world applications, the quality of the input facial images is not always good due to various factors such as imaging equipment and outdoor environment (i.e. low resolution facial image). Therefore, the robustness of a face recognition approach under this condition is essential. In this experiment, the input images of both the ORL and FERET databases were downsampled to 32×32 pixels before processing, the rest of the settings were the same as Section 3.1. The recognition rates of different approaches under this environment are listed in Table 2.

Table 2 echoes the robustness of different methods against the condition of low input image resolution. It is shown that the proposed method maintains the highest recognition rate among all the compared methods. Its robustness under such condition is clearly illustrated.

Table 2. Performance of different approaches under the low resolution condition in the ORL and FERET Database. Results of the proposed methods are listed in Rows 4 – 5.

Methods	Recognition Rate %	
	ORL	FERET
1. PCA [4]	66.00	58.41
2. LDA [5]	69.50	64.85
3. LBP [1]	78.50	73.63
4. ELBP	83.00	80.22
5. ELBP + AMDGM	88.50	85.47

4 Conclusion

This paper proposed a novel approach to automatic face recognition. Motivated by the importance of capturing the anisotropic features of facial images, we propose the ELBP feature, which is more general and powerful than the conventional LBP. Also, to embed gradient information based on the definition of ELBP, a new feature AMDGM is proposed. The AMDGM feature encodes the spatial information of the neighboring pixels with respect to the reference center pixel, which is essential for the ELBP. Experimental results based on the ORL and FERET databases demonstrate the effectiveness and robustness of our proposed method when compared with three widely used methods.

References

1. Ahonen, T., Hadid, A., Pietikainen, M.: Face recognition with local binary patterns. In: Pajdla, T., Matas, J.(G.) (eds.) ECCV 2004. LNCS, vol. 3021, pp. 469–481. Springer, Heidelberg (2004)
2. Vapnik, V.N.: Statistical Learning Theory. John Wiley and Sons, New York (1998)
3. Wiskott, L., Fellous, J.-M., Kuiger, N., von der Malsburg, C.: Face recognition by elastic bunch graph matching. *IEEE PAMI* 19(7), 775–779 (1997)
4. Turk, M., Pentland, A.: Eigenfaces for recognition. *J. Cogn. Neurosci.* 3, 71–86 (1991)
5. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE PAMI* 19(7), 711–720 (1997)
6. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE PAMI* 24(7), 971–987 (2002)
7. Phillips, P.J., Wechsler, H., Huang, J., Rauss, P.: The feret database and evaluation procedure for face recognition algorithms. *IVC* 16(5), 295–306 (1998)
8. Blanz, V., Vetter, T.: Face recognition based on fitting a 3D morphable model. *IEEE PAMI* 25(9), 1063–1074 (2003)
9. Shashua, A., Riklin-Raviv, T.: The quotient image: Class based re-rendering and recognition with varying illuminations. *IEEE PAMI* 23(2), 129–139 (2001)
10. Tian, T.-I., Kanade, T., Cohn, J.F.: Recognizing action units for facial expression analysis. *IEEE PAMI* 23(2), 97–115 (2001)