

# A Novel Feature Extraction Method for Facial Expression Recognition

Xiaoyi Feng, Baohua Lv, Zhen Li, Jiling Zhang<sup>1</sup>

<sup>1</sup>School of Electronic and Information, Northwestern Polytechnic University, Xi'an, China

## Abstract

In this work, a novel facial feature extraction method is proposed for automatic facial expressions recognition, which detecting local texture information, global texture information and shape information of the face automatically to form the facial features. First, Active Appearance Model (AAM) is used to locate facial feature points automatically. Then, the local texture information in these feature points and the global texture feature information of the whole face area are extracted based on the Local Binary Pattern (LBP) techniques, and also the shape information of the face are detected. Finally, all the information are combined together to form the feature vector. The proposed feature extraction method is tested by the JAFFE database and experimental results show that it is promising.

**Keywords:** Feature extraction, AAM, texture information, shape information.

## 1. Introduction

Facial expression recognition plays an important role in human-computer interaction. In recent years, numerous algorithms for facial expression analysis from static images have been proposed [1]-[3], which are quite different in the facial features.

Lyons et al. [4] adopted a wavelet-based face representation. Input images were convolved with the Gabor filters of five spatial frequencies and the amplitude of the complex-valued filter responses were sampled on 34 manually selected fiducial points and combined into a single vector, containing 1020 elements. Zhang et al. [5] used a similar representation while they applied wavelet of 3 scales and 6 orientations. They also considered geometric position of the 34 fiducial points as features. Shinohara and Otsu [6] integrated Higher-order Local Auto-Correlation (HLAC) features and Fisher weight map to form a feature vector. First, facial area was cropped manually and resized to 32×40 pixels. Next, illumination correction and histogram equalization

was performed. Then, HLAC features are computed at each pixel in an image and are integrated with a weight map to obtain a feature vector. In Zheng's work [7], 34 landmark points were manually located from each facial image and then were converted into a labeled graph (LG) vector using the Gabor wavelet transformation method. The whole face's features were extracted in Fasel's method [8].

In general, facial features contain three kinds of information: texture information [4], [6], [7], shape information and the combination of texture and shape information [5]. In most cases, the above information is extracted from several facial feature points, which should be located correctly. Since it is not easy to locate these feature points correctly, these points are usually marked manually. Besides of this, information in several feature points is not enough for face's presentation.

In our previous work [9] [10], the Local Binary Pattern (LBP) technique was used to extract facial texture features, which could be easily obtained from the whole face area after face image preprocessing. Experimental results showed that the feature vector could describe face efficiently for expression recognition. Two questions existed in our work: First, since the LBP based feature was face's global texture feature, it was effective to reflect typical expressions, but not good at describing small and local expression changes. Second, since the position of two pupils was used for face normalization in preprocessing, it was still needed to locate pupils manually in our method.

In this paper, a novel facial feature extraction method is proposed to solve the above questions: First, AAM is used to detect some feature points of the face automatically, then the center of eyes, mouth are calculated, finally the local texture information, global information and shape information are combined together to form the feature vector.

The rest of the paper is organized as follows. AAM based feature location method is introduced in section 2. The new feature vector is presented in section 3. In section 4, we introduce the expression recognition results. Finally in section 5, we conclude the paper.

## 2. Feature Points Detection Based on AAM

The Active Appearance Model (AAM) is a powerful tool [10] to extract face shape and texture information, which uses Principal Component Analysis (PCA) to model both shape and texture variations by the training set and uses parameters  $c$  controlling the shape and texture according to

$$\begin{aligned} s_i &= \bar{s} + Q_s c_i \\ g_i &= \bar{g} + Q_t c_i \end{aligned} \quad (1)$$

Here  $Q_s$  and  $Q_t$  are truncated matrices describing the principal modes of combined appearance variations, which are derived from the training set.  $\bar{s}$  and  $\bar{g}$  are the mean shape and texture of samples in the training set.  $c_i$  is the vector of appearance parameters simultaneously controlling the shape  $s_i$  and texture  $g_i$ .

In our work, 70 points (16 points in eyes, 16 points in eyebrows, 7 points in nose, 22 points in mouth, and 9 points in face contour) are selected as feature points to model face shapes (see Fig.1).

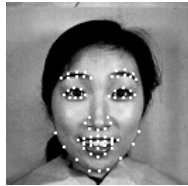


Fig.1: Location of feature points.

To make our method be effective under expressions, we produce 7 models from the training set, each corresponding to one kind of expressions. In the searching procedure, the most matched one is selected as the final search result.

## 3. Feature Vector Extraction

The Local Binary Pattern (LBP) technique is used here to extract texture information, and direction between each pair of feature points is used as shape information.

### 3.1. The Local Binary Pattern Operator

Fig.2 is an illustration of the basic LBP operator [12]. The original  $3 \times 3$  neighborhood at the left is threshold by the value of the center pixel, and a binary pattern code is produced. The LBP code of the center pixel in

the neighborhood is obtained by converting the binary code into a decimal one. Some examples of local texture primitives detected by LBP are shown in Fig.3.

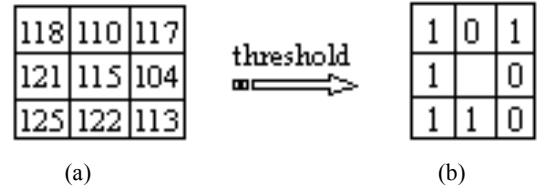


Fig.2. The basic LBP operator.

(a) Original neighbourhood

(b) Local binary pattern

Binary code: 11110010 LBP code:242

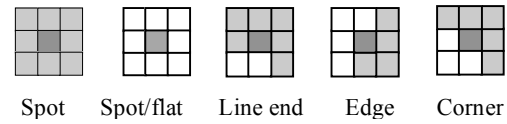


Fig.3. Different texture primitives detected by LBP (white box represent ones and light gray zeros).

Based on the operator, each pixel of an image can be labeled with an LBP code.

## 3.2. Local Texture Information Extraction

The feature vector of local texture information (supposed as  $v_{ll}$ ) is formed by the LBP code of each feature points detected by AAM, so the dimension of it is 70.

## 3.3. Global Texture Information Extraction

Global texture (suppose as  $v_{tg}$ ) is extracted with the following steps:

1. Divide the face image into small regions. After pre-processing, each image's size is  $150 \times 128$ . In our experiment, each normalized image is divided into 80 ( $10 \times 8$ ) non-overlapping blocks, with 240 ( $15 \times 16$ ) pixels in each block. In fact, 76 of the 80 blocks belong to the non-masked area and are used for expression recognition, while the other 4 blocks are corresponding to masked area and are left out of.

2. Calculate the LBP histogram of each region.

3. Concatenate the LBP feature histograms of each region into a single feature vector, with the dimension of 19456.

## 3.4. Shape information Extraction

Suppose positions of two feature points are  $(x_i, y_i)$  and  $(x_{i+1}, y_{i+1})$ , then the direction between them is defined as:

$$\alpha_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \quad (2)$$

Shape feature vector  $v_s$  is formed by the direction between each pair of feature points as shown in Fig.1 and its dimension is 140.

As a result, the new feature vector  $v$  is composed as

$$v = \{v_{tg}, v_{tl}, v_s\} \quad (3)$$

## 4. Experimental Results

The Japanese Female Facial Expression (JAFPE) Database [4] is used in our experiment, which contains 213 images of ten expressers posed 3 or 4 examples of each of the seven basic expressions (happiness, sadness, surprise, anger, disgust, fear, neutral). Examples from the database are shown in Fig.4.

Suppose  $e_i = \{(x_i, y_i) | i = 0, 1, \dots, 7\}$  describe the searched shape of left eye and  $(x_{el}, y_{el})$  denotes its center. We have

$$\begin{aligned} x_{el} &= \frac{1}{8} \sum_{i=0}^7 x_i \\ y_{el} &= \frac{1}{8} \sum_{i=0}^7 y_i \end{aligned} \quad (4)$$

Center of the right eye and mouth can be calculated in the similar way.

Since eyes center and mouth center is determined by several feature points in our method, they can be detected correctly enough for face normalization, even if several feature points are not located quite exactly.



Fig.4: Samples from original Japanese Female Facial Expression Database. The resolution is 256X256 pixels.

After location of eyes and mouth, the images are normalized and cropped with an elliptical mask to exclude non-face area from the image (see Fig.5).



Fig.5: Sample of preprocessing images. The resolution is 150 × 128 pixels.

The nearest neighborhood classification with the weighted Chi square statistic is used here for expression classification.

There are mainly two ways to divide the JAFPE database. The first way is to divide the database randomly into 10 roughly equal-sized segments, of which nine segments are used for training and the last one for testing. The second way is to divide the database into several segments, but each segment corresponds to one expresser.

To compare our results to other methods, we use both of above division methods respectively. First, the database is divided into 10 roughly equal-sized sets: nine sets are used for training and the remaining one for testing. The above process is repeated so that each of the ten roughly equal-sized sets is used once as the test set. Average result over all 10 cycles is considered as the recognition rate of one trial. The above process is repeated so that each of the nine partitions is used once as the test set. The average of recognizing the expression of novel expressers is 72%.

Now we compare the recognition performance to other published methods using the same database. Methods in [4] and [9] used the first way to divide the database. In [4], a result of 75% was reported, which needed to locate 34 fiducial points manually. In [9], a result of 77% was obtained, which needed to locate two eyes center manually. Methods in [5] and [10] used the second database division. In [5], a recognition

accuracy of 90.1% was obtained after locating 34 fiducial points manually. In [10], a result of 93.8% was obtained after manually located eyes center during preprocessing. While in our method, the recognition procedure is fully automatic.

It can be seen from experimental results that the proposed feature vector can represent face more efficiently. Examples of some correct recognition result are shown in Fig.6, which are wrongly recognized in our previous work.



Fig.6; Examples of right recognition.

From left, the recognized expressions are disgust and sadness with the new method, while they are recognized as sadness and neutral with the old method.

## 5. Conclusion

How to present face effectively is one of the key tasks in facial expression recognition. The Local Binary Pattern technique was used to represent face global texture effectively in our previous work and experiments show it is effective to describe classical and obvious expression features. To improve its performance, a novel feature vector is proposed in this paper, which take the local texture feature, global texture feature and shape feature into consider for forming feature vector. To make the whole recognition procedure automatically, AAM is used for feature points' location, and the center of eyes and mouth is calculated based on them. Experimental results on the JAFFE database demonstrated that our method performs well.

## 6. Acknowledgement

The author would like to thank Dr. M. Lyons for providing the facial expression database.

Financial support for this work was obtained from the natural science foundation of Shaanxi province and the NWPU talent program, which are greatly acknowledged.

## 7. References

[1] M.Pantic and L.Rothkrantz, "Automatic Analysis of Facial Expressions: The State of the Art", *IEEE Trans. on Pattern Analysis and*

*Machine Intelligence*, Vol. 22, 2000, pp. 1424-1445.

[2] B.Fasel and J.Luetttin, "Automatic Facial Expression Analysis: A Survey", *Pattern Recognition*, Vol.36, 2003, pp. 259-275.

[3] W.Fellenz, J.Taylor, N.Tsapatsoulis, and S.Kollias, "Comparing Template-based, Feature-based and Supervised Classification of Facial Expression from Static Images", *Computational Intelligence and Applications*, 1999.

[4] M.Lyons, J.Budynek, and S.Akamastu, "Automatic Classification of Single Facial Images", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.21, 1999, pp. 1357-1362.

[5] Z.Zhang, M.Lyons, M.Schuster, and S.Akamatsu, "Comparison Between Geometry-based and Gabor-Wavelet-based Facial Expression Recognition Using Multi-layer Perceptron", *Proc. 3<sup>rd</sup> Int. Conf. Automatic Face and Gesture Recognition*, 1998, pp. 454-459.

[6] Y.Shinohara and N.Otsu, "Facial Expression Recognition Using Fisher Weight Maps", *IEEE Conf. on Automatic Face and Gesture Recognition*, 2004, pp. 499-504.

[7] W. Zheng, X.Zhou, C. Zou and L. Zhao, "Facial expression recognition using kernel canonical correlation analysis (KCCA)", *IEEE Trans. on Neural Networks*, Vol.17, 2006, pp. 233-238.

[8] B. Fasel, "Head-pose invariant facial expression recognition using convolutional neural networks. In: *Fourth IEEE Conference on Multimodal Interfaces*. 2002, 529- 534

[9] X.Feng, A.Hadid, and M.Pietikainen, "A Coarse-to-Fine Classification Scheme for Facial Expression Recognition", *Image Analysis and Recognition, ICIAR 2004 Proceedings, LNCS 3212*, Springer, pp.668-675.

[10] X.Feng, A.Hadid & M.Pietikainen, "Facial Expression Recognition with Local Binary Patterns and Linear Programming", *Pattern Recognition and Image Analysis*, Vol. 15(2), 2005, pp. 546-549.

[11] T. F. Cootes and P. Kittipanya-ngam, Comparing variations on the active appearance model algorithm. In *BMVC 2002*, pp.837-846.

[12] T. Ojala, M. Pietikäinen and T. Mäenpää. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002. Vol. 24(7), pp. 971-987.