A Facial Expression Recognition Method Based on Feature Blocks

Wencheng Wang*, Faliang Chang**

** College of Information and Control Engineering, Weifang University, Weifang, 261061, China, (e-mail: wwcwfu@126.com).*
*** College of Control Science and Engineering, Shandong University, Jinan, 250014, China, (e-mail: flchang@sdu.edu.cn)*

Abstract: In this paper, a novel method for human facial expression recognition (FER) is proposed. It adopts the idea that the appearance of a region of interest can be well characterized by the distribution of its local features. Considering the importance of the eyes and mouth for FER and the outstanding performance of local binary pattern (LBP) to extract local textures, a representation model for facial expressions based on feature blocks (FB) and LBP descriptors is proposed. The strategies of FER including face normalization, feature-block acquisition, and LBP feature extraction are explained in detail. A principal component analysis (PCA) method is implemented to learn the structure of the expression in the LBP feature space. A recognition experiment is conducted on the JAFFE facial expression and TFEID databases using the nearest neighbor classifier. Experimental results confirm that the method is simple and demonstrates competitive performance.

Keywords: Facial expression, feature block, local binary pattern, principal component analysis.

1. INTRODUCTION

Facial expression is an important factor in human communication. It is one of the most natural and immediate means for human beings to communicate their emotions and intentions. As early as 1971, (Ekman and Friesen, 1971) postulated six primary emotions, each of which possesses a distinctive content together with a unique facial expression. These prototypic emotional displays are named happiness, sadness, fear, disgust, surprise, and anger. With the development of computer technology, automatic facial expression recognition (FER) has attracted considerable attention for its wide range of potential applications in areas such as image understanding, psychological study, synthetic face animation, and intelligent human-computer interaction. Automatic facial expression is a challenging task in intelligent human-computer interaction and several methods have been proposed to address this issue. The Facial Action Coding System (FACS) is a typical example for representing and understanding human facial expressions (Fasel and Luettin, 2003). Using this as a basis, several systems were successfully developed for facial expression analysis and recognition. (Turk and Pentland, 1991) employed Principal Component Analysis (PCA) to calculate feature sets called Eigenfaces. (Abboud and Davoine, 2004) proposed a bilinear factorization expression classifier for the recognition. (Lyons et al., 1999) proposed a method for classifying facial images automatically based on labeled elastic graph matching and 2-dimensional Gabor wavelet representation. (Rosenblum et al., 1996) used a system of networks where the complexity of recognizing facial expressions was divided into three layers of decomposition. Several other novel approaches have been applied in FER including Artificial Neural Networks (ANN), Support Vector Machines (SVM), Wavelet Analysis, Hidden Markov Model (HMM), and Optical Flow (Wang et al., 2010; Domaika and Davoine, 2008; Kotsia and Pitas, 2007; Valstar and Mehu, 2012; Yu et al., 2014; Khan et al., 2014; Wan and Aggarwal, 2014). Though enormous efforts have been made by many researchers and remarkable achievements have been realized in FER, recognizing facial expressions with high accuracy remains difficult owing to the subtlety, complexity, and variability of facial expressions.

In FER, how to extract a feature is a key problem. The goal of feature extraction is to obtain a more compact representation of the data with limited loss of information and therefore make the result more suitable for classification. The more reasonable the feature selection, the higher the recognition rate will be. In general, there are two categories of feature representation: geometric features and appearance features (Yang et al., 2009; Zhu et al., 2006; Liu and Wechsler, 2002; Wang and Chang, 2014; Hadid et al., 2004). Appearance features have been demonstrated to be superior to geometric features because geometric features are sensitive to noise, especially illumination noise. Therefore, appearance features are frequently used for representing facial expression. There are two popular appearance-based approaches to extract facial features: Gabor translation and local binary pattern (LBP). Gabor appearance feature extraction is frequently used to describe local appearances; it can achieve a high recognition rate. It suffers, however, from the disadvantages of excessive computation and a high dimension of feature space (Danisman et al., 2004; Shan et al., 2013). Compared to Gabor translation, LBP is excellent because of its low computation cost and texture description ability. It has gained increasing attention in facial image analysis owing to its robustness to challenges such as pose and illumination changes (Ying et al., 2009; Zhao and Zhang, 2011; Zhan and Cheng, 2010; Feng et al., 2005). However, during the course of using LBP, it has been observed that some local facial regions contain more useful information for expression classification than others from the original face images. That is, different sub-regions have a different contribution to the classification. This motives us to propose in this paper a method for FER that crops the feature blocks from the original image and extracts the LBP features.

The remainder of this paper is organized as follows. Section 2 describes the LBP operator and the method used to obtain the feature blocks. The steps for feature extraction and dimension reduction based on feature blocks, LBP, and PCA are introduced in Section 3. Section 4 describes the classification rule and system framework. The experiments performed on two facial expression databases are presented in Section 5. The comparisons of the proposed algorithm with various other methods of FER are also included in Section 5. The paper is concluded in Section 6.

2. LOCAL BINARY PATTERN AND FEATURE BLOCK

2.1 Local Binary Pattern

The original LBP operator is a powerful method for texture description. It was introduced by (Ojala et al., 2002) and describes the surroundings of a pixel by generating a bit-code from the binary derivatives of the pixel in an image. At a given pixel position (x_c, y_c), LBP is defined as an ordered set of binary comparisons of pixel intensities between the central pixel and its neighbor pixels. Taking 3×3 pixels for an example, the resulting LBP pattern at the pixel can be expressed as follows:

$$LBP(x_c, y_c) = \sum_{n=0}^{7} s(i_n - i_c) \times 2^n \tag{1}$$

where i_c is the gray value of the central pixel (x_c, y_c), i_n is the gray values of the eight surrounding pixels, and if $i_n - i_c > 0$, then $s(i_n - i_c) = 1$, else $s(i_n - i_c) = 0$. Fig. 1 is an example of how to compute the LBP value.

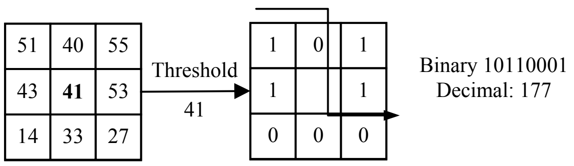


Fig.1. Example of basic LBP operator.

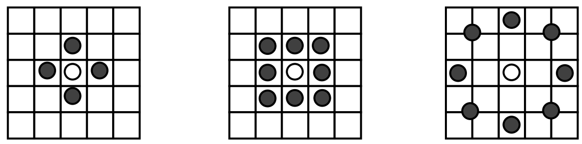
To improve the robustness and generalization ability of the original LBP operator and capture the large-scale structure possible with the dominant features of some textures, LBP was extended to use neighborhoods of different sizes. One of the successful extensions to the original operator is called uniform patterns, which contain at most two bitwise transitions from zero to one or vice versa when the binary string is considered circular. For example, 00000000, 00011110, and 10000011 are uniform patterns. Based on a large number of image statistics, the majority of patterns in images are uniform patterns. Ojala's experiments confirmed that uniform patterns account for almost 90% of all patterns in texture images. Therefore, using uniform patterns loses only minimal image information. Another extension of the multi-resolution LBP operator is defined as $LBP_{P,R}^{u2}$. This uses the operator in a neighborhood of P sampling points on a circle of radius R. Superscript $u2$ implies using uniform patterns and labeling all the remaining patterns with a single label. The formula is defined as:

$$LBP_{P,R}^{u2} = \sum_{j=0}^{p-1} S(g_j - g_c) \times 2^j \tag{2}$$

where g_j and g_c are the gray values of the $j-th$ pixel and the central pixel, respectively. $S(x)$ is a unit step function that is defined as:

$$S(x) = \begin{cases} 1 & if\,(x \geq 0) \\ 0 & if\,(x < 0) \end{cases} \tag{3}$$

Multi-resolution analysis can be achieved by choosing different values of R and P. Fig. 2 illustrates three different radiuses of LBP operators. From left to right, they are $LBP_{4,1}^{u2}$, $LBP_{8,1}^{u2}$, and $LBP_{8,2}^{u2}$ operators.



(P=4,R=1) **(P=8,R=1)** **(P=8,R=2)**

Fig. 2. Three different LBP operators.

After labelling an image with the LBP operator, a histogram of the labelled image $f_l(x, y)$ can be defined as:

$$H_i = \sum_{x,y} I(f_l(x, y) = i), \quad i = 0, \ldots, n-1 \tag{4}$$

where n is the number of different labels produced by the LBP operator and:

$$I(A) = \begin{cases} 1 & A\ is\ true \\ 0 & A\ is\ false \end{cases} \tag{5}$$

This LBP histogram contains information regarding the distribution of the local micro-patterns such as edges, spots, and flat areas. For the entire image, it can be used to describe the image characteristics statistically (Ahonen et al., 2004; Maalej and Amor, 2011).

2.2 Feature Block

The traditional approaches for FER based on the entire face are highly sensitive to noise caused by human face contour, hair, and other factors. To reduce the influence of such noise, the original image requires preprocessing by removing the unnecessary parts such as hair, background, and contour, retaining the main areas of the face only. However, the preprocessed face image continues to include redundancies. Using the nose as an example, when a person expresses an emotion, the shape and position of the nose tip and sides change only marginally, making this feature of limited value for classification. However, it will generate noise for different people having different nose shapes. Based on the above, a method is proposed to separate the main expression characteristic areas into feature blocks(FB). Fig. 3 illustrates an example of facial expression feature-block separation. The feature blocks include the left eyebrow, right eyebrow, left eye, right eye, area between the eyebrows, nose, and entire mouth.

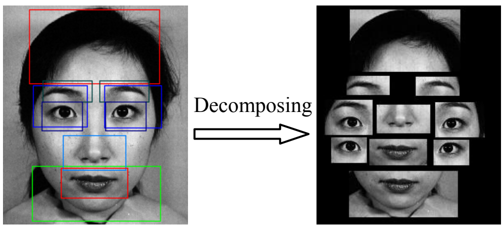


Fig. 3. Example of feature-block decomposing.

If the human face is used as an object, then the eyes, nose, mouth, and forehead will be components of the object and the area that contains the different components will be called a feature block. Therefore, the face can be well described using a small number of feature blocks and their relations. Suppose a human face can be decomposed into n_b feature blocks and n_r corresponding relations. Then, the face can be defined as the model:

$$F = U(B_i, R_j)\ i = 1,\dots,n_b\ \ j = 1,\dots,r_j \tag{6}$$

where B_i is the i - th feature block, and R_j is the j - th relation.

In this paper, the mouth and eye sections are separated primarily, which are the key feature regions of a human face, for FER. This retains the characteristic information and eliminates a significant amount of the noise. The model is defined as:

$$F = U(B_{\text{eyes}}, B_{\text{mouth}}, R_{em}) \tag{7}$$

where B_{eyes} is the eye feature block, B_{mouth} is the mouth feature block, and R_{em} is the relation of above two feature blocks.

3. FEATURE EXTRACTION

3.1 Face Image Preprocessing

To improve the efficiency of extracting facial features, additional facial image preprocessing including feature region localization and face normalization is performed before detecting facial feature points. The ideal output of the processing is a pure facial expression image with normalized intensity and uniform size and shape. In this work, all the images have two eyes aligned and the same sized mouth. The detailed strategy is as follows. To begin, the two eyes and mouth centers are selected manually, determine the corresponding coordinates of these three points, and align the left eye center and right eye center on the same level by rotation. Then, the center coordinate (x_c, y_c) of the pure face is obtained according the equation:

$$\begin{cases} x_c = (x_{le} + x_{re})/2 \\ y_c = (y_{re} + y_m)/2 \end{cases} \tag{8}$$

where (x_{re}, y_{re}), (x_{le}, y_{le}), and (x_m, y_m) are the coordinates of right eye, left eye, and mouth, respectively.

Then, the region is separated with the coordinate (x_c, y_c) as the center and $2 \times (y_m - y_{re})$ pixels as the height and $1.6 \times (x_{re} - x_{le})$ pixels as the width. Lastly, the separated image is resized to 133×126 pixels. The distribution of the key coordinates is indicated in Fig. 4.

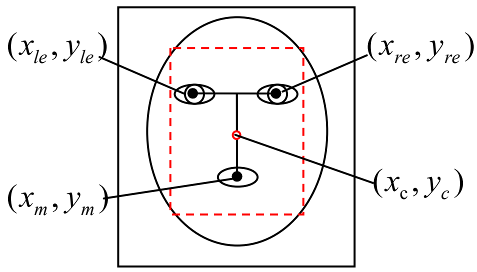
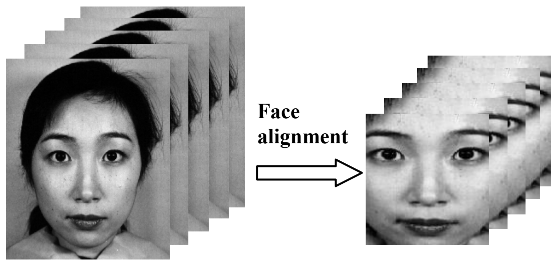


Fig. 4. Distribution of key coordinates.

Unlike a traditional face verification or recognition method, the proposed method requires no other preprocessing steps such as histogram equalization or Gamma correction. After this image adjustment, that LBP is assumed a type of invariant feature. Fig. 5 presents face images with background information and normalized faces.

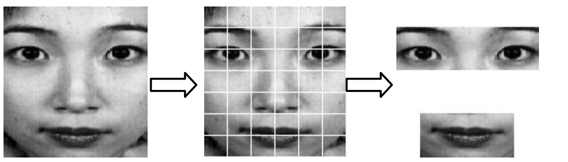


(a) Original images (b) Normalized images

Fig. 5. Facial image preprocessing.

3.2 Acquiring Feature Blocks

It is a fact that some local facial regions contain more useful information for expression classification than others. Considering that the human eyes and mouth contribute significantly to classification based on human experience, the eye and mouth region are cropped as Feature Blocks(FB). To facilitate the operation for calculating LBP features, each face image is first divided into several non-overlapping possible sub-regions. It must be recognized that there are several methods for division; the division scheme of the LBP operator has an influence on the effectiveness of LBP. See Fig. 6 for the steps of acquiring feature blocks. Fig. 6(a) is the gray image. As indicated in Fig. 6(b), the face image is divided into 6×7 non-overlapping sub-regions with a fixed dimension of 19×21 pixels. Then, from Fig. 6(c), it can be seen that an eye feature block image with 126×38 pixels and mouth feature block image with 84×38 pixels are obtained. In this manner, a description of the facial expression can be obtained effectively.



(a) Face image (b) Sub-block (c) Feature blocks

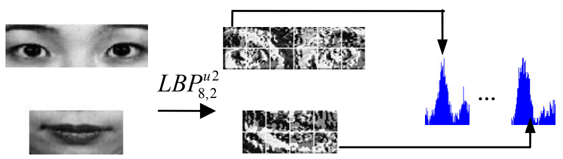
Fig. 6. Acquiring feature blocks.

3.3 LBP Feature Extraction and Selection

An LBP histogram computed over the entire face image encodes only the occurrences of the micro-patterns without any indication regarding their locations. Therefore, to consider the shape information of faces, face images are divided into small equal regions $R_0, R_1, ..., R_m$ to extract the LBP histograms. The LBP features extracted from each sub-region are concatenated into a single, spatially enhanced feature histogram defined as:

$$H_{ij} = \sum_{x,y} I\{f_l(x,y) = i\} I\{(x,y) \in R_j\}$$
$$i = 0,...,n-1,\ j = 0,...,m-1 \tag{9}$$

Fig. 7 is an example of LBP feature extraction where Fig. 7(a) is the original gray image. An LBP operator is used to compute the LBP patterns for each block of the image as illustrated in Fig. 7(b). Then the histogram of the LBP patterns is calculated for each block as presented in Fig. 7(c). The LBP histogram of each square can reflect the edge change, sharpness or flatness of the region, existence of special points, and other characteristics. Finally, the histograms of all the blocks are organized to form a long series of histograms as the feature vector.



(a) Images (b) Block division (c) LBP histogram

Fig. 7. Example of LBP feature extraction.

The extracted feature histograms represent the local texture and global shape of face images. Parameters can be optimized for improved feature extraction. Two of these are the LBP operator and the number of separate regions.

3.4 Feature Dimension Reduction with PCA

Because high-dimensional LBP features obtained using a high-dimension LBP operator cannot significantly improve the classification result, a PCA method is adopted. The purpose of PCA is to reduce the large dimensionality of the data space to the smaller intrinsic dimensionality of the feature space. In this work, the LBP histogram will be processed as a vector. Suppose there are M training samples of size N $\{\chi_1, \chi_2, \chi_3, ..., \chi_M\}$,where the average vector is:

$$\overline{\chi} = \sum_{i=1}^{M} \chi_i \tag{10}$$

Then, the average vector subtracted from the original faces is:

$$\psi_i = \chi_i - \overline{\chi} \tag{11}$$

By ordering the matrix $A = [\psi_1, \psi_2, \psi_3, ..., \psi_M]$, the covariance matrix C is calculated according to:

$$C = \frac{1}{M} AA^T = \frac{1}{N} \sum_{i=1}^{M} \psi_i \psi_i^T \tag{12}$$

According to K-L transform theory, a feature sub-space coordinate system is composed of the eigenvectors corresponding to the nonzero eigenvalues from matrix AA^T . Because it is difficult to acquire the eigenvalues and orthonormalized eigenvectors directly, based on the singular value decomposition principle, the eigenvalues and the eigenvectors of AA^T can be obtained by solving the $A^T A$ eigenvalues and eigenvectors. Therefore, the orthonormalized eigenvectors can be obtained:

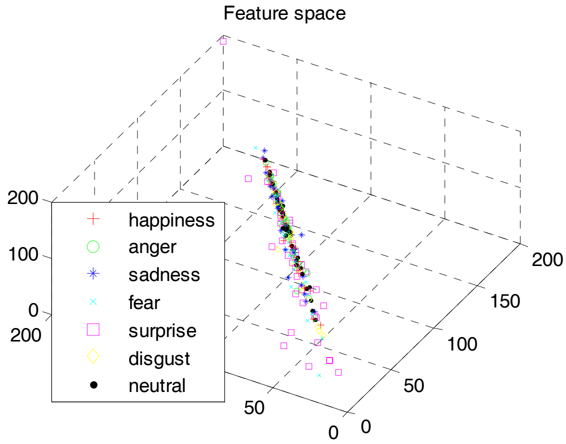
$$v_i = \frac{1}{\sqrt{\lambda_i}} A u_i \ \ (i = 1,2,3,...,k) \tag{13}$$

Then, we select the eigenvectors $(v_1, v_2, ..., v_d)$ corresponding to the front d large eigenvalues ($\lambda_1 \geq \lambda_2 ... \geq \lambda_d$) to build the features sub-space:

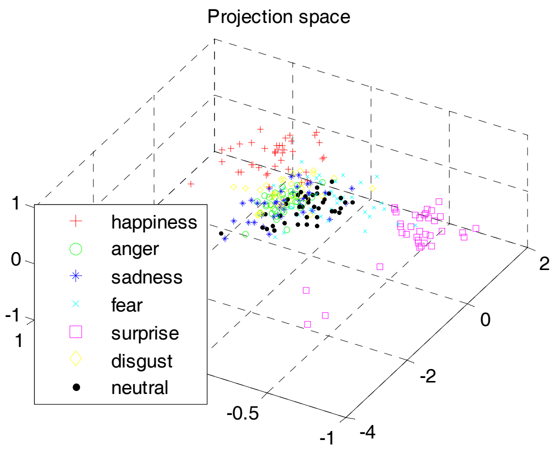
$$X = (v_1, v_2, ..., v_d) \tag{14}$$

With PCA, a linear transform is performed in the high-dimensional high-density LBP feature space and retain the features corresponding to the components with the greatest eigenvalues.

Fig. 8 presents the scattering plots of the training samples in the 3-dimensional (3D) spaces. Fig. 8(a) corresponds to the scattering plots of the training data in the LBP feature space with three random features. The 3D projection space plots in Fig. 8(b) are obtained by applying the PCA method where the first three principal components were plotted. As can be seen from these plots, the between class discrimination increases after the feature dimension reduction using PCA. These features can retain the high-energy information of the high-density LBP feature space and effectively lower the dimensions resulting in lower computational classifying cost.



(a) Distribution in feature space



(b) Distribution in projection space

Fig. 8. Distribution plots of training samples in different 3D spaces.

4. CLASSIFICATION RULE AND SYSTEM FRAMWORK

The feature blocks and PCA-based approach derives the LBP features of training and test face images as discussed in the previous section. After the feature vectors $\{y_i, i = 1,...,t\}$ are extracted, the nearest neighbor classifier is adopted to perform the classification task. The most probable class $\tilde{c}$ of a query face is determined by identifying the neighbor with the minimum distance between the query feature and all prototypes. Let y and y_{cs} denote the feature vectors of the query and all prototypes $\{y_{cs}, 1 \le c \le C, 1 \le s \le M_s\}$, respectively, where C is the number of classes (expressions) and M_s is the number of training images of class c . The minimum distance, which indicates the similarity of two vectors, is determined as:

$$d(y, y_{\tilde{c}\tilde{s}}) = \min_{1\le c\le C,\, 1\le s\le M_s} \|y - y_{cs}\| \qquad (15)$$

where $\|\cdot\|$ computes the distance of two vectors.

In this paper, the Euclidean distance measure is used. Given an arbitrary face image x represented by LBP feature vector set y , which is described as $y = [a_1, a_2, ..., a_n]$, the similarity measured by distance $d(y_i, y_j)$ is:

$$d(y_i, y_j) = \sqrt{\sum_{r=1}^{n} (a_{ir} - a_{jr})^2} \qquad (16)$$

where a_{ir} and a_{jr} denote the value of the r - th attribute of the two feature vectors y_i and y_j , respectively.

The output of a recognition system is a list of sorted reference images in descending order by similarity with the probe (testing) image. That is, the reference image on the top of the list has the highest similarity (minimum distance) to the testing image. The recognition rate is defined as follows:

$$ratio = \frac{Num_{correct}}{Num_{all}} \times 100\% \qquad (17)$$

where $ratio$ is the correct recognition rate, $Num_{correct}$ is the number of correct matches, and Num_{all} is the total number of test images.

Combining the above sections, the procedure for the proposed facial expression is presented in Fig. 9.

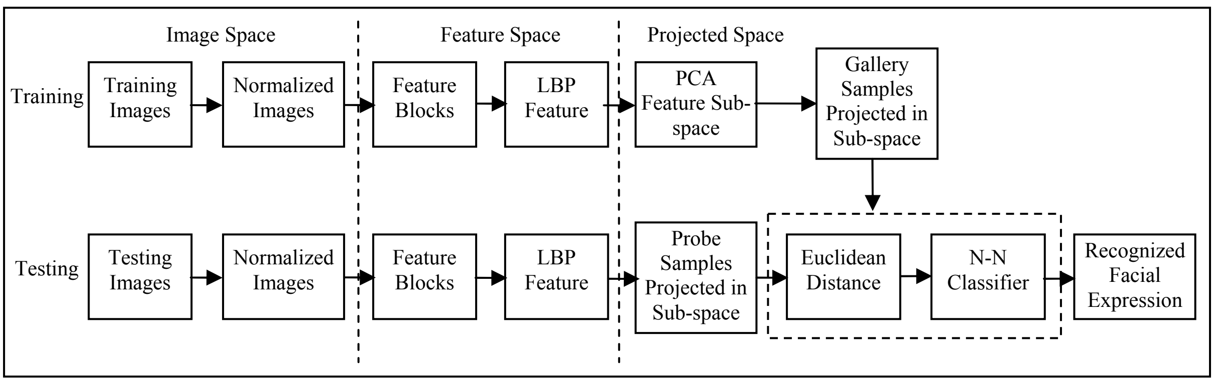


Fig. 9. Diagram of proposed facial expression recognition system.

From Fig. 9, it can be seen that after the feature blocks, which primarily contain eyes and mouth, are obtained, the LBP operator is used for LBP feature extraction. Then, PCA is used to project the images onto a lower dimensional space. Finally, the nearest neighbor classifier based on Euclidean Distance Similarity is used for expression classification.

5 EXPERIMENTAL RESULTS

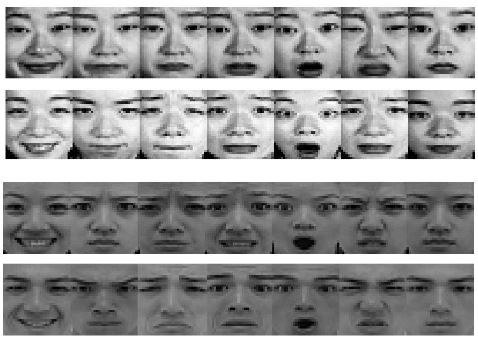
In this section, experiments were performed on international open databases to verify the effectiveness of the proposed algorithm for FER. The proposed algorithm was implemented in MATLAB2008a on a Windows XP platform. All simulations were conducted using a 2.80 GHz Pentium processor and 1 GB memory.

5.1 Facial Expression Databases

The proposed method is tested on the JAFFE and TFEID databases.

JAFFE: This database is a set of 213 images of seven different expressions posed by ten Japanese females with two to four images for each expression. The seven expressions were happiness, anger, sadness, fear, surprise, disgust, and neutral (Kamachi et al., 1998). All the images were cropped manually to ensure the eyes and mouth were at the same positions and resized to 133×126 pixels. Cropped samples are presented in Fig. 10(a).

TFEID: The Taiwanese Facial Expression Image Database (TFEID) was open to public access on 11th December 2007 (Chen and Yen, 2007). In this database, eight categories of images with 40 models (20 males and 20 females) have been collected, each with eight facial expressions: neutral, anger, contempt, disgust, fear, happiness, sadness, and surprise. In our experiments, the contempt expression is excluded and the experiment is focused on the six basic expressions and neutral face. It is found that the downloaded TFEID database to be incomplete, i.e., for some particular persons in the database, some of the expression images were missing. Therefore, about 268 images are used in the TFEID database for our experiments. All the images were cropped such that the background information was removed and the size of the images was normalized to be uniform. Example face images, which have been cropped and aligned, from the TFEID dataset are presented in Fig. 10(b).



(a) JAFFE

(b) TFEID

Fig. 10. Examples of seven expressions of JAFFE and TFEID databases.

5.2 Experimental Results on JAFFE

A cross-validation technique is used to verify the proposed algorithm; the experiment was person-dependent. The entire image set was randomly divided into five groups of approximately equal size. Four groups of the images were used for training; the remaining group was used for testing. For comparison, the results of other FER methods with their highest recognition rate are displayed in Table 1, where the word “FB” is the abbreviation of feature blocks. From this table, it can be seen that the FER performance based on LBP features was significantly superior to those methods based directly on gray level images. The recognition of the proposed method out-performed the other methods with the highest recognition rate attaining 90.14%. Moreover, the number of feature vectors was lower, reducing the computing complexity.

Table 1. Recognition rate.

Methods	Feature Dimensions	Recognition Rates
Gray+PCA	16758	77.00%
LBP+PCA	2478	87.79%
FB+Gray+ PCA	7980	79.81%
FB+LBP+ PCA	1180	90.14%

For above person-dependent experiments, the training samples and testing samples had the same person with different images. Though the recognition rate was high, it did not represent its generalization ability. Therefore, another experiment was designed to test the generalization ability of the proposed algorithm on the JAFFE database. FER rates changing with the dimensions of the feature space are presented in Fig. 11(a).

From Fig. 11(a), it can be seen that the FER rates increased with the dimensions of the feature space at the beginning. For the proposed method, when the dimension m of the feature space approached 35, the expression recognition rates attained their maximum values. When $m > 35$, the recognition rates decreased with m. It should be noted that the best selection for the dimensions of the feature space m in this experiment may or may not be the best choice for other tests. Compared with the person-independent experiment, the recognition rate of the person-dependent experiment was relatively lower; the reason may be that the facial expression database, which contained only ten faces, was small. The experiment results indicate that FB+LBP is also an effective combination for FER.

The recognition rate of proposed method is also compared with the methods in the other references. The numerical comparisons on JAFFE database is shown in Table 2.

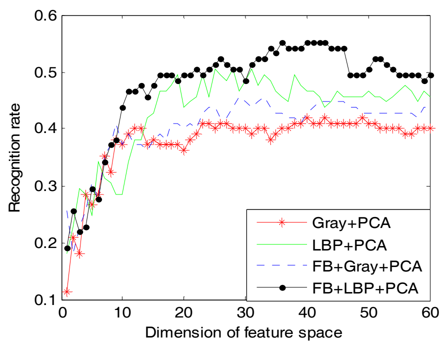
From Table 2, it can be seen that the recognition rate of proposed method is the highest. It indicates that the feature blocks of eye and mouth are more related with the facial expression and have more information for classification.

Table 2. Comparison of different methods.

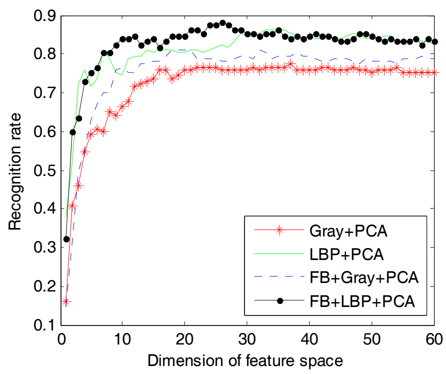
Methods	Recognition Rates
LBP + Template Matching (Wang et al., 2013)	79.1%
Geometrical Analysis + Naive Bayesian Classification (He et al., 2005).	73.2%
Gabor Filter + SVM (Shan et al., 2009)	76.9%
Multiple Local Binary Patterns (Jiang et al., 2013).	83.3%
Complete Local Binary Pattern (Singh et al., 2012)	87.14%
Local Directional Pattern (Jabid et al., 2010).	90.0%
Gabor Parameter Matrix + Adaboost (Yang and Zhang, 2014)	89.67%
Proposed Method	90.14%

5.3 Experimental Results on TFEID

The TFEID database contains only one image per person for each category of expression. Therefore, it can only be used for performing a person-independent experiment. In this experiment, the entire image set was divided into two groups of approximately equal size, with 20 persons (ten males and ten females) for training and the other 20 persons for testing. The training set consisted of 121 images and the testing set consisted of 127 images. The FER rates changing with the dimensions of the feature space are presented in Fig. 11(b).



(a) JAFFE database



(b) TFEID database

Fig. 11. Expression recognition rate changing with dimension of the feature space.

From Fig. 11(b), it is clear that the LBP-based methods were significantly superior to those methods based directly on gray level images and the proposed method continued to demonstrate its competitive performance. This is consistent with the result from JAFFE database. An approximately 88% correct recognition rate occurred when 26 features were used. The relationship between the feature dimension after PCA analysis and the energy is illustrated in Fig. 12.

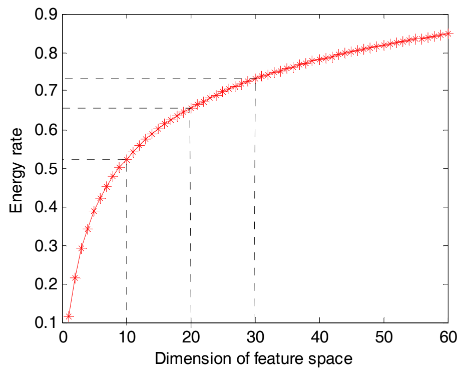


Fig. 12. The energy rate and the dimension of feature space.

From this, it can be seen that the energy rate increased with the dimensions of the feature space, approaching a rate of 74% when the dimension rate was 30.

5.4 Comparison of Different Feature Blocks

Another experiment was designed for testing the performance of different feature blocks (eye blocks, mouth block, eyes + mouth blocks). The comparison curves for the recognition rates under the different dimensions of the feature space are presented in Fig. 13. It is clear that the combination of eye blocks and mouth block indicated superior performance compared to each of them working independently. Further, the recognition rate based on the eye blocks was similar to that based on the mouth block. The reason may be that the eye blocks and mouth block each included important feature information and their combination increased the overall amount of information resulting in a higher FER rate.

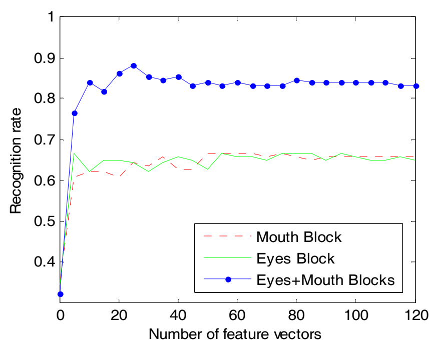
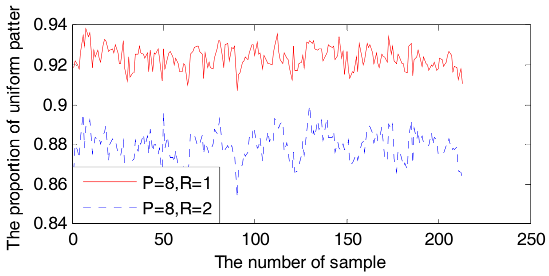


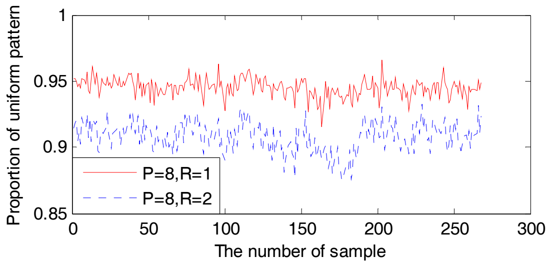
Fig. 13. Comparison of FER rates based on different feature blocks.

5.5 Experiment Results Comparison on JAFFE and TFEID

From above two groups of experiments, it can be determined that FER performance based on LBP features is significantly better than those methods based directly on gray level images. The difference between them is that the recognition rate from the TFEID database is higher than that from the JAFFE database. The reason may be that the number of persons in the TFEID database was greater than that of the JAFFE database. Further, some models from the TFEID database majored in drama or other related fields and others were familiar with facial performance skills; hence their expressions were more consistence and accurate. Another reason could be that the information derived from the original gray image had gaps between the two databases. To verify the guess, the proportion of uniform pattern was studied on the facial expression images from both databases. The result is displayed in Fig.14.



(a) JAFFE database



(b) TFEID database

Fig. 14. Proportion of uniform LBP pattern.

The upper red curves reveal the proportion of uniform pattern of each facial expression image filtered by the $LBP_{8,1}^{u2}$ operator. The proportion of uniform pattern averaged up to 92.37% on the JAFFE database and 94.51% on the TFEID database. The results of the $LBP_{8,2}^{u2}$ operator are displayed in the lower blue curves, accounting for an average of 87.93% of the total information on the JAFFE database and 90.79% on the TFEID database. From this data, it can be concluded that the information extracted from the original images of the TEEID database was richer than that of the JAFFE database.

5.6 Experimental Results Analysis

From the discussion above, a number of experimental findings can be extracted from the results:

(1) FER performance based on LBP features is significantly better than those methods based directly on gray level images.

(2) The proposed algorithm performed well on the JAFFE database and better on the TFEID database with a higher recognition rate.

(3) It is clear that the combination of eye blocks and mouth block had a superior performance than either of them separately.

6. CONCLUSION

In this paper, a novel algorithm for FER was proposed based on feature blocks and LBP. The method introduced utilized the well-known framework of linear space analysis: feature blocks contain the main information for recognition, the LBP feature is invariant to illumination and rotation, and PCA can reduce the dimensionality of features in linear space. The effectiveness of the proposed method was demonstrated through experimentation and excellent results were obtained on the JAFFE and TFEID databases. Compared with the nonlinear model algorithm, the proposed algorithm was computationally simpler without overfitting. It was effective when applied to images with a variety of different illumination conditions. Though the proposed method can improve the total recognition ratios, it requires additional research and improvement, especially in the method to optimally combine the eye blocks with the mouth block. A future goal is to implement this method for selecting and cropping additional feature blocks resulting from the feature representation of face images and perform further research for an optimal joint strategy of feature blocks and applying the method for facial expression classification. In summary, FER rates remain relatively low for practical applications and there is considerable improvement required for FER to achieve satisfactory recognition rates.

ACKNOWLEDGMENTS

This work is supported by National Nature Science Foundation of China (No. 61403283, 61273277), Shandong Provincial Natural Science Foundation, China (No. ZR2013FQ036) and Technology Development Plan of Weifang City(No. 201301015). It would also like to express our thanks to Brain Mapping Laboratory (National Yang-Ming University), they offer us the TFEID database for testing our method.

REFERENCES

Abboud B. and Davoine F.(2004). Appearance factorization for facial expression recognition and synthesis. *Proceeding of International Conference on Pattern Recognition*, pp.163–166.

Ahonen, T., Hadid, A. and Pietikainen, M.(2004). Face recognition with local binary patterns, *8th European Conf. Computer Vision*, pp.469–481.

Chen, L. and Yen, Y.(2007). Taiwanese facial expression image database. *http://bml.ym.edu.tw/download.html, 2007.*

Danisman, T., Bilasco, I.M., Martinet, J. and Djeraba, C.(2013). Intelligent pixels of interest selection with application to facial expression recognition using multilayer perceptron. *Signal Processing*, 93(6), pp.1547–1556.

Domaika, F. and Davoine, F.(2008). Simultaneous facial action tracking and expression recognition in the presence of head motion, *International Journal of Computer Vision*, 76(3), pp.257–281.

Ekman, P. and Friesen, W.V.(1971). Constants across cultures in the face and emotion. *Journal of Personality Social Psychology*, 17(2), pp.124–129.

Fasel, B. and Luettin, J.(2003). Automatic facial expression analysis: a survey, *Pattern Recognition,* 36, pp.259–275.

Feng, X., Pietikäinen, M. and Hadid, A.(2005). Facial expression recognition with local binary patterns and linear programming. *Pattern Recognition and Image Analysis*, 15(2), pp. 546–548.

Hadid, A., Pietikäinen, M. and Ahonen, T.(2004). A discriminative feature space for detecting and recognizing faces. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

He, X., Yan, S. and Hu, Y.(2005). Face recognition using laplacianfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3), pp. 328–340.

Jabid, T., Kabir, M. and Oksam, C.(2010). Facial expression recognition using local directional pattern. *Proceeding of IEEE International Conference on Image Processing,* Hongkong, pp.1605–1608.

Jiang, R. Xu, J. and Zhang, A.(2013). Facial expression recognition based on improved lbp. J*ournal of Zhejiang sci-Tech University*. 30(4), pp.546–549.

Kamachi, M., Lyons, M. and Gyoba, J.(1998). The Japanese female facial expression database. *http://www.kasrl. org/jaffe.html*.

Khan, R.A., Meyer, A., Konik H. and Bouakaz S.(2013). Framework for reliable, real-time facial expression recognition for low resolution images. *Pattern Recognition Letters*, 34(10), pp.1159–1168.

Kotsia, I. and Pitas, I.(2007). Facial expression recognition in image sequences using geometric deformation features and support vector machines. *IEEE Transactions on Image Processing*, 16(1), pp.172–187.

Liu, C. and Wechsler, H.(2002). Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Transaction on Image Processing*, 11(4), pp.467–476.

Lyons, M., Budynek, J. and Akamatsu, S.(1999). Automatic classification of single facial images. IEEE *Transactions on Pattern Analysis and Machine Intelligence*, vol.21, no.12, pp.1357–1362.

Maalej, A. and Amor, B.(2011). Shape analysis of local facial patches for 3D facial expression recognition. *Pattern Recognition*, 44(8), pp.1581–1589.

Ojala, T., Pietikainen, M. and Maenpaa, T.(2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7), pp. 971–987.

Rosenblum, M., Yacoob, Y. and Davis, L.(1996). Human expression recognition from motion using a radial basis function network architecture. *IEEE Transactions on Neural Networks*, 7(5), pp.1121–1138.

Valstar, M. and Mehu, M.(2012). Meta-analysis of the first facial expression recognition challenge. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 42(4), pp.966–979.

Wan, S. and Aggarwal, J.K.(2014). Spontaneous facial expression recognition: A robust metric learning approach. *Pattern Recognition*, 47(5), pp.859–1868.

Wang, J., Feng, J. and Han, Z.(2013). Locally preserving PCA method based on manifold learning and its application in fault detection. *Control and Decisi*on, 28(5), pp. 683–687.

Wang W., Chang F., Zhao J. and Chen Z.(2010). Automatic Facial Expression Recognition Using Local Binary Pattern. *WCICA2010*, pp. 6375– 6378.

Wang, W. and Chang, F.(2014). Adaptive ratio local binary pattern with application to eye location. *Journal of Multiple-valued Logic and Soft Computing*, 23(1), pp.209–234.

Wang, W. and Chang, F.(2014). Facial expression recognition based on feature block and local binary pattern. *Communications in Computer and Information Science*, 472, pp.447-451.

Shan, C., Gong, S. and Mcowan, P.(2009). Facial expression recognition based on local binary patterns: a comprehensive study. *Image and Vision Computing*, 27(6), pp. 803–816.

Singh, S., Maurya, R. and Mittal, A.(2012). Application of complete local binary pattern method for facial expression recognition. *Proceedings of IEEE International Conference on Intelligent Human Computer Interaction*, Kharagpur, pp.1–4.

Turk, M. and Pentland, A.(1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3, pp.71–86.

Yang, F. and Zhang L.(2014). Facial expression recognition based on Gabor Parameters Matrix and improved Adaboost. *Journal of Computer Application*, 34(4), pp.1134–1138.

Yang, P., Liu, Q. and Metaxas, D.(2009). Boosting encoded dynamic features for facial expression recognition. *Pattern Recognition Letters*, 30(2), pp.132–139.

Ying, Z., Zhang, Y. and Li, J.(2009). Manifold learning approach to facial expression recognition on local binary pattern features. *ICMLC2009*, pp.405–409.

Yu, K., Wang, Z. and Hagenbuchner, M.(2014). Spectral embedding based facial expression recognition with multiple features. *Neurocomputing*, 129, pp.136–145.

Zhao, X. and Zhang, S.(2011). Facial Expression Recognition Based on Local Binary Patterns and Kernel Discriminant Isomap. *Sensors (Basel)*, 11(10), pp.9573–9588.

Zhan, Y. and Cheng, K.(2010). A new classifier for facial expression recognition: fuzzy buried markov model. *Journal of Computer Science and Technology*, 25(3), pp. 641–650.

Zhu, J., Su, G. and Li, Y.(2006). Facial expression recognition based on Gabor feature and Adaboost. *Journal of Optoelectronics Laser*, 17(8), pp.993–998.