International Journal of Hybrid Information Technology
Vol.8, No.6 (2015), pp.57-64
http://dx.doi.org/10.14257/ijhit.2015.8.6.06

Feature Extraction for Facial Expression Recognition based on Ensemble Learning Algorithm

Shang-fu Gong, Juan Chen, Peng-tao Jia

College of Computer Science and Technology, Xi'an University of Science and Technology
Xi'an 710054, China
mamawork@sina.com

Abstract

The accuracy of facial expression feature extraction directly influences the recognition rate of facial expression. In order to extract facial expression feature effectively, this paper puts forward a new way of facial expression feature extraction ensemble learning algorithm based on ensemble thinking. The superposition method of heteromorphy ensemble learning is used to construct an ensemble learning model, two-d gabor wavelet, block local binary patterns and two-directional two-dimensional principal component analysis as the single learning device. Firstly, two-d gabor wavelet was used to get image texture information at each level. Then the image was divided into some blocks to acquire the eigenvectors of block local binary patterns, and reduced its dimensionality. Next, the dimensionality was further reduced by two-directional two-dimensional principal component analysis, extracted the effective characteristic features at the same time. Finally, it is classified by nearest neighbor classifier on the extracted feature library. Experimental results on JAFFE expression database show that this ensemble learning model gets higher recognition rate and better generalization ability than single learning device.

Key words: *Feature extraction, Ensemble learning, Facial expression recognition*

1. Introduction

Facial expression recognition is the process of using the computer to extract feature from facial expression information, and classify them. From the facial expression information, computers can conclude the people's state of mind, so as to realize human-computer intelligent interaction [1]. In recent years, facial expression recognition technology has a variety of applications in many fields such as intelligence robot research, safety tests, computer games, educational experiments and terrorists' identification [2]. It becomes one of the hot research topics in pattern recognition.

In the process of facial expression recognition, the capability of expression feature extraction has direct effect to the final recognition rate and identifying performance, it is the core element. At present, there are some common methods to extract expression feature, such as principal component analysis (PCA) [3,4], Gabor wavelet [5,6], geometrical feature extraction, local binary patterns (LBP) [7,8], singular value decomposition (SVD) [9,10] and so on. These methods obtained the certain result in practice, existed some shortage at the same time. PCA bases on K-L transformation, reserves the original data's difference to the largest extent. PCA is used for feature extraction and dimensionality reduction. But this algorithm's shortage is that the dispersion of sample classes increased the dispersion within sample class increased all along with it [11]. The method of gabor wavelet can preserve the expressional details well, and it is not sensitive to the variation in light intensity, but the dimension is higher after sample. Geometrical feature extraction reduces the input feature information, compresses the information, and poor quality of accuracy. LBP is suitable for extracting texture characteristic and keep enough expression image information. Unfortunately, there is no fixed method to divide the pattern statistical panes. SVD can extract characteristics under the layer of rough-classification quickly and simply. But the facial expression recognition based on SVD depends on the training set selection. Although many scholars improved these algorithms, and achieved certain effect, it is still difficult to overcome the self-defect of the single learning algorithm. If we can integrate these ways to learn form each other, it can improve the facial expression recognition rate effectively.

This paper introduces the ensemble thinking into the filed of expression feature extraction, the superposition method of heteromorphy ensemble learning is used to construct an ensemble l- earning model. Using two-d gabor wavelet, block local binary patterns (Block-LBP) and two-directional two-dimensional principal component analysis (2D-2DPCA) as the single learning device. Experimental results on JAFFE expression database show that this ensemble learning method improves the facial expression recognition rate and the generalization ability.

2. Research on the Feature Extraction Ensemble Learning Algorithm

2.1. Ensemble Learning

Ensemble learning is a new kind of machine learning pattern, it uses multiple learning devices to solve the same problem [12, 13]. Originally, scholars found the performance of integrating a set of neural network was higher than a single neural network. So, the earliest ensemble learning method called Bayesian Averaging appeared. After that, ensemble learning has attracted many people's attention. It can learn groups of algorithms in some way, then it significant increases the generalization ability of machine learning algorithm [14, 15], and enhances the accuracy and stability of the predict results. Its advantage makes it become the head of machine learning research rapidly and the focus in international machine learning research. It has broad application in some fields such as information filtration and biometric features recognition [16].

Ensemble learning can be divided into homomorphic ensemble learning and heteromorphy ensemble learning on the basis of the essence of learners. Homomorphic ensemble learning uses the same class learners. Heteromorphy ensemble learning integrates different kinds of learners. There are some common and different things in kinds of algorithms of expression feature extraction. This paper chooses suitable feature extraction method as the single learner that based on the heteromorphy ensemble learning. And then fuses the each learner’s result to strengthen the advantage of extracting common features and weaken the bad effects of shortcomings.

2.2. Ensemble Learning Algorithm of Feature Extraction

In the process of feature extraction of facial expression recognition, it demands to carry with effective expression feature information as much as possible, reinforce the information content of feature. On the other hand, try to reduce the redundant information, speed up the information processing. Through studying the facial feature, we found, texture feature contains a lot of expression information about gray level, outline change and so on, it is very suitable as the main expression feature to be extracted. In the field of texture feature extraction, we compared every feature extraction algorithm. From different directions, two-d gabor wavelet is able to detect the edge feature whose gray level is dramatic changing, and catch lots of texture feature information. This meets the requirement that extracts large scale of feature information. But because after two-d gabor wavelet extracts feature, the dimension is too high to improve the processing speed. Block-LBP is not only get texture information rapidly from the grayscale images, but also represent the relationships between the features. It increases the feature information content, and also realizes the fast extraction, reduces the dimensions to some extent at the same time. 2D-2DPCA can retain the most of facial information and strengthen the effective feature extraction. It satisfies the feature extraction's demand about effective characteristics, meanwhile 2D-2DPCA reduces the feature dimension from the row and column two directions, cuts down the redundant information in feature extraction and increases the processing speed.

By the requirements of feature extraction and analysis of the above, this paper extracts texture features via the heteromorphy ensemble learning base on the three different kinds of feature extraction methods. In the side of feature extraction, two-d gabor wavelet achieves the purpose of large quantity of information, Block-LBP reaches fast and enhance information, 2D-2DPCA realizes that takes out effective characteristics from numbers of characteristics. In the side of dimensionality reduction and processing speed, Block-LBP and 2D-2DPCA can complementary handle the high-dimensional problem after two-d gabor wavelet extracted feature. This relationship of progressive and complementary conforms to the superposition method of heteromorphy ensemble learning, what distributes basic learners on several levels, and then finishes the learning tasks by multilayer learners [17]. The meaning is the result of previous learner enters the next learner, and then the performance of the system is better than a single learner.

In conclusion, this article selects three methods like two-d gabor wavelet, Block-LBP and 2D-2DPCA to integrate in series superposition. On the one hand, the integrated algorithm gradual increases the extraction of expression texture features, on the other hand, after feature extraction, it can complementary reduce dimension to handle the high-dimensional problem. So, it can improve the recognition rate and the generalization ability in the end.

This paper centers on extracting texture features, adopts the superposition method of heteromorphy ensemble learning to construct the model of feature extraction ensemble learning algorithm for facial expression recognition. The model is shown in Figure 1.

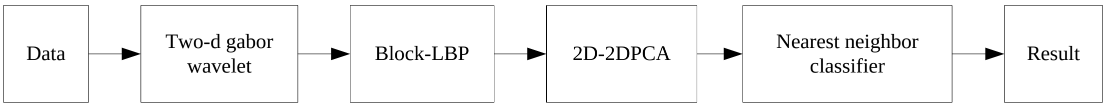

Figure 1. Structure of Feature Extraction Ensemble Learning Algorithm

Feature extraction has the following steps:

(1) Two-d gabor wavelet gets the expression texture features over all scales, and amplifies some local features.

Input: Expression images after pretreatment

Output: Characteristic image after the gabor filters processed

Step1 This paper sets three center frequencies ($v = 0,1,2$) and four directions ($u = 0,1,2,3$) to sample, and constructs twelve filters by formula (1).

$$G_j(\vec{x}) = \frac{\left\| \begin{pmatrix} k_v \cos\phi_u \\ k_v \sin\phi_u \end{pmatrix} \right\|^2}{\sigma^2} \exp\left(-\frac{\left\| \begin{pmatrix} k_v \cos\phi_u \\ k_v \sin\phi_u \end{pmatrix} \right\|^2 \|\vec{x}\|^2}{2\sigma^2} \right) \left[\exp\left(i \begin{pmatrix} k_v \cos\phi_u \\ k_v \sin\phi_u \end{pmatrix} \vec{x} \right) - \exp\left(-\frac{\sigma^2}{2} \right) \right], \tag{1}$$

$$(k_v = 2^{-\frac{v+2}{2}}\pi, \phi_u = u\frac{\pi}{8})$$

Step2 In order to improve the computing speed, we uses 2-D Fast Fourier Transformation to make gabor transformation into the mode of Fourier.

Suppose there is an image after the pretreatment, and $I(x, y)$ is the grayscale of a point that in the image. The definition of linear convolution is $y(n) = h(n) * x(n)$, according it, we can obtain the information about the image that after gabor filters processed, and the information is as follows.

Real part:

$$\mathrm{Im}ag_{u,v}(x,y) = I(x,y) * IG_{u,v}(x,y) \tag{2}$$

Imaginary part:

$$\mathrm{Re}al_{u,v}(x,y) = I(x,y) * RG_{u,v}(x,y) \tag{3}$$

Where the $IG_{u,v}(x,y)$ and $RG_{u,v}(x,y)$ represent the gabor function components after decomposition. Then the filtered image is as follows.

$$J_{u,v}(x,y) = \sqrt{\mathrm{Im}ag_{u,v}(x,y)^2 + \mathrm{Re}al_{u,v}(x,y)^2} \tag{4}$$

The result is $J_{u,v}(x,y)$, it means the characteristic image after the gabor filters processed.

(2) Gabor wavelet transformation brings the problem of high-dimension after extracts texture features. LBP not only can rapid measure and extract texture features in local adjacent area from the grayscale images, and keep enough expression image information, but also can decrease dimension, at some level. So after the two-d gabor wavelet transformation, the Block-LBP is used to extract texture features again, and decrease dimension.

Input: Characteristic image after the gabor filters processed is $J_{u,v}(x,y)$

Output: The Block-LBP's feature vector of the characteristic image $J_{u,v}(x,y)$

Step1 This paper employs uniform LBP operator that is $LBP_{P,R}^{u2}$. It describes the most of texture features in the image effectively, and cuts down the number characteristics at the same time. First of all, the characteristic image $J_{u,v}(x,y)$ after the gabor filters processed is divided into several no-repeat regions like $J_0, J_1, \cdots, J_{m-1}$. Considering the processing of the images in the database and the final recognition rate, this article chooses the mode of blocks is 16*4. Then the LBP histogram of each region is calculates, and the mathematical expression is as follows.

$$H_{i,t} = \sum_{x,y} I\{J_{u,v}(x,y) = i\} I\{(x,y) \in J_t\}, (i = 0,1,\cdots,n-1, t = 0,1,\cdots,m-1) \tag{5}$$

Step2 The LBP histogram $H_{i,t}$ of each region is stringed together in a specified order to form a integral and long string of histogram. The formula (6) is the calculation formula.

$$Y = [Y, H_{i,t}] \tag{6}$$

Where Y records the Block-LBP's feature vector of the characteristic image what after the gabor filters processed.

(3) Although in the concept of Block-LBP, the more blocks of the image, the performance is better, the dimension of feature vector is increasing with the increasing of the number of blocks. Therefore, the effect of reducing dimension is weakened. So, this paper introduces 2D-2DPCA that can directly base on the image matrix, and reduce the feature dimension from the row and column two directions. On the one hand, 2D-2DPCA decreases the number of features, makes the expressions of the features more compact, and reduces the computation complexity. On the other hand, it extracts the effective feature once more that is beneficial to classification and identification.

Input: The image matrix Y_i is constituted by the Block-LBP's feature vector Y

Output: The characteristic matrix after the 2D-2DPCA processed

Step1 The feature vector Y after the Block-LBP processed constructs the image matrix $Y_i(\{Y_1, Y_2, \cdots, Y_M\})$, and the covariance matrix of the image is as follows.

$$G_t = \frac{1}{M} \sum_{i=1}^{M} (Y_i - \bar{Y})^T (Y_i - \bar{Y}) \tag{7}$$

Where the $\bar{Y}$ is the mean of all the training samples. Suppose the first d maximal eigenvalues in the G_t are $\lambda_1, \lambda_2, \cdots, \lambda_d$, and the corresponding eigenvectors are $u_1, u_2, \cdots, u_d$, then the matrix $U = [u_1, u_2, \cdots, u_d]$ is the best transformation matrix. The characteristic matrix of each sample is as follows.

$$B_i = \left(Y_i - \bar{Y}\right)U \tag{8}$$

Step2 According to the formula (8), we get the new train sample set that is $B_i\left(\{B_1, B_2, \cdots, B_M\}\right)$. The corresponding covariance matrix is as follows.

$$G_t^B = \frac{1}{M}\sum_{i=1}^{M}\left(B_i - \bar{B}\right)\left(B_i - \bar{B}\right)^T \tag{9}$$

Where the $\bar{B}$ is the mean of the new training samples. Through calculation, the first k maximal eigenvalues in the G_t^B corresponding eigenvectors are $l_1, l_2, \cdots, l_k$. So the optimal transformation matrix V is constituted. We can get the linear transformation expression of the 2D-2DPCA as follows.

$$P = V^T\left(Y\text{-}\bar{Y}\right)U \tag{10}$$

Finally, the P not only is the characteristic matrix after 2D-2DPCA processed, but also is the characteristic matrix after the three feature extraction algorithms processed.

This paper makes the three feature extraction algorithms integrate in series superposition learning. In the side of feature extraction, it realizes the progressive strengthen extraction of the texture features in the expression images, highlights the eigenvalues, increases the amount of feature information and the validity of feature extraction. In terms of feature dimension, on the basis of guaranteeing the recognition rate, it cuts down the feature dimension as much as possible successively in order to increase processing speed.

3. Experiment and Results Analysis

3.1. Introduction of the Database

This paper did the experiments based on the Japanese female facial expression database (JAFFE). The database contains 213 expression images that are come from 10 Japanese women. And every woman shows 7 kinds of expression. They are anger, disgust, fear, happy, sad, surprise and neutral. Every woman poses the same type of expression would have 2 or 3 images.

This paper uses MATLAB as the simulation platform, and selects the 7 kinds of expression (anger, disgust, fear, happy, sad, surprise and neutral) in JAFFE to do the experiment. For every woman, from every kind of expression images, 1 to 2 images was chose to be the training samples, and the other images as the test samples.

3.2. Feature Extraction Ensemble Learning for Facial Expression Recognition

According to the process of facial expression recognition, the procedural details of the experiment based on the JAFFE are as follows.

(1) The acquisition of image: Untreated facial expression images in JAFFE database are shown in Figure 2.

Figure 2. Untreated Facial Expression Images in JAFFE Database

(2) Image preprocessing: Through the analysis of the image's grayscale distribution, the linear gray transformation was adopted to adjust the image's grayscale that could be well separated with the facial image and the background. Then, clipped the image and did some other processes to achieve the purpose of standing out the face. Finally, the gray level of the image was equalized in order to eliminate the influence of light. The images after preprocessing are shown in Figure 3.

Figure. 3. The Images after Preprocessing

(3) Expression feature extraction:

① Two-d gabor wavelet transformation

The before and after images of processed by two-d gabor wavelet transformation are shown in Figure 4.

Figure 4. The Images after Processed by Two-d gabor Wavelet Transformation

② Block-LBP

The images after processed by two-d gabor wavelet transformation would get their characteristics' histogram by Block-LBP. The expression image called "Anger" was shown in figure 3 were used as an example, the image was processed by two-d gabor wavelet transformation firstly, then it got its Block-LBP characteristics' histogram what is shown in Figure 5. According to the order like top to bottom and left to right, the Block-LBP characteristics' histogram shows the LBP eigenvalue of each pixel in every region. In figure 5, the X axis means the pixels in the image and they were arrayed by the order like top to bottom and left to right, the Y axis means the LBP eigenvalue of the corresponding pixel.

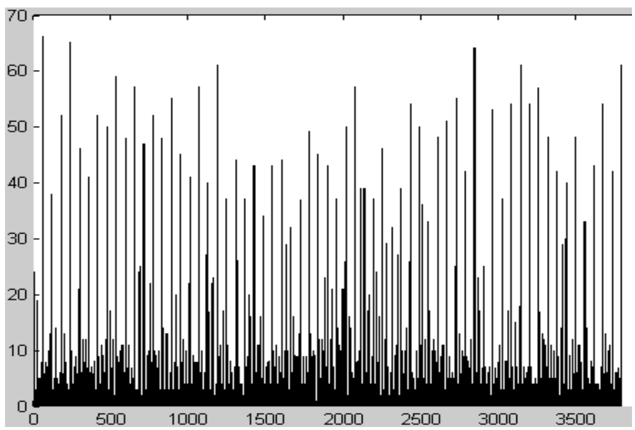

Figure 5. The Characteristics' Histogram after Block-LBP Processed

③ 2D-2DPCA

By 2D-2DPCA, the original high-dimensional feature would drop to the stated low dimensional feature, and then used in the classification and recognition.

(4) Classification and recognition: This paper uses the nearest neighbor classifier to classify and identify, the distance metric is Euclid distance. Calculated the Euclid distance between the test sample's feature and training sample's feature. Then from the training sample, got the label of the nearest point, and the label would be the label of the test sample.

3.3. The Analysis of Experiment Results

With the above analysis and experiment, the comparative results are shown in table 1.

Table 1. The Comparative Results of the Experiment

Method / Label	Two-d gabor wavelet	Block-LBP	2D-2DPCA	Ensemble learning
Anger	80.00%	100.00%	100.00%	100.00%
Disgust	90.91%	100.00%	100.00%	100.00%
Fear	50.00%	91.67%	66.67%	100.00%
Happy	66.67%	75.00%	83.33%	91.67%
Neutral	90.00%	100.00%	100.00%	100.00%
Sad	72.73%	81.82%	72.73%	90.91%
Surprise	100.00%	100.00%	100.00%	100.00%
Average rate	77.63%	92.11%	88.16%	97.51%

From Table 1, the experiment used the 7 kinds of expressions in the JAFFE database. The recognition rates for the expressions of anger, disgust, fear, neutral and surprise are all 100% by the algorithm of the feature extraction ensemble learning that is put forward in this paper. The recognition rates for the expressions of happy and sad reach 91.67% and 90.91% respectively. They are all higher than the recognition rates of the other three single methods. The average recognition rate of the ensemble learning algorithm is 97.51% that is clearly higher than the same rates of the other three single feature learners. So the ensemble learning algorithm achieves better recognition effect, verifies the effectiveness and the generalization ability of the method. But the algorithm of the feature extraction ensemble learning costs more time than the single method.

4. Conclusion

This paper analyzes the feature extraction methods of facial expression recognition. Then bases on the serial learning idea of superposition method of heteromorphy ensemble learning, we make two-d gabor wavelet, Block-LBP and 2D-2DPCA integrate in series superposition to increase the extraction of expression texture features gradually and reduce dimension. The nearest neighbor classifier is adopted to realize classification and identification. Experimental results on JAFFE expression database show that this ensemble learning algorithm has higher recognition rate and better generalization ability, and verifies the effectiveness. But this ensemble learning algorithm brings some problem, such as：

(1) The new algorithm gets better classification results in these experiments, but they have a little long of the execution time of the system, how to improve the efficiency of processing speed under the condition of ensure the identification, even improve the identification efficiency;

(2) The new algorithm is proposed in this paper is used in the static expression database, the future work is to use it into the reality, and satisfy the demand of people's lives.

Acknowledgements

The project is supported by the Scientific Research Program Funded by Shaanxi Provincial Education Department (Program NO. 2013JK0870)

References

[1] W. Li, "Research of Human Facial Expression Recognition," Shaanxi: Shaanxi Normal University, **(2011)**, pp.5-10.
[2] S. Giripun, P. Bajaj, "Recognition of Facial Expressions for Images using Neural Network," International Journal of Computer Applications, Vol. 40, No. 11, **(2012)**, pp. 3-7.
[3] Miao, "A Hierarchical Multiscale and Multiangle System for Human Face Detection in a Complex Background Using Gravity- Center Template," Pattern Recognition, Vol. 32, No. 7, **(1999)**, pp. 1237–1248.
[4] M.-W. Huang, "Face Expression Recognition Base on PCA and SVM," Computer Knowledge and Technology, Vol. 9, No. 19, **(2013)**, pp. 4485-4487.
[5] J. Daugman, "Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two dimensional visual cortical filters," Journal of the Optical Society America A, **(2011)**, pp. 1160—1169.
[6] J. Zhang, Y.-Z. Zhan, Q.-R. Mao, "Facial Expression Recognition Based on Gabor Wavelet and Sparse Representation," Computer Engineering, Vol. 38, No. 6, **(2012)**, pp. 207-209.
[7] T Ojala, M Pietikainen, D. Harwood, "A comparative study of texture measures with classification based on featured distributions," Pattern Recognition, Vol. 29, No. 1, **(1996)**, pp. 51-59.
[8] W.-F. Liu, S.-J. Li, Y.-J. Wang, "Facial expression analysis using LBP features," Computer Engineering and Applications, Vol. 47, No. 2, **(2011)**, pp. 149-152.
[9] Z.-Q. Hong, "Algebraic feature extraction of image for recognition," Pattern Recognition, Vol. 24, No. 3, **(1991)**, pp. 211-219.
[10] L.-X. Li, L.-H. Su, "Double level classification method of facial expression recognition," Fujian Computer, Vol. 10, **(2010)**, pp. 94-96.
[11] T. GONG, "Research on facial expression recognition," Zhejiang: Zhejiang University of Technology, **(2009),** pp. 9-12.
[12] X. Li, C.-J. Fan, "Research of ensemble learning in brain computer interface classification algorithm," Journal of electronic measurement and instrument, Vol. 25, No. 11, **(2011)**, pp. 940-945,.
[13] U. Tariq, K.-H. Lin, Z. Li, "Emotion recognition from an ensemble of features," Proceeding on Automatic Face & Gesture Recognition and Workshops, **(2011)**, IEEE International Conference, pp. 872-877.
[14] S. Ding, "An Ensemble Learning Algorithm Based on the Feature Extraction by Wavelet Transform," Henan: Henan University, **(2010)**, pp. 22-28.
[15] M.B. Mariappan, M. Suk, B. Prabhakaran, "Facial Expression Recognition Using Dual Layer Hierarchical SVM Ensemble Classification," in Proceeding on Multimedia, IEEE International Symposium, **(2012),** pp. 104-107.
[16] L.-K Feng, "The Research of Cover Algorithm based on Ensemble Learning," Anhui: Anhui University, **(2010)**, pp. 14-18.
[17] F. Wang, "Ensemble Learning and its Application in Genetic Data Analysis," Nanjing: Nanjing Normal University, **(2010)**, pp. 25-30.

