Review of Feature Extraction Methods Based on Facial Expression Recognition

Li Yang

Civil Aviation Flight University of China,
Guanghan 618307, Sichuan, China

Abstract
This paper introduce the methods of facial expression recognition, and describe different methods of feature extraction in expression recognition, including the method based on geometric feature, the method based on feature face, the method based on wavelet analysis, more features fusion and combination method, and simple analysis of various methods. Look ahead to the future research emphasis and difficulties of facial expression recognition.
Key words: FACIAL EXPRESSION RECOGNITION, FEATURE EXTRACTION, METHODS

1. Introduction

Facial expression recognition is divided into three core areas: face detection, facial expression feature extraction and facial expressions classification of emotion. After face is detected location, according to different face paint method using different facial expression feature extraction method to extraction information, then according to the extraction information to classification. Facial feature extraction is the key step in facial expression recognition, is the key to the recognition technology, facial feature extraction algorithm needs of face features of human organs, texture area, location and extracted the predefined feature points. In face detection at home and abroad have done lots of research, and existing relevant effective methods and results reported, but the research of expression feature extraction algorithm is still in the exploration. Due to the face image has more information and complexity mode, dimensionality of the image and decorrelation transformation such as PCA, ICA and wavelet analysis, symmetry transformation, optical flow analysis method etc. has been widely used. At present, the method of feature extraction is principal component analysis PCA[1], LDA[2]independent component analysis ICA[3], CDA[4], the sorting PCA +LDA[5], MPP-CA[6] Etc. Although the domestic and foreign researchers have proposed some feature extraction methods, but as a result of many factors that affect the performance of facial expression recognition, such as environment, illumination, age, posture, factors such as image resolution, and imaging noise, so the methods of feature extraction need to be improved.

Based on the introduction and comparison of various facial expressions feature extraction methods, introduced some kinds of method, and new progress on the basis of the comparison, and prospected the future research direction.

The following details several extraction methods of face feature in recent years.

2. The method based on geometric features

2.1 The principle of facial expression extraction method based on geometric features

Geometric feature method extraction by faces important features such as eyes, nose, mouth and chin point position, scale, and the ratio of each other, expression face with a group of geometric feature vector. But the result of this method is more sensitive to expression changes, is not the mainstream, only combined with other methods can achieve good results.

2.2 Many feature integrated classifier of facial expression recognition

The process of the many feature integrated classifier algorithm of facial expression recognition is, first of all, the facial expression of the image after preprocessing, through three different feature extraction method to extract the characteristics of different types of expression, then according to different characteristics building different classifiers, finally construct an integrated classifier based on neural network model, according the output of the three classifiers to decision fusion, to realize expression recognition. In the JAFFE facial expression database test results show that, recognition effect of proposed algorithm is superior to the individual characteristics and the single classifier.

The method put forward facial expression recognition system based on the many feature integrated classifier, make full use of the information of face image, for the expression of the image after pretreatment, through three different feature extraction method to extract the characteristics of different types of expression, include improve the Gabor feature, $(2D)^2$DLDA characteristic, 2DDLPP characteristic. For three kinds characteristics use nearest neighbor classifier, neural network classifier, the minimum distance classifier for a preliminary classification, and then use integrated classifier based on neural network model fusion the output of multiple classifiers, realize the final facial expression recognition.

Facial expression recognition based on Gabor feature

The improvement of Gabor feature extraction methods is simplify the selection problem of direction and scale to feature selection problem, using the incremental learning algorithm based on group (PBIL) choose scale and direction of Gabor nuclear, to get the optimal scale and the direction of the filter set, and on this basis extracting Gabor feature, reduce the feature dimension, reduce the redundant information in the feature.

Based on the PBIL Gabor nuclear scale and direction selection algorithm process is described below.

1. Set necessary parameters in PBIL algorithm, including population size M, individual vector length L, mutation rate Pm, learning rate α, probability variable learning rate α_{m}, the optimal number of individual μ. The individual consists of fixed length of binary strings; L is the full number of filter group.

2. Initialize the probability vector p(x). According to the probability vector generate the first generation M individuals. Set up training times t to 1

3. According to the fitness function evaluate each individual adaptive value.

4. Select μ individual of highest adaptive value and according to the following formula fixed probability vector p:

$$p_{l+1}(x) = (1-\alpha)\, p_l(x) + \alpha(1/\mu)\sum_{k=1}^{\mu} x_l^k \quad (1)$$

$p_l(x)$ as the first generation of probability vector; $x_l^1, x_l^2, ..., x_l^\mu$ for selection of μ individuals.

5. According to the following formula modified probability model

$$p_{l+1}(x) = (1-\alpha_m)\, p_l(x) + \alpha_m U(0,1) \quad (2)$$

U(0,1)represents a vector, each component is uniformly distributed random number between 0 and 1.

6. According to the new probability model generate the next generation of individuals.

7. If the termination condition is met, the algorithm terminates, output optimal solution; Otherwise, turn to step③, t = t + 1.

Based on $(2D)^2$DLDA algorithm of facial expression recognition

$(2D)^2$DLDA [7] algorithm from the horizontal and vertical directions of image matrix to perform direct linear discriminate analysis, compression the columns and rows of two dimension, greatly reduce the number of feature extracting. Horizontal processing is 2DDLDA algorithm, vertically processing is alternate 2DDLDA algorithm.

Hypothesis in training set there are C sample categories: $\omega_1, \omega_2, ..., \omega_c$, each type has N_i samples, all the total number of training samples are M, matrix $A_j^i (j = 1, 2, ..., N_i = 1, 2, ..., C)(m \times n$ image matrix) representative i class j training sample images, all the training sample average image is $\bar{A}$, i class average image is $\bar{A}(i = 1, 2, ..., C)$.

1. 2DDLDA algorithm. Discrete degree

matrix between training samples of class G_b, discrete degree matrix in class G_w.

$$G_b = \frac{1}{M}\sum_{i=1}^{c} N_i(\bar{A}_i - \bar{A})^T(\bar{A}_i - \bar{A}) \quad (3)$$

$$G_w = \frac{1}{M}\sum_{i=1}^{c}\sum_{j=1}^{N_i}(A_j^i - \bar{A}_i)^T(A_j^i - \bar{A}_i) \quad (4)$$

Trying to find a matrix diagonalization G_b and G_w, this matrix is a horizontal projection matrix X.

2. Alternate 2DDLDA algorithm. Discrete degree matrix between training samples of class $H_b$and discrete degree matrix in class H_w.

$$H_b = \frac{1}{M}\sum_{i=1}^{c} N_i(\bar{A}_i - \bar{A})(\bar{A}_i - \bar{A})^T \quad (5)$$

$$H_w = \frac{1}{M}\sum_{i=1}^{c}\sum_{j=1}^{N_i}(A_j^i - \bar{A}_i)^T(A_j^i - \bar{A}_i)^T \quad (6)$$

Similar to 2DDLDA, find a matrix diagonalization H_b and H_w this matrix is the vertical projection matrix Z. After X and Z is determined, given face image A, make $Y = Z^T AX$, can get the characteristics matrix Y of image A.

3. 2D- DLPP algorithm

2D-DLPP algorithm through increase a between the discrete 'degree of the constraint in 2D-LPP objective function, find a minimizing distance within the class, to maximize the distance between the classes, can distinguish different kinds of subspace, this algorithm emphasizes the discriminant information, is more suitable for identification. The objective function of 2D-DLPP algorithm is,

$$J(Y) = \frac{\sum_{c=1}^{c}\sum_{i,j=1}^{n_c}(Y_j^c - Y_j^c)^T(Y_i^c - Y_j^c)S_{ij}^c}{\sum_{i,j=1}^{c}(M_i - M_j)^T(M_i - M_j)W_{ij}} \quad (7)$$

$$M_i = \frac{1}{n}\sum_{k=1}^{n_i} Y_k^i, M_j = \frac{1}{n_j}\sum_{k=1}^{n_j} Y_k^j \quad (8)$$

$Y_i^c$and Y_j^c on behalf of the two projection image matrix in the class c, S^c as belonging to the class c weighting matrix between any two samples; W_{ij} for any two classes of weighting matrix between the mean value matrix, $M_i$and M_j representing the i classes and the j class average matrix of the projected image. A is $m \times n$ image matrix. I is a transformation matrix, linear transformation is Y = AI. Simplify (7),

$$J(I) = \frac{I^T A^T LAI}{I^T F^T HFI} = \frac{I^T P_s I}{I^T P_w I} \quad (9)$$

$$P_s = \frac{1}{2}\sum_{c=1}^{c}\sum_{i,j=1}^{n_c}(A_i^c - A_j^c)^T(A_i^c - A_j^c)S_{ij}^c,$$

$$P_w = \frac{1}{2}\sum_{c=1}^{c}\sum_{i,j=1}^{n_c}(F_i - F_j)^T(F_i - F_j)W_{ij},$$

$$F_i = \frac{1}{n}\sum_{k=1}^{n_i} A_i^k$$

Eigenvalue and eigenvector of solve the equation $A^T LAI = \lambda F^T HFI$. L=D-S, $D_{ii} = \sum_j S_{ji}$, F_i is the mean value matrix of i class. Projection matrix I=[$a_1, a_2, ...a_d$], constitute by d smallest nonzero eigenvalues feature vectors. For any image A_i, $A_i \to Y_i = A_i I.Y_i$ is the characteristic matrix after projection.

Facial expression feature extraction method based on geometric features, Wang Xiaoxia, Li Zhenlong, Xin Le, "Human facial expression recognition based on mixed features and hierarchical nearest neighbor method" this research according to the characters of facial expression based on classification tree categorizing expressions for three layers. The entire identification process from coarse to fine is a mixture of the geometric features and frequency domain features. The experimental results show that the method is effective.

3. The method based on characteristics face

3.1 The principle of characteristics face features extraction method

Characteristics Face method derived from principal component analysis is a facial recognition and description technique. On the basis of the traditional characteristics face method, the researchers noticed the bigger characteristic value of characteristic vector, not the best direction classification performance, and for K-L transformation, external factors of the image differences and face itself is unable to distinguish. The experimental results show that, characteristic face method with the introduction of the factors such as light, angle and size, recognition rate fell sharply. So there are defects in theory. In recent years, development a variety methods to improve characteristics face method, such as Gemini space method, by ambiguity analysis method, Fisher face method, etc. All in all, characteristic face method is a simple, rapid and practical, based on the characteristics of transform coefficient algorithm, but because it is in essence depends on the training set and test image gray correlation, and it requires a similar test image and the training set, so it has great limitations.

3.1.1 Facial expression recognition based on CMAC neural network

In research of facial expression recognition based on CMAC neural network, researchers first preprocessing the sample image, extracting interested facial region, through the K-L transform the processed image to extract eyes, mouth and nose, and other important feature point's location and local geometry as recognition feature expression of interest area. Finally use the distance of under test face and standard expression as a CMAC neural network's input, expression type as the network output, identification the 7 kinds of typical face expression.

(1) K-L is the optimal transformation in minimum mean square error of data compression. Based on K-L transform the expression of features eyes and features mouth inherited the vast majority of expression information. Size of N×N eyes sub image set $T_1, T_{2,} ... T_M$.

$$\mu = \frac{1}{M}\sum_{i=1}^{M} T_i . C = \frac{1}{M}\sum_{i=1}^{M}(T_i - \mu)(T_i - \mu)^T \quad (10)$$

Covariance matrix C, for the value of characteristic λ_i, and the corresponding eigenvectors μ_i, the feature vector μ_i is the characteristics eye, after transformation got M feature vectors. Using the K-L transform can select one of the largest feature vectors. Make (11) established μ_i composition the characteristics eye spatial. A similar approach can get space of characteristics mouth.

$$\sum_{i=l}^{l} \left|\lambda_2\right| \geq \alpha \sum_{i=1}^{M} \left|\lambda_2\right| \quad (11)$$

(2) Expression classification algorithm based on CMAC network

In this research the Euclidean distance are introduced to under test expressions, the Euclidean distance of actual expression and expression library of under test standard expression as network input, to simplify the network structure. Set under identify sample normalized

$$D_k = \left\| Y - Y_k \right\| = \sqrt{\sum_{i=1}^{m} \left[\left[a_i \right] - \left[a_{ki} \right] \right]^2} \quad (12)$$

$Y = (a(1), a(2), ..., a(m)), m$ for a sample number of pixels, the standard sample normalized $Y_k = \left[a_k(1), a_k(2), ..., a_k(m) \right]$, k is the number of standard sample, and it can work out of the Euclidean distance between under test sample and the *k*th standard sample, normalized

$$DD_k = \frac{D_k}{\sum D_k} \quad (13)$$

DD_k is CMAC network's input s_i, through the network learning to achieve the purpose of pattern recognition[8].

One study of automatic segmentation of facial expression recognition based on the characteristics area, in view of the present three-dimensional face expression region segmentation method is complicated and time-consuming problem, put forward a kind of method, automatic segmentation region of facial expression, detection by using the method of projection and curvature calculation part of the face feature points, on the basis of the feature points automatic segmentation facial expression area. Combined with the facial expression recognition encoding rules expand the extraction characteristic matrix. Use classifier for facial expression recognition. Through the recognition of three-dimensional facial expression database sample results show that the method can achieve higher recognition rate [9].

4. The feature extraction method based on wavelet analysis

Analysis compared with other methods of several typical transform, processing data quantity is less, can satisfy the system real-time requirements. And wavelet transform is not sensitive to illumination changes, can tolerate a certain degree of image rotation and deformation, improve the robustness of the system. L. Wiskott et al. [10] using Gabor wavelet transform coefficient as the feature points, for accurate positioning of facial feature points. Jing fu Ye et al. use two-dimensional Gabor wavelet function to defined Gabor wavelet extract the features in the gray image, in order to analyze the relationship between Gabor wavelet vector and different expressions, study the expression characteristics and light conditions, the relationship between factors such as personal characteristics, in different lighting conditions and different testing conditions, extraction Gabor expression character. Experiment shows that the Gabor wavelet transform can effectively extract the characteristics related to the expression changes, this feature is not sensitive to illumination changes, and to shielding the influence of personal characteristics differences, realization expression feature extraction.

In research of facial expression recognition based on wavelet transformation and neural network ensemble, proposes a new method of facial expression recognition. First of all, through the image decomposition based on wavelet transform and K-L transform extract effective identification features in facial expression areas, after using the neural network integration technology identification of six kinds of typical expression. In

the experiments on CMU face database show that the method to achieve a high recognition rate and no sensitivity to illumination change.

K-L is the optimal transformation in minimum mean square error of data compression. It can be spread on a set of variable information concentrated to certain composite indicator, in fact play the role of a data dimension reduction, and to ensure that in the process of dimension reduction, maximum keep the difference of the raw data. Study shows that the most abundant areas of expression are eyes and mouth, so K-L transform features of eyes and mouth are inherited the most part of expression information. The algorithm of characteristics eye can be expressed as follows: Size of $N_1 \times N_2$ eye expression image sets $T_1, T_2, T_3, ..., T_M$. The average of sub image set is

$$\mu = \frac{1}{M}\sum_{i=1}^{M} T_i \text{ , } \{T_i\} \quad (14)$$

$\{T_i\}$ covariance matrix

$$C = \frac{1}{M}\sum_{i=1}^{M} (T_i - \mu)(T_i - \mu)^T \quad (15)$$

The covariance matrix eigenvalues λ_i and the corresponding eigenvectors u_i, the feature vectors u_i is characteristics eye. In practical application, u_i makes the formula (16) establish, u_i constitute a characteristics eye space, similar approach can get space of characteristics mouth.

$$\sum_{i=1}^{1} |\lambda_1| \geq 0.95 \sum_{i=1}^{M} |\lambda_1| \quad (16)$$

Use the low frequency information after wavelet transform as the research object, due to the low frequency information compare with the high frequency information can be more comprehensive description of the object, including shape, outline, etc. The expression changes mainly embodied in these aspects, therefore to extract low frequency information not only reduced the amount of calculation, drive out the influence of noise, and hold the main characteristics of the facial expression change, is a kind of effective method[11].

5. Based on a variety of feature fusion method

A variety of methods adapt to the different environment and have their advantages and disadvantages, only use a feature extraction method is hard to get good robustness and high precision. So the researchers put forward the theory of data fusion, comprehensive the different method, foster strengths and circumvent weaknesses, in order to achieve good results.

In research of "combination method in the application of facial expression recognition" the researchers believe that, the study of facial expression recognition problem should be effective to extract the face expression characteristics, eliminate irrelevant information. The traditional Gabor filter in facial expression feature extraction process, aimed at the shortcomings of extract features time consuming and data redundancy, put forward the Gabor and PCA feature extraction of combining, and through the support vector machine (SVM) for facial expression recognition method. This method first preprocessing expression of face image to get pure expression, using Gabor extracting expression feature, deal with redundant data with PCA, using support vector machine (SVM) to identify facial expressions and simulation. The simulation results show that, compared with the traditional method of Gabor, this method not only improve the accuracy of facial expression recognition, and accelerate the speed of recognition.

5.1. Gabor feature extraction

When making facial expression feature extraction, generally composed of different scales and directions Gabor filter. Large scale Gabor filter describes the stronger information, and it can remove the influence noise of image, small scale Gabor filter can describe the fine feature of image, but the influence of noise is big. So when making facial expression feature extraction, should try to select small scale Gabor filter.

The study used by the five dimensions(v=0,1,…,4) and eight directions ($\mu = 0,1,...,7$) Gabor filter group to extract the image different frequency scale and the direction of the texture information, constitute feature vector. Through a set of Gabor wavelet filter transformation, make the input image $I(x, y)$ convolution with every Gabor wavelet filter. The convolution of the Gabor filter is defined as follows:

$$G_{v\mu} = \psi_{v\mu}(x, y) \times I(x, y) = \int \psi \int (x, y) \times I(x, y) dxdy \quad (17)$$

(x, y) as facial expression image pixel coordinates, I(x, y) as the corresponding value of pixels. Therefore, for any point in the image after two-dimensional Gabor filter, you can get a forty-dimensional feature vector.

5.2 Feature dimension reduction

After two-dimensional Gabor extract feature vector to the facial expression of the image, the dimension has been significantly reduced, but the feature dimension is still quite high, if identification directly, identification process not only time-consuming and accuracy is low, so must dimension, research using principal component

analysis (PCA) to reduce the dimensions of the feature vector.

Suppose N individual sample image face expression, after two-dimensional Gabor extraction get m dimensional feature vector, $x_i = (i = 1, 2, ...N)$, through PCA method map original m dimensional feature vector to f dimensional subspace, and $f \le m$ .The specific steps as follows;

1. Data normalization. In the first place, each study sample x_i minus the average, and then divided by the variance to implementation normalization. Make x_i normalized between 0 and 1, get x_i', $x_i' = \frac{x_i - \overline{x_i}}{\sigma}$, $\overline{x_i}$ is average, σ is variance.

2. After normalization of the samples to composition a new matrix, $X = \left[x_1', x_2', ..., x_N' \right]$

3. So the covariance matrix is $Q = XX^T$, according to the characteristic equation for the characteristics vector of the covariance matrix $\lambda E = QE$, λ is sample characteristic value, E is sample feature vector.

4. Decided principal components. The samples ordered by the size of the characteristic value. According to the cumulative variance contribution ratio of the principal component, select the cumulative contribution rate of the principal component $\sum_{i=1}^{f} \alpha_i \ge 85\%$ the smallest integer f, to determine the front f principal component.

5. New sample feature vector, $y_i = W_{pca}^T x_i (i = 1, 2, ..., N)$, W_{pca} is linear transformation matrix, $W^T{}_{pca}$ representative transpose, after dealing with the PCA, the characteristic dimension of the original have fallen sharply, while maintaining the most important identifying information, improve the speed of recognition[12].

According to the problems existing in the human face feature extraction method and the future research trend, made a summary and outlook as follows.

6. Conclusions

The face of the movement including the rigid motion of the head and the flexible movement of facial, if extract the facial movement data without separation of the two kinds of movement, easy to cause the failure of expression feature extraction, in order to solve this problem, the three-dimensional expression information should be introduced.

Facial expression information is very complicated, feature extraction methods suitable for different environment, and each method has advantages and disadvantages, simple to use a feature extraction method to extract facial expression information it is not realistic. Therefore using the theory of data fusion, comprehensive use of various methods, foster strengths and circumvent weaknesses, in order to achieve good extraction effect, is the inevitable trend in the future research. At present facial feature extraction methods are not yet provided a good solution the problem of light, extraction low-level visual features which were not sensitive to illumination change; (horizontal grain characteristics, wavelet transform coefficient, etc.) would be worth studying one of the solutions in the future.

Considering the robustness and practicality, the resolution of the images in real application, system processing speed, all is need to solve the problems of the feature extraction methods study in the future.

Acknowledgements

This work was supported by

1. J2013-19(The formation mechanism and intervention strategy of airport mass disturbances research based on social psychology. Civil Aviation Flight University of China, surface project) .

2. 2014GXS2B009 (Public safety aviation passenger behavior recognition research. National soft science project).

3. MHRDZ201102 (Based on the behavior pattern recognition passengers screen detection theory and application research. The civil aviation administration of China soft science project).

References

1. L Ma, K Khorasani (2004) Facial expression recognition using constructive feed-forward neural networks. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 34(3), p.p. 1588-1595
2. S Mitra, Y Liu(2004) Local facial asymmetry for expression classification. *Proc. Conf. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Washington DC, USA*, p.p. 889-894
3. G Donato, et a1.(1999) Classifying facial actions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(10), p.p. 974-989
4. X Chen, T Huang (2003) Facial expression recognition: a clustering based approach. *Pattern Recognition Letters*, 24(9-10), p.p. 295-302
5. S Dubuisson, F Devoine, MMasson (2002)

A solution for facial expression representation and recognition. *Signal Processing: Image Communication*, 17(9), p.p. 657-673
6. L Zalewsk, S Gong(2004) Synthesis and recognition of facial. Expressions in virtual 3D views. *Proceedings of IEEE 6th International Conference on Automatic Face and Gesture Recognition*, Seoul Korea, p.p.493-498.
7. Zheng Qiumei, L Xing-hui, Shi Gong-xi.(2010) Two directional two dimensional direct linear discriminate analysis for facial expression recognition. *Journal of China University of Petroleum,* 34(5), p.p.179-182
8. Ye Fangfang, Xu Li (2010) Facial Expression Recognition Based on CMAC Neural Network. *The computer simulation*, No.8, p.p.262-265
9. Zhang Tengfei, Min Rui, Wang Baoyun (2011) Facial Expression Recognition Based on Feature Regions Automatic Segmentation. *Computer Engineering*, No.5, p.p.146-148
10. Wiskott Laurenz (1997) Face Recognition by Elastic Bunch Graph Matching. *IEEE transactions on PAMI*, 19(7), p.p.775-779
11. XU Zheng-guang, Shen Xiao-dong, Zhang Li-xin (2006) Facial Expression Recognition Based on Wavelet Transformation and Neural Network Ensemble. *Microelectronics and computer*, No.7, p.p.144-150
12. Hong Weiming (2011) Study on Facial Expression Recognition Based on Combination Method. *The computer simulation*, No.2, p.p.337-340