

Facial expression recognition research based on blocked local feature

Zhang Erdong¹, Xu Shuo^{1,2}, Zhang Peng¹

(1. School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200072;

2. Shanghai Key Laboratory of Intelligent Manufacturing and Robotics, Shanghai 200072)

Keywords: Gabor features, block Procrustes analysis, PCA, fused features

Abstract. Focus on the key feature extraction and selection of facial expression recognition, this paper firstly extracts the global Gabor features and facial key expression as local features. Then centering, rotating and scaling every feature block with Procrustes analysis to reduce the effect of position and size inconsistent. At last, reduce the dimension of global Gabor features and local features with PCA algorithm and combine them to fused features. Experimental results show that whether the single features or fused features, blockProcrustes can obviously improve the expression recognition performance in most cases, especially increase the stability and maximum of recognizing accuracy.

1 Introduction

As an important part of emotional computing, facial expression recognition has surpassed human face recognition and become the focus of computer vision research. In the article ^[1], the authors found that facial expression contributed 55% to the emotional expression, while the language and sound were 7% and 38%. It means that facial expression is the key to emotional expression.

At present, the research on facial expression recognition mainly uses the pattern recognition and machine learning methods to classify the extracted expression images features. The expression images features can be divided into two types, global features (affected by light, color, race) and local features which are consisting of some facial key points coordinates such as the contour of the mouth, eyes, etc. The global features are based on the whole face, focusing on the whole information, while the local features could reflect the local changes better. Due to the fully respective advantages of two methods, some researchers considered to fuse them. Li Jianjun, Zhao Baojun^[2] proposed a method by fusing discrete cosine transform(DCT) and local binary pattern(LBP) feature. They extracted the DCT and LBP feature firstly, then normalized and integrated them. The experimental results showed that the fused feature is superior to the single feature. Luo Yuan and Cui Ye ^[3] fused the global low-frequency DCT features and the LBP features of the eyes and mouth area, the experiment showed better results than the single local features or global features. Wang Yi et al. ^[4] proposed combining DCT and Gabor to extract facial expression features. Experimental results showed that the performance of fusion feature is much better than the single DCT or Gabor feature.

For the local feature of fusion feature, facial expression recognition focus on the shape transformation of facial expression feature, but the position and size of face area in the image had some deviation, so some scholars use Procrustes Analysis [5] to process the extracted local feature for the aim-that eliminates the effect of position and size difference. Po Linzhi and Zhao Jieyu et al. ^[6] process the local feature extracted by ASM with the Procrustes Analysis. Laszlo A. Jeni and Daniel Takacs ^[7] pointed out that facial expression recognition with robustness to different people and postures is the key of scenario analysis and automatic video annotation. Therefore, they used Procrustes analysis to process the extracted feature points of facial expressions, and the SVM classification experiment finally proved that this method is robust to different postures. Sunil Kumar, MK Bhuyan and Senior Member ^[8] pointed out that any expression is the result of deviation degree compared to neutral expression according to Ekman theory, and the analysis for that) is very important. But, the neutral expression between different people exists major variation in size and position. For this problem , they proposed using Procrustes analysis to align, center and rotate the

feature. The Procrustes analysis was utilized to reduce the effect of different size and position and experimental results showed that it can improve the recognition rate. But, as we know, different people's eyes and mouth is also different in size and position, and eye and mouth feature block have a key contribution to facial expression recognition. According to Ekman's theory, facial expression is the results of various facial features block's superposition^[9]. As we all know, the size, position of eye and mouth is also various depend on different people. Therefore, we divide the facial features into three parts: mouth, left eye and right eye feature block. Every feature block of the facial features is processed separately by Procrustes analysis and then fused every block into fused feature. By this method, we can reduce the effect of position and shape deviation in each feature block and the interaction among feature block. This process was called block Procrustes analysis.

The dimension of global Gabor feature is up to tens of thousands which include plenty of noise information and need a lot of computational time. Therefore, the PCA algorithm is used to reduce the dimension and noise of the global feature, and then combined it with the local feature. Based on above research points, we design the experimental framework shown in Fig.1. Firstly, we extract the global Gabor feature of the image and locate the key facial points coordinate by method^[10] as the local feature, then blockProcrustes analysis and fuse the global and local features after PCA analysis in different PCA feature retention rate. We research the performance of blockProcrustes analysis for the fused feature and single feature in different local or global feature retention rate.

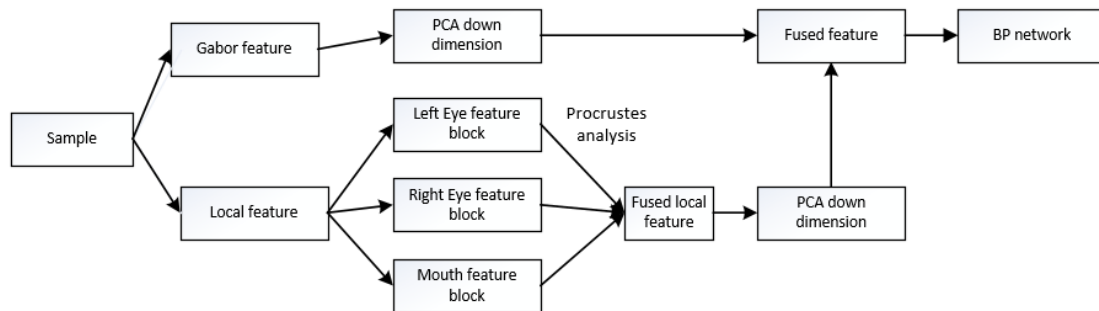


Fig.1 experimental framework

2 Feature extraction and selection

2.1 Gabor features extraction

Gabor wavelet can extract multi-scale and multi-directional texture features of facial expression, which is less affected by illumination change. Researchers^[11; 12] had paid more and more attention on Gabor filter which is designed according to different frequency and bandwidth. The texture information of 8 scales and 5 directions^[13; 14] can extract entire expression information. The extracted Gabor global features are shown in Figure 2 below.

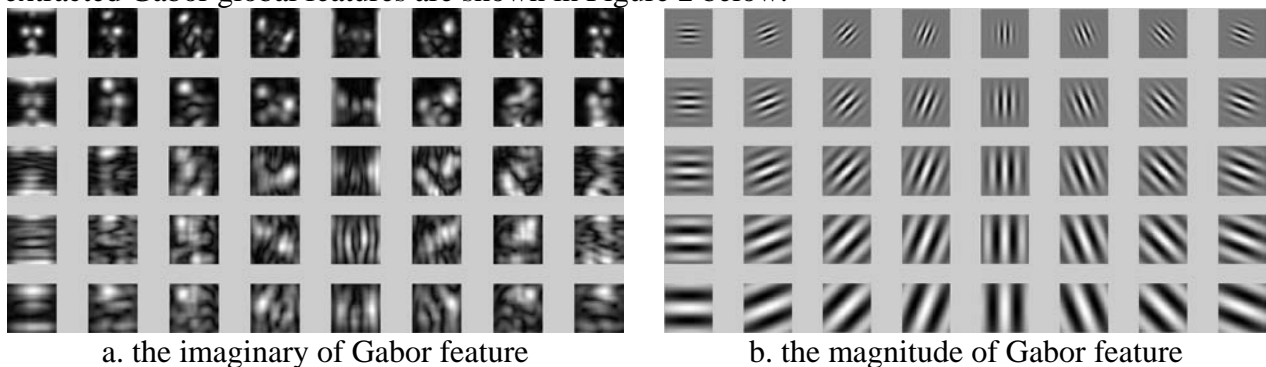


Fig2.Gabor feature

The feature dimension is $64 * 64 * 5 * 8$, down-sample the feature to 2560.

2.2 Local feature and Procrustes analysis

2.2.1 Local feature extraction

In this paper, the method proposed in paper [10] is used to extract facial feature points. This method is a four-layer convolution network cascade, which tackles the problem in a coarse-to-fine manner, each network level is trained to locally refine a subset of facial landmarks generated in previous network levels. This method can accurately extract 83 feature points including facial contour, mouth, left and right eyes, nose and eyebrows. The characteristic points example is shown in Fig.3. In facial expression recognition, mouth and eye feature blocks contribute more than nose and facial contour. Hence we combine mouth, left with right eye into a fused feature including 108 feature points.

The extracted local feature points are fused as follow formula:

$$F_l = [F_{leftEye}; F_{rightEye}; F_{mouth}] = [x_1, y_2, \dots, x_{35}, y_{36}; x_{37}, y_{38}, \dots, x_{71}, y_{72}; x_{73}, y_{74}, \dots, x_{107}, y_{108}]$$

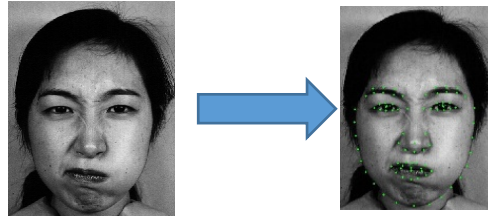


Fig 3. facial key point schematic diagram

2.3 block Procrustes analysis

The step of block Procrustes analysis^[15] :

1. Aligning the shape to the origin: For each feature block, subtract its mean of feature blocks. In this step, we can reduce the effect of difference in position.
2. Scale normalization: Each feature block is divided by its corresponding L2 norm. In this step, we can reduce the effect of difference in size.
3. Eliminating the rotation difference: Calculate the average of all the normalized samples in each feature block as the standard shape, and use the least squares method to calculate the rotation angle between each feature block and the standard shape. Each feature block is rotated using the rotation angle.

Each feature block in the local feature is regarded as an independent to Eliminating the effects of size and position deviation and fusing these feature block to get a new local feature.)

Because the dimension of global Gabor features, local feature is too high and exist some noise information, PCA^[16] algorithm is chosen to reduce the dimension and select the representative information.

3 Experiment and analysis

We choose Jaffe expression library^[12] as our experiment samples, the number of test sample and training samples are shown as follows Table 1:

Table 1. the amount of sample

expression	Anger	Disgust	Fear	Happy	Negative	Sad	Surprise	all
train set	20	19	21	21	20	21	20	142
test set	10	10	11	10	10	10	10	71
all	30	29	32	31	30	31	30	213

In accordance with the experimental framework of Figure 1, the following two experiments was executed:

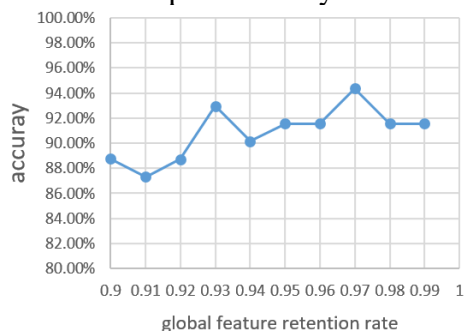
1. The global and local features were tested separately, and the local features were processed by Procrustes and block Procrustes analysis.
2. Local features are processed by Procrustes and block Procrustes analysis. Then we combined global features and local features into fused feature. In this experiment, we compare the

performance between fused feature, Procrustes analysis, block Procrustes analysis.

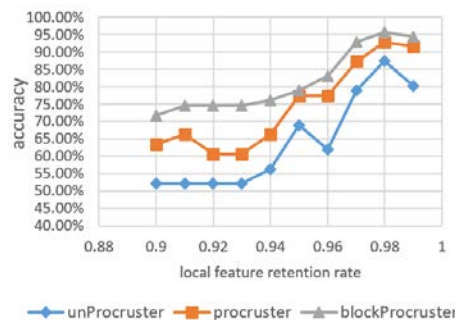
In above experiment, the PCA feature retention rate was between 0.9~0.99, the experiment was carried out under different global and local feature PCA retention rates. The features were classified by Bp neural network^[18].

3.1 Single feature experiment

The experimental results about global feature are shown in Fig.4 (a) and experimental results about local feature processed by Procrustes, block Procrustes analysis are shown in Fig.4 (b).



a. the global feature experiment under different retention rate



b. the local feature experiment under different retention rate

Fig4. experiment result about global or local feature as classification characteristics individual

The dimension of global feature is reduced by PCA and it can retain 90% of the features in 35 dimensions. When the dimension closes to 100, the retention rate is close to 100%. From about Fig.4 (a), we can see that expression recognition rate didn't reduce significantly due to the reduction of the dimension, recognition accuracy is up to 94.37% when the feature retention rate is 0.97, and the feature dimension is 67.

As we can see from Fig.4 (b) that with the increase of PCA feature retention rate, the recognition accuracy rate tends to increase in general. When the retention rate is low, the dimension down to about 5, and the recognition accuracy get relatively low, but reached the highest value at 0.98, then decreased. Compared (with) unProcrustes (local feature without Procrustes analysis process), the local feature's recognition rate of Procrustes analysis has a remarkable improvement under any feature retention rates. This improvement gets better with block Procrustes analysis process than the two cases mentioned above) except in the lowest and highest feature retention rate. We can conclude that block Procrustes analysis have a better robustness ability than Procrustes analysis to the effect of different size and position.

3.2 Fused feature experiment

The dimension of global and local features had been reduced in different PCA feature retention rates, and then combine the local and local features to fused feature. We get three hundred results due to the three different local feature process method.

In order to compare the experimental results, we statistics the results in follow two manners: 1. Under different global feature retention rate, calculate the mean, maximum, and minimum values of ten group local feature points under different global feature retention rates. 2. Under different global feature retention rate, calculate the mean, maximum, and minimum values of the ten (group) global feature points under different local feature retention rates. The statistics diagram is shown in the Fig5 and Fig6.

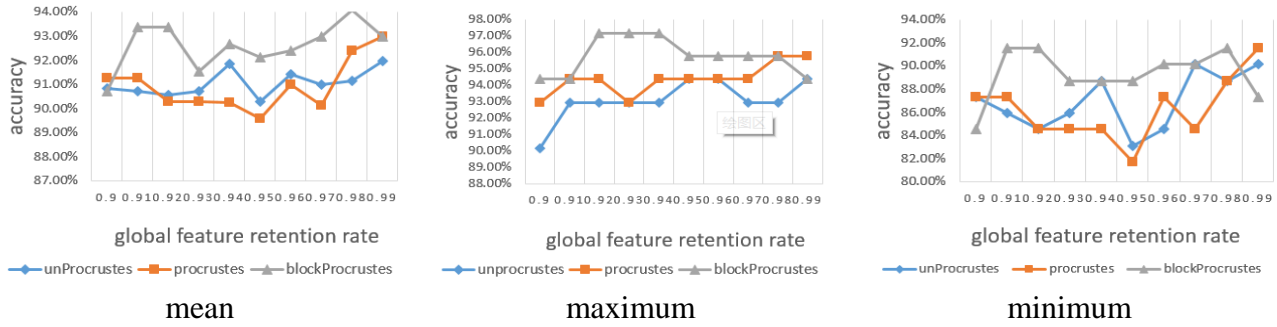


Fig5. The mean, maximum, and minimum values of the ten local feature points under different global feature retention rates

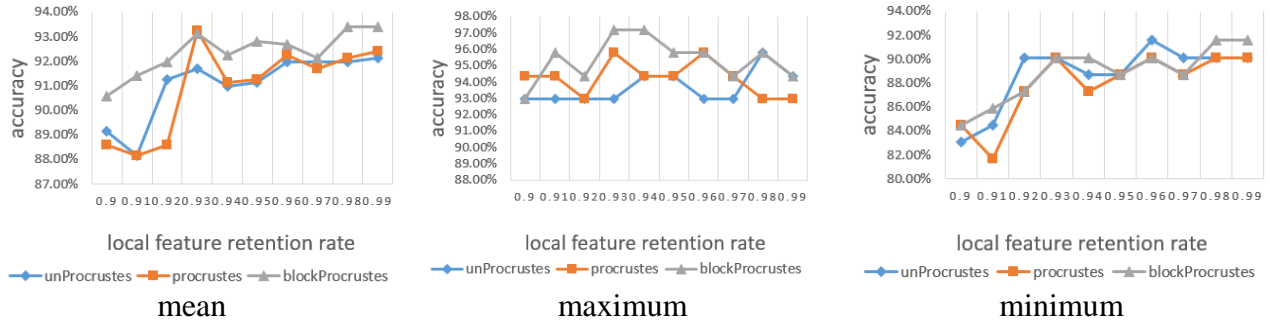


Fig 6. The mean, maximum, and minimum values of the ten global feature points under different local feature retention rates

As we can see from Fig5 and Fig6, block Procrustes performs much better than unProcrustes and procrustes in terms of average comparison. Basically, all accuracy are higher than the other two cases. Procrustes can also enhance accuracy but the performance is not outstanding. In the maximum comparison, block Procrustes accuracy are basically higher than the other two cases except for two points. In the minimum comparison, we can see from Fig5 that the minimum value of block Procrustes are higher than the other two cases except for the two edge points. In Fig6, block Procrustes does not perform better than the other two cases, but the gap between the three cases is minor.

The 300 groups experiments results are statistically. We analysis the number of accuracy increase(in), equal(eq) and reduce(re) among the three cases one to one.

Table2. The increase, equal and reduce number among the three cases

algorithm	unProcrustes			procrustes			blockProcrustes		
		in	eq	re		in	eq	re	
unProcrustes		39	16	45		59	19	22	
procrustes						61	16	23	
blockProcrustes									

As we can be seen from above table, Procrustes and block Procrustes can improve recognition rates at certain feature retention rates compared with unProcrustes, but-The increased number of blockProcrustes is 20 more than Procrustes and the reduced number down to 22 from 45. The increased number up to 61 when to compare blockProcrustes with Procrustes. Besides, the block Procrustes' best accuracy 97.18% is better than unProcrustes' 94.37% and procrustes' 95.77%. Overall, block Procrustes have better performance than the other two cases.

4 Conclusion

Focus on increasing the robustness to difference size and location of different person's expression samples. we first extract facial key point's coordinate as local feature and divide the local feature to left eye, right eye and mouth feature block. Then each feature block of the facial local features (left eye, right eye and mouth) are processed separately by Procrustes analysis and then fused every block into a new local feature. Combine the local features with Gabor global feature at different PCA feature retention rates and use Bp network to test the performance.

Experiment result shows that blockProcrustes had better performance than unProcrustes and Procrustes no matter in stability and highest accuracy.

Acknowledgement

In this paper, the research was sponsored by the Nature Science Foundation of china (Project No. 201112400450401).

Reference

- [1] A M. Communication without words[J]. Psychology Today., 1987, (2): 53-56.
- [2] J.K. Li, B.J. Zhao. Fusing DCT and LBP features for face recognition[J]. Transaction of Beijing Institute of Technology, 2010, 30(11).
- [3] Y. Luo, Y. Cui, Facial expression recognition based on fusion feature extraction of DCT and LBP[J]. Semiconductor Optoelectronics, 2014, 35(2).
- [4] Y. Wang. Facial expression recognition based on Gabor and two times DCT [J]. Microelectronics & Computer, 2009, 26(5).
- [5] Ross A. Procrustes Analysis[J]. Department of Computer Science and Engineering University of South Carolina , SC 29208.
- [6] L.Z. Piao, J.Y. Zhao. Facial expression recognition using statistical shape analysis[J]. Journal of Ningbo University (NSEE), 2009, 22(2):196-201.
- [7] Lorincz L a J D T A. High quality facial expression recognition in video streams using shape related information only[J]. Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on, 2011: 2168-2174.
- [8] Kumar , M.K.Bhuyan S. Neutral expression modeling in feature domain for facial expression recognition[J]. 2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS), 2015: 224-228.
- [9] Ekman P F W. Facial action coding system (FACS)[J]. A Technique for the Measurement of Facial Action Consulting, 1978.
- [10] Erjin Zhou H F. Extensive Facial Landmark Localization with coarse to fine convolutional network cascade[J]. IEEE International Conference on Computer Vision Workshops, 2013.
- [11] K. Crammer Y S. On the learnability and design of output codes for multiclass problems[J]. In Computational Learning Theory, 2000: 35-46.
- [12] Michael Lyons S A. Coding Facial Expressions With Gabor Wavelets[J]. IEEE Computer Society, 1998: 200-205.
- [13] T L. Image representation using 2-D Gabor wavelets[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 1996, 18(10): 959-971.
- [14] Z.H. Liu, J. Yin, Z. Jin. An adaptive feature and weight selection method based on Gabor image for face recognition[J]. Acta Photonica Sinica, 2011, 40(4) : 637-638.
- [15] Jain K B B F K a K. Component-Based Representation in Automated Face Recognition[J]. IEEE Transactions on Information Forensics and Security, 2013, 8(1): 239-253.
- [16] K. G. The biplot graphic display of matrices with application to principal component analysis[J]. Biometrika, 1971, 58(3): 453-467.
- [17] X.L. Deng, C. Kong. Detection of citrus HuangLongBing based on principal component analysis and back propagation neural network[J]. Acta Photonica Sinica, 2014, 43(4).
- [18] Wang F C Z. Research on a Method of Facial Expression Recognition [J]. The Ninth International Conference on Electronic Measurement & Instruments 2009: 225-228.