Facial Expression Recognition Based on Fusion of Multiple Gabor Features

WeiFeng Liu, ZengFu Wang
Department of Automation, University of Science and Technology of China
Hefei, Anhui, 230027, P.R.China
zfwang@ustc.edu.cn

Abstract

In order to accomplish subject-independent facial expression recognition task, a multiple Gabor features based facial expression recognition method is presented in this paper. Different channels of Gabor filters have different contributions on the facial expression recognition and reasonable combination of these features can improve the performance of a facial expression recognition system. NN based data fusion method is designed for facial expression recognition in this paper. Experimental results show that the facial expression recognition rate can be improved by using multiple channel features and neural network fusion.

1. Introduction

Facial expression recognition has many applications in areas such as image understanding, psychological studies, and smarter human-computer interfaces. In recent years, a number of techniques have been presented in the literature of face expression recognition. According to the properties of images used, the methods can be divided into two classes: video frames based method [5],[6],[7],[11],[12] and still image based method [1],[2],[3],[4]. The former deals with a sequence of intensity image frames and extracts their dynamics for expression recognition. The successful methods include Optical Flow models [5] and Hidden Markov Models (HMM) [11],[12]. The latter analyzes single still image and gives a recognition result based on spatial analysis of features extracted. This class of methods can be grouped into two subclasses: holistic spatial analysis, such as Principle Components Analysis (PCA), and Fisher linear discriminates (FLD), and local spatial analysis, such as Gabor wavelet and local PCA [14].

In [14] Donato etc. pointed out that a good performance could be obtained with Gabor wavelet decomposition. It employs gray-level texture filters with properties of spatial locality. These filters can mimic the response properties of visual cortical neurons.

This paper investigates the Gabor features for expression recognition. From experiments, we find that different channels of Gabor filters have different contributions on the facial expression recognition and reasonable combination of the features can improve the performance of a facial expression recognition system. Based on the observations, we use multiple Gabor features based reliability weight to improve the performance of system. Given a still image with landmarks selected, we use Gabor wavelets of these landmarks obtained from the outputs of different channels of Gabor filters as features. Then the Gabor PCA method is employed to obtain evaluations of different facial expressions by using Gabor wavelets of different channels. Meanwhile a reliability table, which represents the contributions of the different channel features to different facial expressions, is obtained by training samples. Finally, a reliability weighted data fusion method is used to give the result of facial expression recognition. Experimental results strongly support our observations and show the validity of our new method.

2. Gabor feature extraction

2.1. Gabor wavelets

Gabor kernels are similar to the receptive field profiles in cortical simple cells, which are characterized as localized, orientation selective, and frequency selective. A family of Gabor kernel is the product of a Gaussian envelope and a plane wave, defined as [4],[8],[9],[10]:

$$\Psi_{\mu,\nu}(z) = \frac{\| k_{\mu,\nu} \|^2}{\sigma^2} e^{-\|k_{\mu,\nu}\|^2/\sigma^2} [e^{ik_{\mu,\nu}z} - e^{-\sigma^2/2}] \quad (1)$$

Here,$z = (x, y)$ is the variable in spatial domain and $k_{\mu,\nu}$ is the frequency vector, which determines the scales and the orientations of Gabor kernels. In our system, $k_{\mu,\nu}$ is defined as follows,

$$k_{\mu,\nu} = \frac{k_{max}}{f^\nu} e^{i\phi_\mu} \quad (2)$$

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

where, $k_{max} = \frac{\pi}{2}, f = \sqrt{2}$ and $\phi_\mu = \frac{\mu\pi}{8}$, while μ and ν are orientation factor and scale factor respectively. Different selection of subscript μ and ν gives different Gabor kernel. In our system, we choose $\mu = 0, 1, 2, ..., 7$ and $\nu = 0, 1, 2, ..., 4$, thus totally we have 40 Gabor functions to be used. The real part of the Gabor kernels used in our experiments is shown in Figure 1.

Given an image $I(z)$, its Gabor transformation at a particular position can be computed by a convolution with the Gabor kernels:

$$G_{\mu,\nu} = I(z) * \Psi_{\mu,\nu}(z) \tag{3}$$

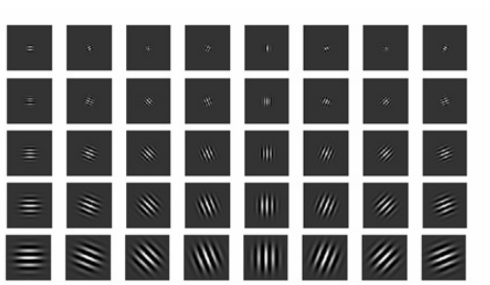
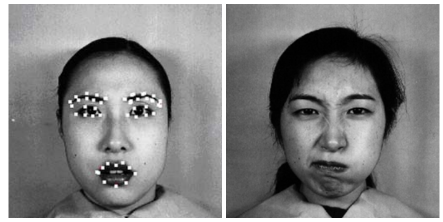
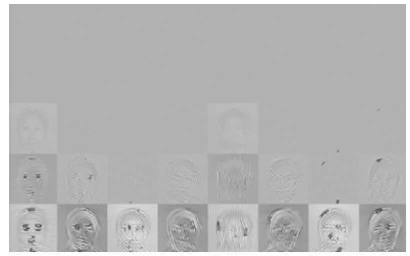


Figure 1. Real part of the 5×8 Gabor filters.



(a) Fiducial points of face graph model.

(b) A facial image.



(c) The responses of the Gabor filters to the facial image of (b).

Figure 2. Fiducial points of face graph model and the responses of the Gabor filters to the facial image.

2.2. Gabor features extraction

The experiments show that mouth contributes the most to facial expression, while canthus and eyebrows follow it. Hereby we designed a face graph model with 42 nodes for facial expression analysis. As shown in Figure 2(a), the 42 nodes distributes in peripheries of the mouth, canthus and eyebrows as critical fiducial points. For each fiducial point, totally 40 Gabor features can be obtained when 40 Gabor filters are used. Here the real part of outputs of the ith fiducial point forms a vector f_{pi} . We use it as local feature. In such way, the face image can be represented by a Gabor feature Group $G^T = [f_{p1}^T, f_{p2}^T, ..., f_{p42}^T]$. An example facial image and its response of Gabor filters are shown in Figure 2(b) and 2(c).

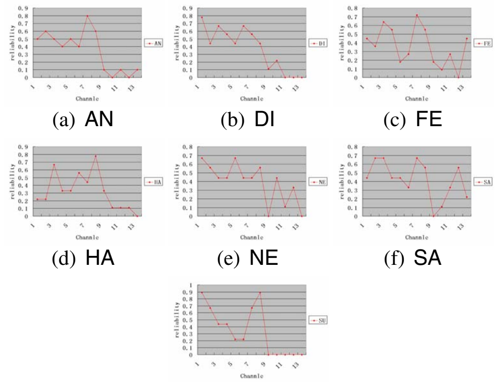
P. Ekman [13] divided human emotions into 6 classes (anger, disgust, fear, happiness, sadness, surprise). In research field of facial expression recognition, 7 classes including neutral one are often used. In order to verify the validity of the above Gabor features selected, we have made an experiment to analyze the recognition rates of Gabor features with different scale or orientation to different facial expressions. In our experiment, we divided all of the Gabor filters into several groups. A group of Gabor filters with the same scale or orientation is called a channel. All of the Gabor filters are then divided into 13 channels (corresponding to 5 scales and 8 orientations respectively). Figure 3 shows the results, where x-axis denotes the channel, and y-axis denotes the recognition rate.

From Figure 3 we can see that the Gabor features of different channels give different recognition rates to different facial expressions. In addition, we can also see that none of the channels can give satisfying recognition rate in all cases. The Gabor features of some channels give good recognition rate to some facial expressions, while the Gabor features of other channels give acceptable recognition rate to other facial expressions. Therefore, reasonable combination of the features belonging to different channels should improve the total recognition rate of a system.

3. Expression recognition based on data fusion

As mentioned before, different Gabor channel-features have different contributions to different facial expressions recognition. In order to perform the facial expression recognition task, we have developed an expression recognition system based on Fusion of Multiple Gabor Features. Firstly a reliability table, which represents the contributions of the different channel features to different facial expressions, is obtained in the training stage. Then for a given test image, PCA based method is adopted to recognize the facial expression on each of the 13 Gabor channel-feature vectors. Finally, a neural network based data fusion method is de-

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE



(a) AN (b) DI (c) FE

(d) HA (e) NE (f) SA

(g) SU

Figure 3. The recognition rate of the Gabor features belonging to different channels to facial expressions of anger, disgust, fear, happiness, neutral, sadness and surprise.

signed to give the decision.

The JAFFE database [10] is used in the study. It includes 3 or 4 examples for each of the six basic facial expressions and a neutral face image for each person, a total of 219 images of 10 persons. Two images of each expression for all of the people are used as training examples. The reliability table is shown in Table 1, which represents the contributions of the different channel features to different facial expressions (denote anger, disgust, fear, happiness, neutral, sadness and surprise expression as AN, DI, FE, HA, NE, SA, and SU, respectively).

For simplicity of description, denote $R(c, e)$ be the reliability of correctly recognizing facial expression e by using the Gabor features of channel c. For example, $R(2, HA$ means the reliability of correctly recognizing facial expression HA by using the Gabor features of channel 2. From Table 1, $R(2, HA) = 0.22$.

Table 1. Expression-Channel relation table

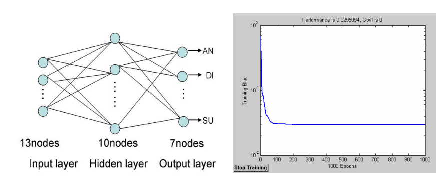
CH	AN	DI	FE	HA	NE	SA	SU
1	0.50	0.78	0.45	0.22	0.67	0.44	0.89
2	0.60	0.44	0.36	0.22	0.56	0.67	0.67
3	0.50	0.67	0.64	0.67	0.44	0.67	0.44
4	0.40	0.56	0.55	0.33	0.44	0.44	0.44
5	0.50	0.44	0.18	0.33	0.67	0.44	0.22
6	0.40	0.67	0.27	0.56	0.44	0.33	0.22
7	0.80	0.56	0.72	0.44	0.44	0.67	0.67
8	0.60	0.44	0.55	0.78	0.56	0.56	0.89
9	0.10	0.11	0.18	0.33	0.00	0.44	0.00
10	0.00	0.22	0.09	0.11	0.44	0.11	0.00
11	0.10	0.00	0.27	0.11	0.11	0.33	0.00
12	0.00	0.00	0.00	0.11	0.33	0.56	0.00
13	0.10	0.00	0.45	0.00	0.00	0.22	0.00

In the test phase, a PCA based method is employed to obtain the estimations of facial expressions of the test facial image on different Gabor channel features. Denote $r(c, e)$ be the estimation of correctly recognizing facial expression e by using the Gabor features of channel c. For each test image, an estimation table can be obtained. Table 2 gives the corresponding result of $r(c, e)$ for the facial image KA_DI3 (taken from JAFFE database).

Table 2. Expression-Channel estimate table

CH	AN	DI	FE	HA	NE	SA	SU
1	0.05	0.34	0.05	0.09	0.08	0.20	0.19
2	0.04	0.18	0.24	0.14	0.08	0.15	0.17
3	0.10	0.35	0.13	0.06	0.10	0.17	0.09
4	0.03	0.33	0.12	0.09	0.04	0.06	0.31
5	0.07	0.26	0.14	0.11	0.07	0.08	0.27
6	0.20	0.19	0.07	0.16	0.09	0.13	0.06
7	0.04	0.36	0.12	0.14	0.11	0.16	0.07
8	0.06	0.24	0.13	0.19	0.13	0.16	0.09
9	0.08	0.21	0.28	0.13	0.06	0.17	0.07
10	0.01	0.30	0.09	0.21	0.24	0.11	0.05
11	0.10	0.19	0.16	0.11	0.09	0.13	0.12
12	0.30	0.11	0.17	0.09	0.13	0.06	0.17
13	0.30	0.07	0.15	0.12	0.09	0.16	0.11

The final decision can be obtained by a data fusion procedure. In this work, we construct and use a neural network to perform the task. The neural network has a general back propagation structure with three layers. The input layer has 13 nodes for the estimations from 13 Gabor channel-features, followed by a hidden layer of 10 TANSIG neurons. The output layer has 7 nodes which give the recognition result. In implementation, we used TRAINLM (Levenberg-Marquardt) for back propagation network training function, LEARNGDM (Grads Descend on Momentum) for back propagation weight/bias learning function, and MSE for performance function. The structure and the training performance of the neural network are illustrated in Figure 4.



(a) The structure of the NN (b) The performance curve of the NN

Figure 4. The structure and the performance curve of the NN.

4. Experiments and discussion

To demonstrate the effectiveness of our method, experiments were performed on the JAFFE database [10]. The training images are the same as those mentioned in Section 3, and the rest images in JAFFE database are test images. For purpose of comparison, several facial expression recognition methods were performed in the experiments. The baseline is the PCA method on the total Gabor wavelet coefficients (5 scales and 8 orientations). The next are methods based on fusion of different channels. The fusion methods include the Max fusion method

$$R_{res\_max} = \max_{e}[\sum_{c}(R(c,e)\max_{e} r(c,e))] \qquad (4)$$

the Sum fusion method

$$R_{res\_max} = \max_{e}[\sum_{c} R(c,e)r(c,e)] \qquad (5)$$

and the fusion method is based on neural network (NN). The recognition rates for different methods are shown in Table 3.

It can be noticed that the three data fusion methods provided superior performance to the original Gabor PCA method and the NN fusion method performed better than the other two fusion methods, Max fusion and Sum fusion. The experimental results showed that the fusion of Multiple Gabor features from different channels can provide better performance for facial expression recognition.

Table 3. Recognition rate (%) for different method

Ex.	Gabor PCA	Max Fusion	Sum Fusion	NN fusion
AN	60	70	80	95
DI	67	78	87.5	88
FE	80	82	88	100
HA	56	67	70	100
NE	67	72	72	75
SA	56	89	89	90
SU	89	100	100	100

5. Conclusion

In this paper, a multiple Gabor features based facial expression recognition algorithm is proposed. In order to verify the effectiveness of the algorithm, the corresponding algorithms are evaluated on the JAFFE database. The performance of the algorithm mainly relies on the scale and orientation features selected as well as the final combination strategy. The experiments indicate that the fusion methods perform better than the original Gabor method (regarding all channel features as one vector) and neural network based fusion method performs best.

The future work is to find more discriminating features and more efficient fusion method for facial expression recognition.

References

[1] A.J. Calder, A.M. Burton, P. Miller, etc. A principal component analysis of facial expressions. *Vision Research,* 41:1179-1208, 2001.

[2] S. Dubussion, F. Devoine, M. Masson. A solution for facial expression representation and recognition. *Signal Processing: Image Communication,* 17:657-673, 2002.

[3] M. Matsugu, K. Mori, Y. Mitari, Y. Kaneda. Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Networks,* 16:555-559, 2003.

[4] Yongzhao ZHAN, Jingfu YE, Dejiao NIU, Peng CAO. Facial Expression Recognition Based on Gabor Wavelet Transformation and Elastic Templates Matching. *Proceedings of the Third International Conference on Image and Graphics*, 2004.

[5] Y.Yacoob, L.S. Recognizing human facial expressions from long image sequences using optical flow. *IEEE Trans. PAMI,* 16(6): 636-642, 1994.

[6] Y. Zhu, L.C. De Silva, C.C. Ko. Using moment invariants and HMM in facial expression recognition. *Patter Recognition Letters.* 23:83-91, 2002.

[7] T. Hu, L.C. De Silva, K. Sengupta. A hybrid approach of NN and HMM for facial emotion classification. *Pattern Recognition Letters,* 23:1303-1310, 2002.

[8] M. Lades, J. Vorbruuggen, J. Buhmann, J. Lange, W. Konen, C. von der Malsburg, R. Wurtz. Distortion Invariant Object Recognition in the Dynamic Link Architecture. *IEEE Trans. Computers,* 42(3):300-311, 1993.

[9] L. Wiskott, J. Fellous, N. Kruger, C. von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Trans. PAMI,* 19(7):775-779, 1997.

[10] M. Lyons, S. Akamatsu, etc. Coding Facial Expressions with Gabor Wavelets. *Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition,* Nara Japan, 200-205, 1998.

[11] Hui Jin, Wen Gao. The Human Facial Combined Expression Recognition System. *Chinese Journal of Computer,* 23(6):602-608, 2000.

[12] M. Pardas, A. Bonafonte. Facial animation parameters extraction and expression recognition using Hidden Markov Models. *Signal Processing: Image Communication* 17:675-688,2002.

[13] Ekman P., Friesen W.V. Unmasking the Face: A guide to recognizing emotions from facial clues. Consulting Psychologists Press 1975.

[14] G.Donato, t M Bartlet., J.C. Hager, P. Ekman, T. Sejnowski. Classifying Facial Actions. *IEEE TRANS. PAMI,* 21(10):974-989,1999.

[15] J. Kittler, M.Hatef, R. Duin, J. Matas. On Combining classifiers. *IEEE Trans. PAMI,* 20(3):226-239, 1998.

[16] Wei Fan, Yunhong Wang, Wei Liu, Tieniu Tan. Combining Null Space-based Gabor Features for Face Recognition. *Proceedings of the 17th International Conference on Pattern Recognition, Cambridge,* UK(2004).

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE