Modeling Facial Expression Space for Recognition *

Yuwen Wu
National Lab. on Machine Perception
Peking University
Beijing, China
wuyw@cis.pku.edu.cn

Hong Liu
National Lab. on Machine Perception
Peking University
Beijing, China
liuhong@cis.pku.edu.cn

Hongbin Zha
National Lab. on Machine Perception
Peking University
Beijing, China
zha@cis.pku.edu.cn

Abstract – **In this paper, we present a method of modeling facial expression space for facial expression recognition by fuzzy integral. In traditional expression recognition methods using shape features, there are problems in describing both the uncertainty in facial expression classification and the relationship between facial features and facial expressions. Using facial expression space model, those problems can be solved easily. Firstly, we use values of fuzzy integral in different facial expression spaces to describe the uncertainty of facial expression. Secondly, by the fuzzy measure automatically constructed in each facial expression space, we deal with different effects of facial features for facial expression classification. Experiments show this method has a good ability of describing the uncertainty of facial expression and acquires good results of classification.**

Keywords – Facial Expression Recognition, Facial Expression Analysis, Fuzzy Integral, Fuzzy Measure

I. INTRODUCTION

Facial expression plays an important role in cognition of human emotion. For natural and harmonious human-robot communion, it is essential to understand human intent effectively. Facial expression recognition is the base of emotion understanding. At the same time, it is an effective way for understanding human emotion. Human brain can recognize facial expression just by facial shape features [1,2]. Shape features have some characteristics: the dimension of feature vector is small and the computational complex of training and recognizing is low. The aim of this paper is to design a new algorithm that can use shape feature's characteristics effectively and give a good description of expression uncertainty.

There have been more than 20 years since the facial expression recognition was primarily a research subject for computer science [3]. The method of facial expression recognition based on shape feature is always an important research subject. The traditional methods use the characteristic that the shape feature dimensionality is small [4,5,6]. Algorithms those methods used are simple. Recent methods adopt some complex algorithms [7,8,9,10,11]. But most of those methods use shape features to recognize Action Units (AU) [12,13]. There are two problems in using shape feature to recognize facial expression. Firstly, the uncertainty is a remarkable characteristic of facial expression recognition. The uncertainty of facial expression is that one expression belongs to several expression classes with different possibilities. Therefore, the description of uncertainty should base on the result of classification. Secondly, the movement scope of different features is different; such as the movement scope of mouth features is lager than that of eye features. At the same time, different expressions cause different facial movements. How to use those two characteristics will affect results of facial expression classification.

This paper designs a method of modeling facial expression space for expression recognition by analyzing above problems. Firstly, we use an automatic method to construct fuzzy measures on each facial expression space. Those fuzzy measures are models of expression spaces and describe different facial movements caused by different expressions. Secondly, the values of fuzzy integral computed on each expression space are regarded as possibilities of different facial expressions for an input sample. Therefore, with those values of fuzzy integral, our algorithm gives a method for describing uncertain facial expression.

This paper is organized in the following manner. Section II will describe the theory of fuzzy measure, fuzzy integral and how to use fuzzy integral to classify. Section III provides our method of shape feature extraction and representation. The method of modeling facial expression space will be shown in section IV. Section V describes the technique of facial expression recognition using the model of expression space. Experimental results are shown in Section VI. Finally, concluding comments are covered in Section VII.

II. FUZZY MEASURE AND FUZZY INTEGRAL

Fuzzy integral is a method of evidence fusion [14,15]. It nonlinearly combines objective evidences. Underlying the concept of fuzzy integral is that of fuzzy measure. Here, we first show the definition of fuzzy measure.

A fuzzy measure g defined on X is a set function $g : P(X) \to [0,1]$ satisfying the following axioms:

1) $g(\Phi) = 0$, $g(X) = 1$.
2) $A \in B \Rightarrow g(A) \le g(B)$.
3) $\lim_{i\to\infty} g(A_i) = g(\lim_{i\to\infty} A_i)$, if $\{A_i\}_{i=1}^{\infty}$ is an increasing sequence of measure sets.

$P(X)$ indicates the power set of X. Starting from this definition, Sugeno [16] introduced a so-called g_λ fuzzy measure that satisfies the following additional property:

$$g(A \cup B) = g(A) + g(B) + \lambda g(A) g(B) \quad (1)$$

for all $A, B \in X$ and $A \cap B = \Phi$, and for some $\lambda > -1$. In general, the value of λ can be determined owing to the boundary

* This work is supported by National Natural Science Foundation of China (NSFC). Project No: 60175025, P.R.China.

condition of the g_λ fuzzy measure. Let X be a finite index set $X = (x_1,...,x_n)$. This condition reads as $g(X)=1$, hence, the value of λ is determined by solving the following:

$$g_\lambda(X) = \frac{1}{\lambda}\left(\prod_{i=1}^{n}(1+\lambda g^i)-1\right),\ \lambda \neq 0 \tag{2}$$

Put it equivalently,

$$G(\lambda) = \prod_{i=1}^{n}(1+\lambda g^i) - \lambda - 1 = 0 \tag{3}$$

where $\lambda \in (-1,+\infty)$, $\lambda \neq 0$ and g^i is the value of the fuzzy density function, $g^i = g_\lambda(\{x_i\})$. The values can be interpreted as the degrees of importance for the corresponding sources of information x_i. The solution can be easily obtained. With the root of $G(\lambda)$, the density of any set $A = \{x_{a1},...,x_{an}\} \in X$ can be calculated by

$$g_\lambda(A) = \frac{1}{\lambda}\left(\prod_{i=a1}^{an}(1+\lambda g^i)-1\right) \tag{4}$$

Fuzzy integral is integral of a real function with respect to a fuzzy measure. The fuzzy integral of function h computed over X of the function h with respect to a fuzzy measure g is defined in the form

$$\int_X h \circ g = \sum_{i=1}^{n}\left(h(x_{(i)}) - h(x_{(i-1)})\right) g(A_{(i)}) \tag{5}$$

where $\bullet_{(i)}$ indicates that the indices have been permuted so that $0 \le h(x_{(1)}) \le ... \le h(x_{(n)}) \le 1$. Also $A_{(i)} = \{x_{(i)},....,x_{(n)}\}$, and $h(x_{(0)}) = 0$. Thus the calculations of the fuzzy integral with respect to a g_λ fuzzy measure can be realized once we provided with the values of the density function g^i available for individual points.

There are many papers using fuzzy integral for problems of classification [10,17,18]. In any case, the classification problem is as follow: Let $C = \{C_1,...,C_M\}$ be a set of classes of interest. Let $\vec{x} = (x_1, x_2,...,x_n)$, where n is the dimension of feature vector, be a set of groups of features for the recognition of a particular class C_m. Here, a group of features also includes a singleton. Let Q be the object under consideration for recognition. Let $h_m : \vec{x} \to [0,1]$ be the partial evaluation of the object Q for class C_m, that is, $h_m(x_i)$ is an indication of how certain we are in the classification of object Q to be in class C_m using the group x_i. Now corresponding to each x_i the degree of importance, $g_m^i = g_m(\{x_i\})$, for class C_m, must be given firstly. Where g_m is fuzzy measure defined on class C_m. With density function, we can give the classification algorithm by fuzzy integral.

```
BEGIN classification
    DO for each class
        DO for each group of features x
            calculate h_m(x⃗)
            determine g_m(x⃗)
        END DO
        calculate λ_m
        compute fuzzy integral
    END DO
END classification
```

The key problem of classification using fuzzy integral is the calculation of density function. Density function describes the effect of different attribute for each class. It can be regarded as the model of class space. If density function is not correct, the result of classification will be bad. In the following, we will give an effective method of calculating density function.

III. Shape Feature Extraction and Representation

This section describes the feature vector construction used in this paper and the method of feature representation. We use Active Appearance Models (AAM) [19,20] to extract facial key points. In our performance, we use 53 key points. Fig. 1 shows a result of feature extraction. Based on the idea of AU [21], facial expression can be decomposed as some basic action units. We select basic facial movements of six kinds of basic facial expressions shown in Table I. And we design the shape feature vector for classification by those basic movements (shown in table II).

After feature extraction, feature vectors will be normalized. In this paper, normalization is constituted with two parts. Firstly, we use two eye inner corners and philtrum to normalize any expression image of one person with his neutral expression image. Secondly, we use those three points to construct a coordinate. The line, which is across the two eye inner corners, is x-axis. The middle point of two eye inner corners is origin. The distance between two eye inner corners is the unit length of horizontal direction, and the distance between philtrum and the line of two eye inner corners is the unit length of vertical direction. With this coordinates, we can normalize any two images.

We regard density function as the model of facial expression space, because it is defined on special expression class space and reflects the characteristics of this subspace. How to calculate density function is the key problem of recognition using fuzzy integral. It can be either estimated subjectively or obtained from the training data [22]. In this paper, we use Adaboost [23] and the estimate of error rate of classification to estimate densities.

TABLE I
MOVEMENT FOR FACIAL EXPRESSION

Anger	Eyebrows down and squeezed, eyes widened, lips tightened
Disgust	Eyes narrowed, mouth open, lip corners pushed down and inward
Fear	Eyebrows up and squeezed, eyes narrowed, mouth open
Sad	Eyebrow inner corners raised, eyes narrowed
Happy	Eyes closed, mouth widened, lip corners pushed up and outward
Surprise	Eyebrow up, eyes widened, mouth widened

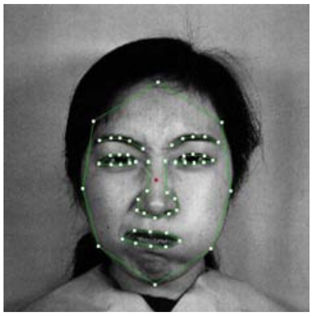

Fig. 1 The result of feature extraction

TABLE II
SHAPE FEATURE VECTOR

1	The displacement of left eyebrow inner corner in vertical direction
2	The displacement of left eyebrow middle point in vertical direction
3	The displacement of right eyebrow inner corner in vertical direction
4	The displacement of right eyebrow middle point in vertical direction
5	The distance between left eyebrow and right eyebrow
6	The height of left eye
7	Left upper eyelid middle point's displacement in vertical direction
8	Left lower eyelid middle point's displacement in vertical direction
9	The height of right eye
10	Right upper eyelid middle point's displacement in vertical direction
11	Right lower eyelid middle point's displacement in vertical direction
12	The width of mouth
13	The height of mouth
14	The displacement of left mouth corner in vertical direction
15	The displacement of right mouth corner in vertical direction
16	The displacement of upper lip middle point in vertical direction
17	The displacement of lower lip middle point in vertical direction
18	The displacement of left mouth corner in horizontal direction
19	The displacement of right mouth corner in horizontal direction

IV. Modeling Facial Expression Space

Adaboost is a learning algorithm that selects a small number of weak classifiers from a large weak classifier pool. Adaboost sorts weak classifiers by ability of classification. Therefore, we can use Adaboost to sort attributes by their effect for special expression class. The details of attribute sorting with Adaboost algorithm is showed as follow:

Given $(\vec{x}_1, y_1),\ldots,(\vec{x}_{N_m}, y_{N_m})$; $\vec{x}_i \in X$, $y_i \in \{-1,1\}$ for class C_m

Initialise weights $D_1(n) = \dfrac{1}{N_m}$

For $i = 1,\ldots,I$

1. Find $H_i^m = \arg\min_{h_j} \xi_j; \xi_j = \dfrac{1}{2}\left[1 - \sum_{n=1}^{N_m} D_i(n) y_n H_n^m(\vec{x}_n)\right]$
2. Set $\alpha_i^m = \dfrac{1}{2}\log\left(\dfrac{1-\xi_i}{\xi_i}\right)$
3. Update $D_{i+1}(n) = \dfrac{D_i(n)\exp(-\alpha_i^m y_n H_i^m(\vec{x}_n))}{Z_i}$, where Z_t is normalization factor so that D_{i+1} is a PDF.

Output the final coefficient α_i^m

Here, $H = sign[f_{att}(\vec{x}) - T]$ is just threshold-type weak classifier, whose output is Boolean value. Where $f_{att}(\vec{x})$ is the attribute value of shape feature vector and T is a threshold. I equals to the dimension of feature vector (It is 19 in this paper). The α_t^m is reliability coefficient, which represents dynamically elicited knowledge that reflects the confidence of each attribute for class C_m.

We use the k-nearest neighbors to get the estimate of the error rate for each class. A facial expression is confusable with some expressions, but entirely different with the others. For example, anger is confusable with disgust, but not confusable with happy and surprise. Therefore, we use a set of error rates for each class to estimate the final error rate. Firstly, we group a class set, $S^m = \{S_1^m,\ldots,S_{n_m}^m\}$, which element is a set of classes, for class C_m. A class set at least includes one element, which is comprised of all classes (six classes in this paper). And the other elements in class set are comprised of class C_m, other classes confused with class C_m. For example, the class set of class anger is comprised of two elements, $S^{an} = \{S_1^{an}, S_2^{an}\}$. S_1^{an} includes six classes and S_2^{an} includes two classes: class anger and class disgust. Secondly, for each element, S_n^m, of class set for class C_m, we calculate the error rate of classification, $(\vec{\varepsilon}^m)^j = \{(\varepsilon_1^m)^j,\ldots,(\varepsilon_n^m)^j\}$, for all attributes, where n is the dimension of feature vector and $1 \le j \le n_m$. Then we calculate the average of error rate, $\vec{\varepsilon}^m = \{\varepsilon_1^m,\ldots,\varepsilon_n^m\}$, on class set as the error rate for classification of attribute about class C_m, where $\varepsilon_i^m = \sum_{j=1}^{n_m} (\varepsilon_i^m)^j / n_m$. With α_i^m and ε_i^m, we can use (6), mentioned in [10,24], to estimate g_m^i of class C_m.

$$g_m^i = \frac{\alpha_i^m(1-\varepsilon_i^m)}{\sum_{j=1}^{n} \alpha_j^m(1-\varepsilon_j^m)} \tag{6}$$

When $\sum_{i=1}^{n} g^i \ge 1$ and n is big enough, the root of $G(\lambda)$ may be very close to –1. For the limitation of the precision range of computer capability, the output root will be –1 in some cases. If $\sum_{i=1}^{n} g^i < 1$, the root of $G(\lambda)$ will be in the open interval $(0,+\infty)$. Then we use (7) to estimate densities.

$$g_m^i = \frac{\alpha_i^m(1-\varepsilon_i^m)}{n\sum_{j=1}^{n} \alpha_j^m(1-\varepsilon_j^m)} \tag{7}$$

Each set of densities is calculated on special class. It describes the characteristics of a subspace, where an expression class locates, of feature vector space. So we regard them as a model of facial expression space.

V. Facial Expression Recognition by Fuzzy Integral

To use fuzzy integral for recognizing facial expression, we need a set of classifiers. In this paper, the fuzzy C-means clustering (FCM) algorithm [25] is used to design those classifiers. FCM can partition a data set of attribute into n fuzzy clusters. The partitioning is based on a least squares minimization of the fuzzy within-group objective function. The FCM adopts Euclidean distance to measure the degree of the membership. It doesn't consider the covariance of class. With different covariance, a point has a bigger distance with a class, but it may be have a higher probability with this class. So we adopt the Mahalanobis distance [26] as the measure of distance and change the objective function of FCM:

$$E_i = \sum_{m=1}^{M}\sum_{l=1}^{N}\left(\mu_{ml}^i\right)^p \left(x_i^l - v_i^m\right)\left(s_i^m\right)^{-1}\left(x_i^l - v_i^m\right) \tag{8}$$

where M is the number of classes, N is the number of training data, x_i^l is the i^{th} attribute of the l^{th} sample, v_i^m is the center of the m^{th} class for the i^{th} attribute, s_i^m is the covariance of the m^{th} class for the i^{th} attribute, u_{ml}^i is the degree of membership for the i^{th} attribute of the l^{th} sample in the m^{th} class. The degree of membership is calculated by (9), and the center and the covariance of every class are updated by (10) and (11).

$$\mu_m^i(\bar{x}_l) = \frac{\left(s_i^m \big/ \left(x_i^l - v_i^m\right)^2\right)^{1/(p-1)}}{\sum_{k=1}^{M}\left(s_i^k \big/ \left(x_i^l - v_i^k\right)^2\right)^{1/(p-1)}} \tag{9}$$

$$v_i^m = \frac{\sum_l \left(\mu_{ml}^i\right)^p \left\|x_i^l\right\|}{\sum_l \left(\mu_{ml}^i\right)^p} \tag{10}$$

$$s_i^m = \frac{\sum_l \left(\mu_{ml}^i\right)^p \left\|x_i^l - v_i^m\right\|^2}{\sum_l \left(\mu_{ml}^i\right)^p} \tag{11}$$

For a new sample, we can use (9) to calculate its degrees of membership for each class. Then the integral function is $h_m(x_i) = \mu_m^i(\bar{x})$.

With $h_m(x_i)$ and the density function, the value of fuzzy integral can be computed by (5) easily. For a face image, we first extract its features and construct its shape feature vector, and then calculate its degree of membership of each class for every attribute by FCM. Using the results of FCM and the densities of all classes, the values of fuzzy integral for each class are computed by (5). Those values are regard as this image's possibilities of each class. For classification, we just select the class whose value of fuzzy integral is maximal as the expression label for inputted face image. Fig. 2 shows the architecture of our expression recognition by modeling facial expression space.

VI. EXPERIMENTS

In this paper, the expression modeling is defined as a six-class problem of anger, disgust, fear, sad, surprise and happy. The expression databases used in our experiments are JAFFE database [27] and FGnet facial expression database [28]. The experimental results are comprised of two parts. Densities of all classes will be shown firstly. And then, some results of classification will be shown. The training data of fuzzy measure estimating are same with the FCM.

Fig. 3 shows density functions of six basic expression classes. Training samples come from FGnet expression database. The number of train sample for each class is 600. As we mentioned in section II, the density g_m^i describes the degree of importance of the i^{th} attribute for the m^{th} class, and it reflects the degree of the importance of the i^{th} attribute for the m^{th} class. From Fig. 3, we can see that an attribute, which has a bigger value of density, for an expression class is included in the movements caused by this expression. In another word, the result shown in Fig. 3 is corresponding with Table. I perfectly. As we mentioned, $g_m(A)$ indicates the degree of importance of an attribute set for m^{th} class. In Fig. 4, we select some attribute sets and calculate their densities for six basic expression classes. The results are corresponding with Table. I perfectly, too.

The PCA algorithm is employed for comparative evaluation with our algorithm. Table. III show the results of classification on two expression databases. The feature of PCA is same with our algorithm. For JAFFE database, we use all images as test sample. For FGnet database, we select 3000 images as test sample. The result on FGnet database is better than the result on JAFFE database. It is because the expressions of test images selected from FGnet are clearer than those ones of JAFFE database. Fig. 5 shows the average correct rate of six categories. The reason that the happy and the surprise have better correct rates is that happy and surprise is not confused with other expressions.

TABLE III
EXPERIMENTAL RESULTS OF CLASSIFICATION ON TWO DATABASES

	JAFFE	FGnet
Our Method	83.2%	91.6%
PCA	74.7%	85.6%

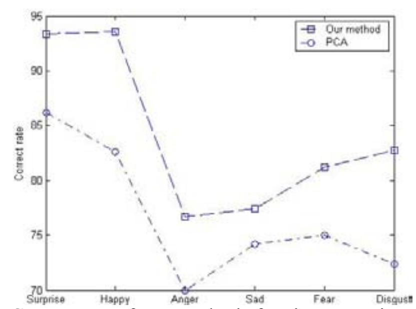

Fig. 5 Correct rate of two methods for six expression classes

The main aim of this paper is giving a reasonable description of expression uncertainty. There are semantic ratings averaged over 60 Japanese female subjects for each image in JAFFE database. We compare those semantic ratings with values of fuzzy integral in Fig. 6. The blue histogram (fuscous one in print document) is values of fuzzy integral and the gray histogram is semantic ratings in JAFFE database. And possibilities calculated by our method have similar distribution with those semantic ratings.

VII. CONCLUSIONS

In this paper, we start from the uncertainty of expression and different facial movements caused by different expressions, and model each facial expression space by fuzzy measure for classification. Experimental results show that our method has good results of classification and present a reasonable description of expression uncertainty. Although this paper focuses on facial expression recognition problems, the framework can be used in other multi-class pattern recognition problems too, if the dimension of feature vector is low.

However, there still has a shortcoming waiting to be overcome. The negative expressions have worse correct rate. The reason is that a negative expression is confusable with another negative expression, but not confusable with a positive expression. In future work, we will use the ideal of grouped feature to design a hierarchical classifier to overcome this shortcoming.

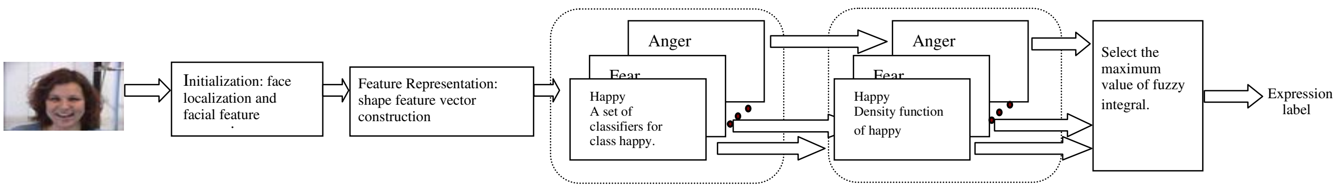

Fig. 2 Architecture for expression recognition by modelling facial expression space

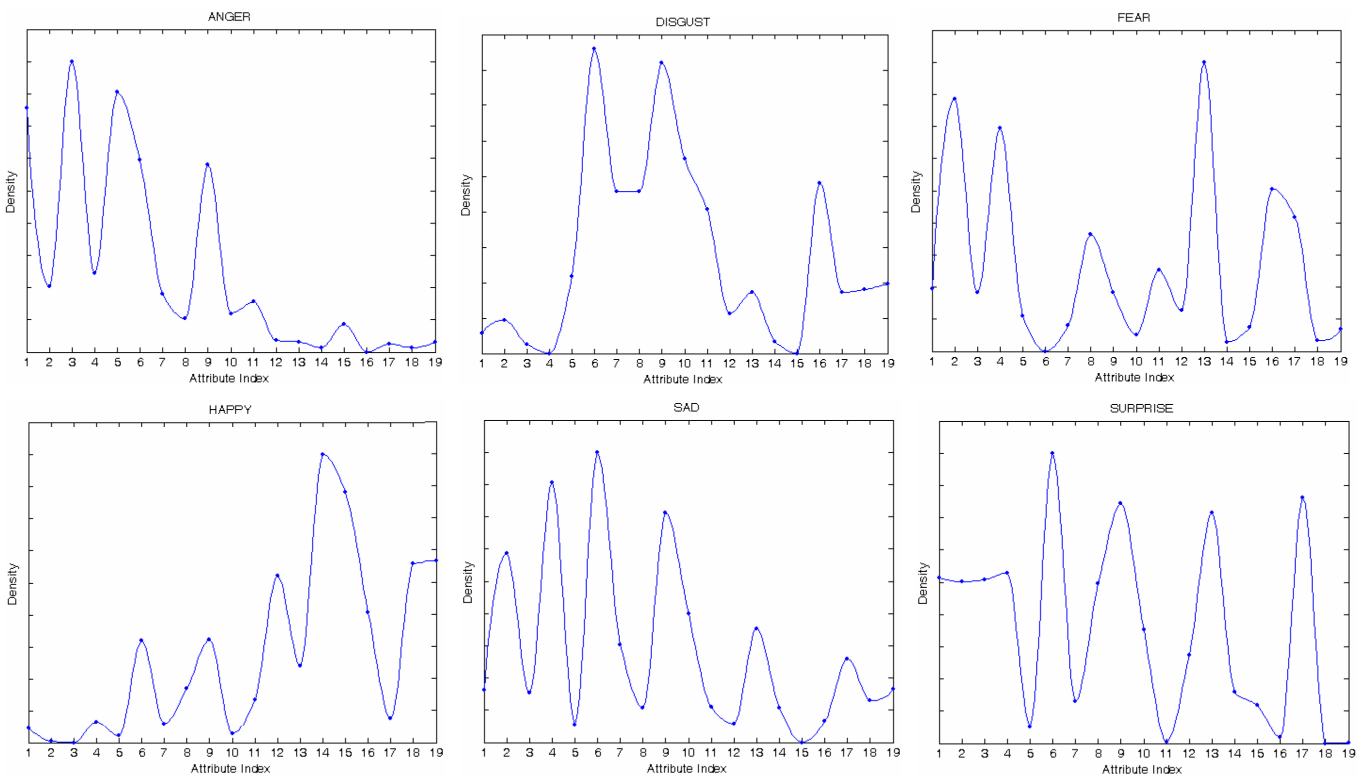

Fig. 3 Density functions for six basic expression classes.

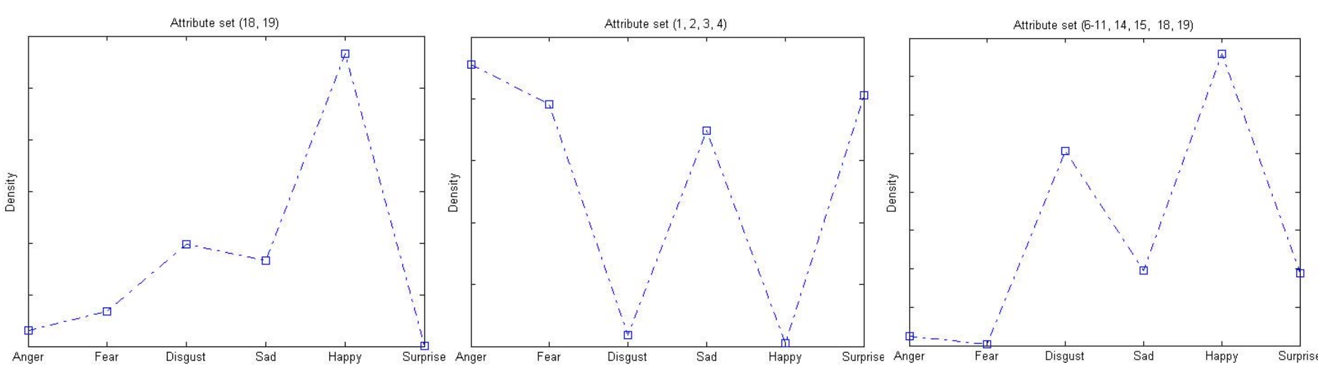

Fig. 4 Densities of attribute sets

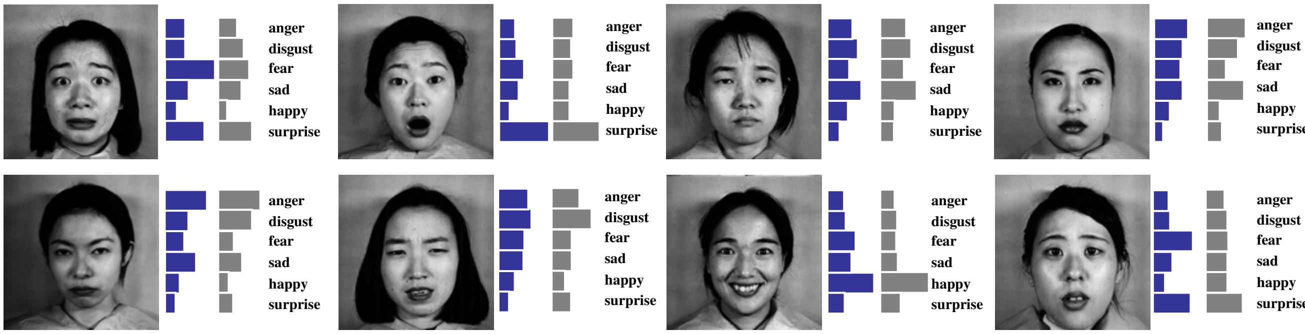

Fig. 6 The comparative results of possibilities and semantic ratings

ACKNOWLEDGEMENT

The work was supported in part by the NSFC Grant (No. 6033010) and NKBRPC (No. 2004CB318000).

REFERENCES

[1] A. J. Calder, R. Duncan, A. W. Young, I. Nimmo-Smith, J. Keane, D. I. Perrett, "Caricaturing facial expressions", *Cognition*, vol. 76, no. 2, pp. 105-146, August 2000.

[2] A. J. Calder, A. M. Burton, P. Miller, A. W. Young, S. Akamatsu, "A principal component analysis of facial expressions", *Vision Research*, vol. 41, no. 9, pp. 1179-1208, April 2001.

[3] B. Fasel, J. Luettin, "Automatic facial expression analysis: a survey", *Pattern Recognition*, vol. 36, no. 1, pp. 259-275, January 2003.

[4] H. Kobayashi, F. Hara, "The recognition of basic facial expressions by neural network", *IEEE Intel. Joint Conf. On Neural Networks*, pp. 460-466, 1991.

[5] A. Ralescu, H. Iwamoto, "Recognition of and reasoning about facial expressions using fuzzy logic", *Proc. 2nd IEEE International Workshop on Robot and Human Communication*, pp. 259-264, 1993.

[6] C. L. Huang, Y. M. Huang, "Facial expression recognition using model-based feature extraction and action parameters classification", *Visual Communication and Image Representation*, vol. 8, no. 3, pp. 278-290, September 1997.

[7] Y. S. Gao, M. K. H. Leung, S. C. Hui, M. W. Tananda, "Facial expression recognition from line-base caricatures", *IEEE Trans. Systems Man and Cybernetics-Part A: Systems and Humans*, vol. 33, no. 3, pp. 407-412, May 2003.

[8] S. Nakamizo, K. Haneda, O. Nakamura, "An effective feature extraction algorithm for the recognition of facial expressions", *Canadian Conf. On Electrical and Computer Engineering (CCECE'04)*, pp. 993-998, 2004.

[9] B. Fasel, "Facial expression analysis using shape and motion information extracted by convolutional neural networks", *Proc. 12th IEEE Workshop on Neural Networks for Signal Processing*, pp. 607-616, 2002.

[10]C. C. Chibelushi, F. Bourel, "Hierarchical multistream recognition of facial expressions", I*EE Proc. -Vis. Image Signal Process*, vol. 151, no. 4, pp. 307-313, August 2004.

[11]F. Bourel, C. C. Chibelushi, A. A. Low, "Robust facial expression recogniton using a state-based model of spatially-localised facial doynamics", *5th IEEE Intel. Conf. On Automatic Face and Gesture Recognition (FGR'02)*, pp. 106-111, 2002.

[12]M. Pantic, L. J. M. Rothkrantz, "Automatic analysis of facial expression: the state of the art", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1424-1445, December 2000.

[13]Y. L. Tian, T. Kanade, J. F. Cohn, "Recognizing action units for facial expression analysis", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 97-115, February 2001.

[14]H. Tahani, J. M. Keller, "Information fusion in computer vision using the fuzzy integral", *IEEE Trans. System Man and Cybernetics*, vol. 20, no.3, May/June 1990.

[15]M. Grabisch, M. Sugeno, "Multi-attribute classification using fuzzy integral", *IEEE Intel. Conf. On Fuzzy Systems*, pp. 47-54, March 1992.

[16]M. Sugeno, "Theory of fuzzy integrals and its applications", PhD Thesis, Tokyo Institute of Technology, 1974.

[17]X. Li, Z. Q. Liu, K. M. Leung, "Detection of vehicles from traffic scenes using fuzzy integrals", *Pattern Recognition*, vol. 35, no.4, pp. 967-980, April 2002.

[18]K. C. Kwak, W. Pedrycz, "Face recognition using fuzzy integral and wavelet decomposition method", *IEEE Trans. System Man and Cybernetics-Part B: Cybernetics*, vol. 34, no. 4, pp. 1666-1675, August 2004.

[19]T. F. Cotes, G. J. Edwards, C. J. Taylor, "Active appearance models", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681-685, June 2001.

[20]http://www.imm.dtu.dk/~aam/

[21]P. Ekamn, W. Friesen, "Facial action coding system: a technique for the measurement of facial movement", Psychologists Press, 1978.

[22]J. M. Keller, J. Osborn, "Training the fuzzy integral", *Approximate Reasoning*, vol. 15, no.1, pp. 1-24, July, 1996.

[23]R. E. Schapire, Y. Singer, "Improved boosting algorithms using confidence-rated predictions", *Machine Learning*, vol. 37, no. 3, pp. 297-336, 1999.

[24]C. C. Chibelushi, F. Deavi, J. S. D. Mason, "Adaptive classifier integration for robust pattern recognition", *IEEE Trans. System Man and Cybernetics-Part B: Cybernetics*, vol. 29, no. 6, pp. 902-907, December 1999.

[25]J. C. Bexdek, "Pattern recognition with fuzzy objective function algorithms", Plenum Press, 1981.

[26]A. Bar-Hen, J. J. Daudin, "Generalization of the Mahalanobis distance in the mixed case", *Multivariate Analysis*, vol. 53, no.2, pp. 332-342, May 1995.

[27]M. J. Lyons, S. Akamatsu, M. Kamachi, et al, "Coding facial expressions with Gabor wavelets", *3rd IEEE Intel. Conf. Automatic Face and Gesture Recognition*, pp.200-205, 1998.

[28]http://www.mmk.ei.tum.de/~waf/fgnet/feedtum.html