Recognition of and reasoning about facial expressions using fuzzy logic

Anca Ralescu[1 2] Hiroshi Iwamoto
Laboratory for International Fuzzy Engineering Research
Siber Hegner Building 4FL.
89-1 Yamashita-cho, Naka-ku, Yokohama 231 JAPAN
Email: anca@fuzzy.or.jp Email: iwamoto@fuzzy.or.jp

Abstract - In the study of the linguistic modeling of facial images we have been previously concerned with deriving qualitative descriptions such as "big eyes, long hair" of face components. To enhance this system we extend our approach at deriving higher level, qualitative descriptions. In particular, we focus on describing facial expressions.

Our approach is that of qualitative modeling based on fuzzy number modeling. The result of this modeling method is a collection of fuzzy if-then rules obtained from input-output data. The input data consists of measurements of the movement of facial parts associated to different facial expressions. The output data consists of scores for face images collected using a questionnaire. In this paper, we show the modeling result obtained from this method for the facial expression "happy".

While the modeling results are satisfactory the initial recognition results are limited, due in part to the absence of the models for the remaining facial expressions.

I. Introduction

The face plays an important role in communication between humans [1, 2]. Similarly, to achieve human friendly computer systems, capable of meaningful communication, the ability for the computer to understand, interpret and express the information conveyed by the human face becomes important. As such, recently, [4, 5, 10, 11], face image processing has shifted from an object recognition and identification problem to an active understanding and interpretation of the face.

As part of a flexible image retrieval system of face photographs based on their linguistic descriptions, our previous work [3, 7], focused on the generation of such descriptions from the image. We required these descriptions to be (i) natural language-like, and (ii) general but also accurate in order to insure an adequate retrieval rate. Using fuzzy logic for representation and reasoning, descriptions like "big eyes, long hair" were produced from measurements of the face components. Since these descriptions are highly subjective their validation was demonstrated by successful retrieval operation [6].

To expand our previous work we aim at higher level descriptions, such as facial expressions, gender, and age. In this paper we present a fuzzy logic based qualitative modeling [8, 9], for recognition of facial expressions.

II. Modeling of Facial Expression

We use a system modeling approach where (1) the input is a collection of face photographs showing different facial expressions, and (2) the output is the impression made by these expressions on human subjects. The method we adopt is the fuzzy logic based qualitative modeling [9].

The outline of our research is as follows: photographs of faces (data from [4, 5]) showing different expressions were given to a sample of users for assessment: each user was asked to assign a score for each photograph within one or more facial expressions. Six basic facial expressions, Surprise, Fear, Disgust, Anger, Happiness, Sad, were used. The score, called *index of impression* is a value in [0, 1] and constitutes the output of the system.

The input is derived directly from the result of image processing: the movement of mouth, eye and eyebrow derived from the difference of the coordinates of the facial characteristic points (FCP) [4, 5].

The qualitative modeling provides (1) the structure identification, i.e. decides the input variables (FCP) needed for detection of facial expressions and (2) the system identification, i.e. (fuzzy IF-THEN) rules relating these input variables to clusters of the output (the index of impression corresponding to different facial expressions).

IEEE International Workshop on
Robot and Human Communication
0-7803-1407-7/93 $3.00 1993 © IEEE

In addition this method gives a linguistic description of the expression (or, alternatively, an explanation of the recognition). (Fig. 1)

III. EXPERIMENTAL RESULT

Initial experimental result were obtained for modeling the facial expression 'happy'. The input variables automatically selected by the method are: "movement of the lower lips", "movement of the lower eyelid" and "distance between the most outer points of eyes" or "inclination of eyebrows". A typical rule for this is: ***"If movement of the lower lip is negative and movement of the lower eyelid is negative and distance between the most outer points of eyes is positive or inclination of eyebrows is negative then the expression is very happy".***

Next we will show in detail the modeling steps.

Input Data

The image data used in our experiment, obtained from [4, 5], consists of photographs of six facial expression for 20 subjects (Fig. 2). Ten of these photographs were selected at random to be used as training data, while the remaining were left for test purposes.

To extract physical features from photographs, the x-y coordinates of FCP for each facial expression were measured. The change/movement of facial part (eyebrow, eye, mouth) between a facial expression and the normal face is used as the input and is given by the difference between the x or y coordinates of FCPs corresponding each facial parts. [4] (see Table 1).

Output Data

After the questionnaire described before, the numerical scores obtained for each photograph were then averaged over the number of users. Thus, for each expression of a photograph we obtained the average *index of impression*, used to quantify the perceived strength of the facial expression.

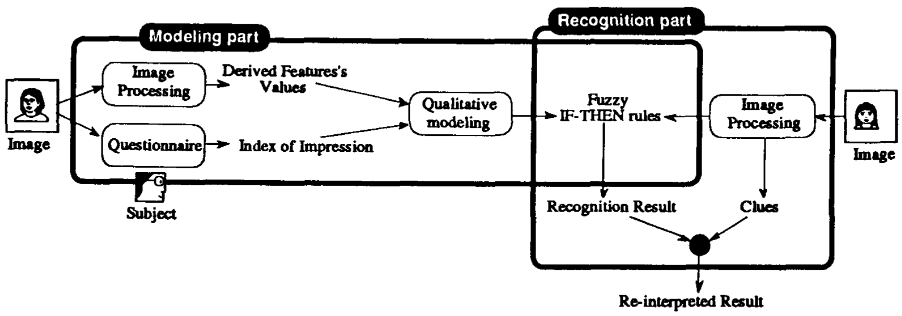

Figure 1 System Diagram

(a) normal (b) surprise (c) happy

Figure 2 Sample of face data.

Table 1 Input data

1	Distance between the most outer points of eyes
2	Openness of eyes
3	Openness of eyes
4	Openness of eyes
5	Movement of upper eyelids
6	Movement of upper eyelids
7	Movement of upper eyelids
8	Movement of lower eyelids
9	Movement of lower eyelids
10	Movement of lower eyelids
11	Distance between inner ends of eyebrows
12	Rise of eyebrows
13	Rise of eyebrows
14	Rise of eyebrows
15	Inclination of eyebrows
16	Width of mouth
17	Openness of mouth
18	Openness of mouth
19	Movement of the upper lip
20	Movement of the upper lip
21	Movement of the lower lip

Fuzzy number modeling

The clustering algorithm of [12], fuzzy c-means method (FCM) is used to cluster the output. To get the optimal number of clusters, we start with two clusters, and apply the algorithm of [13], based on S(c) calculated as:

$$S(U,V,c) = \frac{\sum_{k=1}^{n}\sum_{i=1}^{c}\mu_{ik}^{m}(|x_k - v_i|^2 - |v_i - x|^2)}{Var(x)} \quad (1)$$

where, U is the matrix $U=(\mu_{ik})$ and μ_{ik} is the degree to which the data x_k belongs to the cluster i; $V = (V_1, V_2, \ldots, V_c)$ with V_i being the prototype of i-th cluster; c is number of clusters ; n is the number of data; m is an adjustable weight; x is the average value of x_k; var(x) is the variance of x_k.

The results for our data are shown in Table 2: it can be seen that the smallest value for S(c) is obtained for c=3. The optimum number of clusters found is 3.

Table 2 Results of the FCM algorithm

# of cluster	S(c)
2	-1.784183
3	-1.797683
4	-1.771508

We consider next the task of choosing the input variables among the 21 possible variables. To do this we use as in [9], the regularity criterion (RC) calculated as follows:

The input-output data (21 inputs, one output), sorted by the output, is divided into two representative groups: A, and B. For the divided training data, we project the FCM result (fuzzy clusters, i.e. characterized by a membership function) on the input space. This operation determines fuzzy sets on the range of each input variable. For ease of future calculations the fuzzy sets are approximated by using trapezoid shape. Finally, the RC is given by:

$$RC = \frac{(E_{AB} + E_{BA})}{2}$$
$$= \frac{1}{2}\left(\frac{\sum_{i=1}^{NB}(Y_{Bi} - Y_{ABi})^2}{NB} + \frac{\sum_{i=1}^{NA}(Y_{Ai} - Y_{BAi})^2}{NA}\right) \quad (2)$$

where, NA, NB are the number of elements in A and B respectively; Y_A, Y_B are the output (index of impression) corresponding to A, and B respectively; Y_{AB} is the model output when the model identified using the data of B is used to predict the output corresponding to A; Y_{BA} is the model output when the model identified using the data of A is used to predict the output corresponding to B.

RC measures the average error when the models identified by groups A and B are used on the inputs of B and A respectively. Thus the smaller RC is the better the model.

RC is used to select the input variables to be used in the model: We start with the models corresponding to each input variable. The variable corresponding to the smallest value of RC is selected. Next all the possible combinations of the variable selected and the remaining variables are considered. Again the choice is according to the minimum value of RC. The procedure is repeated increasing the number of input variables by one at each step. The procedure stops when the minimum value of RC fails to decrease. Table 3 shows the results of this procedure for our data:

Table 3. Structure identification

Step	Input variable	RC
Step 1	x21	0.0321
Step 2	x21 & x10	0.0278
Step 3	x21 & x10 & x1	0.0271
	x21 & x10 & x15	0.0271
Step 4	x21 & x10 & x1 & x15	0.0271
	x21 & x10 & x1 & x8	0.0271
	x21 & x10 & x15 & x1	0.0271

Thus we stop at step 3 and select the variables as follows:

x21: Movement of the lower lip
x10: Movement of the lower eyelids
x1 : Distance between the most outer points of eyes
x15: Inclination of eyebrows

After the structure identification the whole training data (merging A and B) is considered again and the membership functions for the selected input variables (X21, X10, X1, X15) are re-approximated. The fuzzy model for the facial expression "happy" is finally obtained. Figure 3 shows the three rules of this model, corresponding to the three clusters of the output values. Note that the model is expressed in terms of unlabeled fuzzy sets. However, the meaning of the relationships expressed by the model can be easily captured by inspection.

IV. Model Evaluation

Given the measurements extracted from a face photograph, fuzzy reasoning is performed on the rules of the model. The result, after defuzzification, is the model's approximation of the index of impression.

To evaluate the model we use it on the inputs corresponding to the training and test data. Ideally the model should recognize all happy faces (highest values of the index of impression for happy in the questionnaire) as happy with a strength which is not lower than that for the faces corresponding to other expressions. In terms of the index of impression it means that the model result should be within some acceptable errors from the questionnaire. In terms of the membership function it means that the result and the index of impression should belong to the same cluster.

The model obtained here has not been tuned. This means that the membership functions used are all as derived from the previously described procedure. The results of applying the model to the training and test data are in Tables 4 and 5 respectively.

Table 4. Questionnaire and Model output for the training data

	Model	Questionnaire
GORO1	0.00585	0.00300
GORO2	0.00585	0.00000
GORO3	0.00585	0.00000
GORO4	0.00586	0.00000
GORO5	0.71636	0.87000
GORO6	0.00585	0.00000
QC1	0.00585	0.00000
QC2	0.04519	0.02700
QC3	0.00585	0.00000
QC4	0.00585	0.00000
QC5	0.71636	0.66200
QC6	0.35856	0.02000
KUMA1	0.71636	0.67000
KUMA2	0.00585	0.00000
KUMA3	0.00585	0.00000
KUMA4	0.00585	0.00000
KUMA5	0.36111	0.74400
KUMA6	0.00585	0.00000

A global evaluation of the model can also be done, on the training and test data separately, by calculating the average square error between the output of the model and the index of impression obtained from the questionnaire. The results are in the Table 6 below:

Table 5 Questionnaire and Model output for the test data

	Model	Questionnaire
MIZO1	0.00585	0.01714
MIZO2	0.00585	0.00000
MIZO3	0.00585	0.00000
MIZO4	0.00585	0.00000
MIZO5	0.00585	0.69571
MIZO6	0.00585	0.00000
MORI1	0.00585	0.00000
MORI2	0.00585	0.07143
MORI3	0.10874	0.09857
MORI4	0.00585	0.00000
MORI5	0.00585	0.70857
MORI6	0.00585	0.00000
PIZ1	0.00585	0.10714
PIZ2	0.36111	0.00000
PIZ3	0.00585	0.00000
PIZ4	0.00585	0.00000
PIZ5	0.00585	0.86143
PIZ6	0.00000	0.00000

Table 6 Average square error of the model

Data set	error
training	0.022032
test	0.1026

V. Other Issues

The basic modeling and recognition method can be expanded by using clues (such as wrinkles, dimples, frowns) from different parts of the face. These clues help reinterpret a facial expression. For example, an upturned crease in the nose-mouth region is often associated with a smiling (happy) face. If, for a particular face, based on the model, we derive a low degree of happiness but we identify this clue we may reinterpret the facial expression. In [7] we presented a method for identifying such clues.

The system can be expanded to distinguish between two mixed expressions such as "a sad yet smiling face", and "smiling yet sad face".

In the modeling stage, the same face could be assigned several expressions. It follows that when we use the model for a given face, several facial expressions could be identified each with a corresponding strength. Here the most important point is that this step requires no additional training.

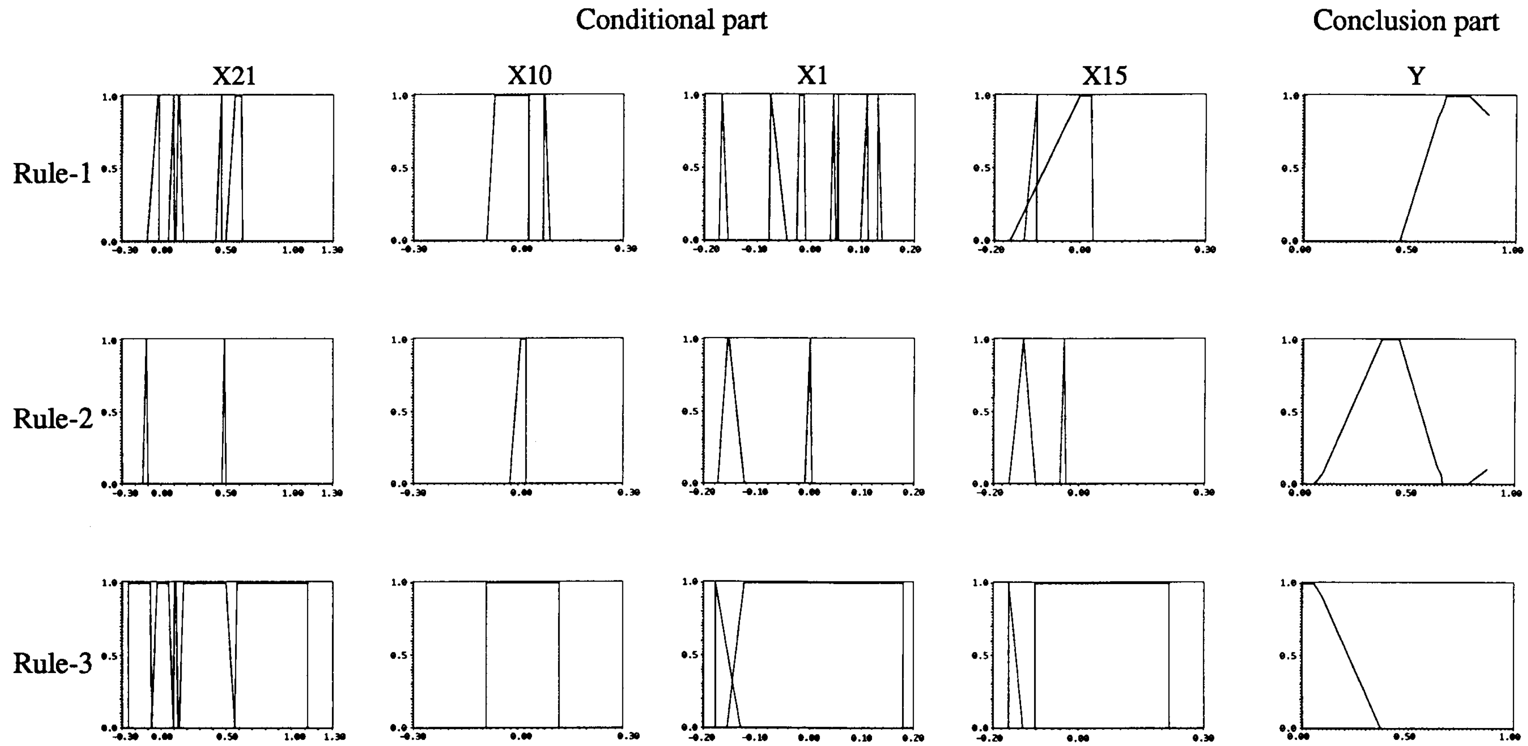

Figure 3. The three rules corresponding to the model of happy

VI. CONCLUSION

Fuzzy logic provides a unified treatment for modeling facial expressions, allowing the link between the quantitative data obtained from measurements of the face and the symbolic descriptions of these as facial expressions.

Concrete results were presented for the model of the expression happy. While the model's tendencies appear to be correct, the recognition results are rather limited: sometimes the model fails to distinguish between the facial expression happy and other facial expressions.

To improve the recognition results the model of the remaining five basic facial expressions must be obtained.

We expect the further expansion of the system to take into account different criteria, such as clues from the face itself, or context information to derive a description, explanation, reinterpretation of a facial expression.

ACKNOWLEDGMENT: The authors wish to thank Professor Fumio Hara of Tokyo Science University for gracefully providing the face image data used for experiments.

REFERENCES

[1] V. Bruce, "What the human face tells the human mind: Some challenges for the robot-human interface," in *Proceedings of R O-MAN'92*, 1992, pp. 44-51.

[2] P. Ekman, W. V. Friesen, *The Facial Action Coding System,*. Palo Alto, California: Consulting Psychologists Press Inc., 1978.

[3] H. Iwamoto, A. Ralescu, "Towards a multimedia model-based image retrieval system using fuzzy logic," presented at SPIE OE/TECHNOLOGY, Boston, MA, , November 15-20, 1992

[4] H. Kobayashi, F. Hara, "Recognition of six basic facial expressions and their strength by neural network," in *Proceedings of RO-MAN'92* , 1992, pp. 381-386.

[5] H. Kobayashi, F. Hara, "Recognition of mixed facial expressions by neural network," in *Proceedings of RO-MAN'92* , 1992, pp. 387-391.

[6] M. Nakayama, K. Miyajima, H. Iwamoto, T. Norita, "Interactive Human Face Retrieval System Based on Linguistic Expression," in *Proceedings of Iizuka '92*, pp.683-686, Fuzzy Logic System Institute, Iizuka, 1992.

[7] M. Nakayama, T. Norita, A. Ralescu, "A Fuzzy Logic based Qualitative Modeling of Image Data," in *Proceedings of IPMU'92*, pp. 615-618, Mallorca, Universitat de les Illes Balears, 1992.

[8] J. Ozawa and M. Sugeno, "Linguistic modeling of numerical data" in *Proceedings of 6th Fuzzy System Symposium*, pp. 159-164, Tokyo September, 6-8, 1990.

[9] M. Sugeno and T. Yasukawa, "A fuzzy logic based approach to qualitative modeling" preprint. Also in IEEE-TFS Vol. 1, February 1993.

[10] M. A. Turk, Interactive-Time Vision: Face Recognition as Visual Behavior. Ph.D. Thesis MIT, September 1991.

[11] M. A. Turk, A. P. Pentland, "Face Recognition Using Eigenfaces," in *Proceedings of CVPR-91*, pp.586-591, IEEE Computer Society Press, Hawaii, 1991.

[12] J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.

[13] Y. Fukuyama, M. Sugeno, "A new method of choosing the number of clusters for fuzzy c-means method", in *Proceedings of 5th fuzzy system symposium*, 1989, pp.247-250.

[1]On leave from the Computer Science Department, University of Cincinnati, Cincinnati, Ohio 45221, U.S.A.

[2]This work was partially supported from the NSF Grant INT91-86032.