2003 International Symposium on Advanced Intelligent Systems
September 25-28, 2003, Jeju, Korea

Design of A Personalized Classifier using Soft Computing Techniques and Its Application to Facial Expression Recognition

Dae-Jin Kim and Zeungnam Bien

Dept. of Electrical Engineering and Computer Science, KAIST
373-1 Guseong-dong, Yuseong-gu, Daejeon, 305-701, South Korea
Tel./Fax.: +82-42-869-8019/8750
E-mail: zbien@ee.kaist.ac.kr, djkim@ctrsys.kaist.ac.kr

***Abstract*— In this paper, we propose a design process of 'personalized' classification with soft computing techniques. Based on human's thinking way, a construction methodology for personalized classifier is mentioned. Here, two fuzzy similarity measures and ensemble of classifiers are effectively used. As one of the possible applications, facial expression recognition problem is discussed. The numerical result shows that the proposed method is very useful for on-line learning, reusability of previous knowledge and so on.**

I. Introduction

According to the definitions in the dictionary [9], the term 'personalize' is used as follows;

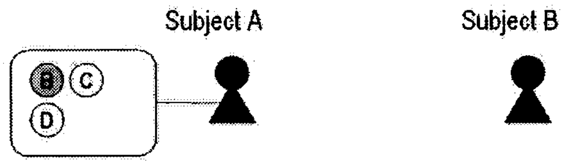
1) If an object is *personalized*, it is marked with the name or initials of its owner.
2) If you *personalize* something, you do or design it specially according to the needs of an individual or to your own needs. (=customize)
3) If you *personalize* an argument, discussion, idea, or issue, you consider it from the point of view of individual people and their characters or relationships, rather than considering the facts in a general or abstract way.

Simply speaking, 'personalized' classification is considered as 'customized' (or 'custom-tailored') classification. For example, Fig. 1(a)-(b) shows a brief view of personalized facial expression recognition process in human's mind. On the contrary to the conventional approach(Fig. 1(a)), the proposed 'personalized' approach starts with many mental models for each individual(Fig. 1(b) shows mental models for subject B, C and D). Human can perform facial expression recognition for so many persons without any difficulties using above mental models.

Based on these basic definitions, we derived key properties of 'personalized' classifier as follows;

- Adaptation
  - A personalized classifier can be adapted to input patterns without heavy change of its parameters. A priori knowledge already learned should be maintained or can be slightly changed.
- Feature Selection
  - A personalized classifier has an ability to select minimal features among given whole feature set. Here, feature selection can be done in view of feature space or in view of classifier itself.

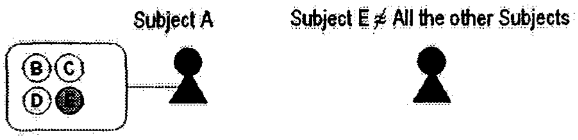
(a) Conventional Facial Expression Recognition



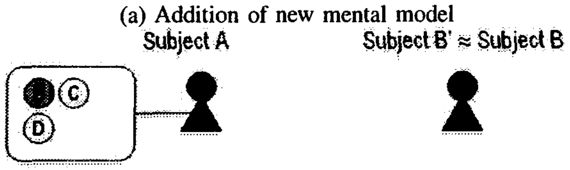
(b) 'Personalized' Facial Expression Recognition

Fig. 1. Conventional Approach vs. 'Personalized' Approach

Above two key properties can be explained by human's thinking way as shown in Fig. 2. When some input patterns are acquired from specific person, it is typical procedure to generate a new classifier based on these input patterns in conventional way. Or the whole classifier is retrained using additionally acquired input patterns plus old data set. On the contrary, as humans do, the personalized way to deal with this situation is to make the classifier as an intelligent agent. That is, human can recognize the other person's facial expressions according to each person's model of facial expressions. When he/she met a new person, corresponding mental model can be generated in one's mind(Fig. 2(a)). After that, the model is used to directly recognize his/her facial expressions or can be slightly changed according to the change of time, life situation and so on(Fig. 2(b)). Thus, we can say that everyone may have a pool of mental model for each individual's set of facial expressions. Besides, as we already explained, everyone performs a kind of adaptation process to keep up with a priori obtained knowledge of the others' facial expressions(Fig.



(a) Addition of new mental model



(b) Adaptation#1

(c) Adaptation#2

Fig. 2. Personalized facial expression recognition process in human's mind

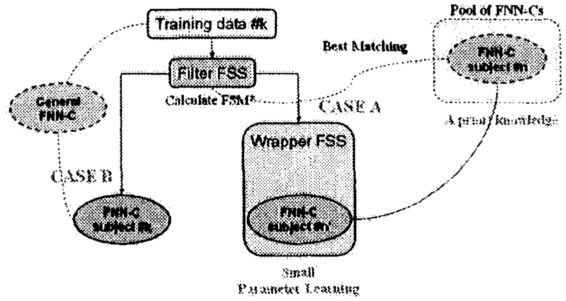


Fig. 3. Proposed personalized classification procedure

2(c)). At last, everyone can minimize (or reduce) the mental model according to the characteristics of each individual.

Thus, in this paper, we propose a novel design procedure for 'personalized' classifier by soft computing techniques. As the underlying classifier, we use a fuzzy neural networks(FNN)-based classifier. The FNN is used due to its capability for implementing effectively the human expert's knowledge-based decision making and learning functions. In fact, we find that FNN can easily handle the knowledge-based findings in the field of psychology such as [1]. Furthermore, due to its transparent inference structure from fuzzy logic, FNN can easily locate causes of error, and implement linguistic rules. Next, as an application of proposed procedure, the facial expression recognition problem is shown with numerical results.

This paper is organized as follows. In Section II, a design procedure of 'personalized' classifier using FNN is mentioned. According to this concept, each necessary process for facial expression recognition is described in following sections. The results of proposed approach are shown in Section III. Finally, concluding remarks are given in Section IV.

II. Design of 'Personalized' Classifier

Based on the human's thinking way, we can set up a classifier(=mental model for facial expressions), a pool of classifier(=set of mental models for different persons) and adaptation process by similarity measurement. Especially, we assume that a mental model can be represented by a FNN-based classifier($FNN-C$). Here, a general $FNN-C$ (fully-connected) is the starting point of all consecutive $FNN-C$s. When a data set(DS)#1 entered the classification system, new $FNN-C$ for subject#1($FNN-C_1$) is generated and it is stored in a pool of $FNN-C$s. After that, when another DS#2 enters the system, a new $FNN-C$ for subject#2($FNN-C_2$) will be generated if the similarity measure between DS#1 and DS#2 is low (case B). Next, $FNN-C_2$ is stored in a pool of FNN-Cs. On the contrary, if the similarity between DS#1 and DS#2 is high (case A), a priori obtained knowledge of subject#1 ($FNN-C_1$) is used to construct a $FNN-C_2$ which is a slightly modified version of $FNN-C_1$. By repeating above process, a personalized machine effectively learns and adapts its structure according to the input patterns of different persons.

To be specific, we will mention the detailed procedure of personalized classifier. Fig. 3 shows a detailed procedure of our personalized classification scheme.

As we already stated, the personalization process consists of two cases: 1) case A is to make an adapted $FNN-C$ in the pool of $FNN-C$s. 2) case B is to make another new $FNN-C$ using given input patterns of specific person. Case A and Case B are determined by checking the similarity between current input patterns and other input patterns in the pool of $FNN-C$s. However, in the real situations, it is almost impossible to store all the input patterns during past learning procedures. So, only two statistical measures are stored with the structure parameters of $FNN-C$ of specific person in the pool of $FNN-C$s. Here, as two statistical measures, we used mean and standard deviation of the distribution of input pattern for specific person.

Next, the similarity between the current input pattern and other input patterns in the pool of $FNN-C$s are measured. When the similarity is high (case A), this pair of input patterns is closely related with each other. Among many possible similar input patterns, we choose the pair of input patterns with the highest similarity value. From this, we can acquire an adapted $FNN-C$ without much effort compared with case B. Additionally, we can make very effective approach not only by choosing the most similar $FNN-C$ from the pool, but also by eliminating redundant input patterns from originally acquired input patterns.

A. Fuzzy Similarity Measure

Using the equation $E(A,B)$ for calculating the fuzzy similarity measure between the fuzzy sets A and B [8], we defined two fuzzy similarity measures based as follows.

$$FSM_1(DS_i, DS_j) = \frac{1}{|\Gamma|}\sum_{fe\in\Gamma} FSM_2(FE_i^{fe}, FE_j^{fe}), \quad (1)$$

$$FSM_2(FE_i^{fe}, FE_j^{fe}) = \frac{1}{N}\sum_{k=1}^{N} E(A_i^{fe}, B_j^{fe}), \quad (2)$$

where Γ is a set of seven facial expressions[1]. DS_i is a data set for whole facial expressions of subject#i in N-dim feature space. $FE_i^{fe} = \{\vec{x}_{i1}, \ldots, \vec{x}_{iN_s}\}$ is an element of DS_i for specific facial expression $fe \in \Gamma$ of subject#i(That is, $DS_i = \{FE_i^{fe}|_{fe\in\Gamma}\}$). N_s is the number of samples of FE_i^{fe}. $\vec{x}_{ij} = [x_{ij1}, \ldots, x_{ijN}]^T \in R^N$ is the j^th feature vector of subject#i. Fuzzy set A_i^{fe} is a bell-shaped membership function with mean $m_{ik}^{fe} = \frac{1}{N_s}\sum_{j=1}^{N_s} x_{ijk}$ and standard deviation $\sigma_{ik}^{fe} = \sqrt{\frac{1}{N_s-1}\sum_{j=1}^{N_s}\left(x_{ijk} - m_{ik}^{fe}\right)^2}$.

FSM_1 measures the similarity between two different data set DS_i and DS_j. On the contrary, FSM_2 measures the similarity between two different elements FE_i^{fe} and FE_j^{fe} for the same facial expression fe.

B Discrimination of case A and case B using FSM

When the data set of subject#i is given, two similarity matrices $M_{FSM_1}(i,j)$ and $M_{FSM_2}^{fe}(i,j)$ are made according to the following equations;

$$M_{FSM_1}(i,j) = FSM_1(DS_i, DS_j), i \le j \quad (3)$$

$$M_{FSM_2}^{fe}(i,j) = FSM_2(FE_i^{fe}, FE_j^{fe}), i \le j \quad (4)$$

Here, in case of $i > j$, we don't need to calculate the similarity measures due to the symmetry of the similarity matrix. Next, for each subject#j, we acquire the most similar subject's ID $MSSID_1(j)$ and $MSSID_2^{fe}(j)$ using following equations.

$$MSSID_1(j) = arg\max_{i<j} M_{FSM_1}(i,j), \quad (5)$$

$$MSSID_2^{fe}(j) = arg\max_{i<j} M_{FSM_2}^{fe}(i,j), \quad (6)$$

$$(7)$$

Using $M_{FSM_1}(MSSID_1(j), j)$, we can categorize whole subjects into two cases(case A and/or case B)[2].

For the case B, we have to train new $FNN-C_j$ with given data set DS_j. However, for the case A, we can minimize the learning process with FSM_2 values not only using a priori knowledge for subject#$MSSID_1(j)$ but also reducing required data set smaller than given data set DS_j. This procedure will be explained in more detail.

To deal with the case A, for each subject#k, we can pick up a set of facial expressions which has $M_{FSM_2}^{fe}(MSSID_2^{fe}(k), k)$ values lower than 0.5. That is, like discriminating case A and case B, we can select sub-data set DSS_k from the original data set DS_k for subject#k. Learning with sub-data set $DSS_k \subseteq DS_k$ (case A) results in very fast and efficient learning than case B.

[1] $\Gamma = \{happy, sad, fear, angry, surprise, disgust, neutral\}$

[2] Here, we used 0.5 as a threshold for discriminating case A and case B.

C. Ensemble of FNN-C: incremental learning of sub-data set

Through the usage of FSM_1 values, current input data set DS_k is categorized as case A and/or case B. In case A, FSM_2 values play a key role to select minimal set of elements DSS_k from DS_k. Thus, for the subject#k, we can reuse $FNN - C_i$ as a basic classifier[3] with additional learning using DSS_k.

Here, the additional learning has to fulfill some requirements as follows[10]:

- It should not forget a prior knowledge.
- It should not require access to the old data.

Most of the conventional incremental learning techniques use a form of 'ensemble' of basic classifier. That is, when an additional data set is available, another new classifier is made and used for classification of the additional data set.

Inspired by above approach, in this paper, we use ensemble of FNN-C to implement an incremental learning scheme. Fig. 4 shows a simple example of our approach with 2-input/2-output FNN-Cs.

In Fig. 4, the left FNN-C is a selected FNN-C from the pool using FSM_1 values and the right FNN-C is a new FNN-C just learned using selected elements DSS_x by FSM_2 values. Instead of merging the outputs of both FNN-Cs, we exclude one of the outputs from the left FNN-C. That is, for the class#2 in Fig. 4, the output value from the right FNN-C is used instead of the output value from the left FNN-C. 2-inputs are commonly used for both FNN-Cs.

III. Application to Personalized Facial Expression Recognition

With predefined five image features[4] [7] and seven output classes, we construct a FNN-based classifier($N_i = 5$ and $N_o = 7$). Here, output classes are defined as six universal facial expressions [1] plus one neutral facial expression (see Fig. 5).

For the training/test of a FNN-based classifier, we make a data set from Ekman's Facial Expression DB which consists of

[3] Here, we assume that $FSM_1(DS_i, DS_k)$ is higher than 0.5.

[4] Here, five image features are degree of eye openness, degree of nasolabial root, degree of nasolabial fold, degree of mouth openness and vertical distance between eyebrows and eyes.

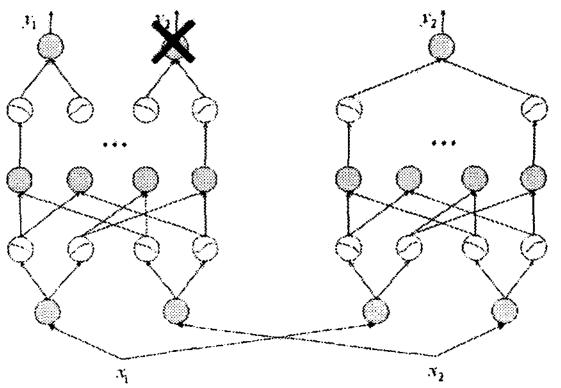


Fig. 4. Ensemble of FNN-C: incremental learning scheme

TABLE I

SELECTED SUB-DATA SET FOR 8 SUBJECTS (CASE A)

Subject ID	Most similar subject's ID	Sub-data set
15	3	Fear
16	6	Happy
17	8	Fear, Angry
18	9	Fear, Angry
19	11	Happy, Sad, Fear
20	11 (or 20)	Happy, Sad, Fear
21	12	Happy, Sad
22	14	Happy, Sad, Angry, Disgust

94 facial photos from 14 different persons [1] plus 56 facial photos from 8 additional persons. According to the feature extraction process in [7], five image features are extracted from each facial image of 150 facial images. Thus, the data set(DATASET#1) consists of 150 image feature vectors[5]. From the DATASET#1, the other data set(DATASET#2) is derived with statistical variation of each feature vector. For each feature vector $\vec{x} = [x_1, \ldots, x_{N_i}]^T \in R^{N_i}$ in the DATASET#1, 100 feature vectors $\vec{z} = [z_1, \ldots, z_{N_i}]^T \in R^{N_i}$ of the DATASET#2 are generated with simple equation such as $z_i = x_i + 2\sigma\left(e^{-(x_i-0.5)^2} - 0.5\right), 1 \leq i \leq N_i$. Thus, the whole feature vectors of the DATASET#2 is 15000.

According to the observations for repetitive trials of making facial expressions with the same/different person, camera noise modelling and so on, $\sigma = 0.05$ is chosen to make the DATASET#2.

As we already explained, the DATASET#2 is divided into two cases: case A(subjects#15,#16,#17,#18,#19,#20,#21 and #22) and case B(subjects#2,#3,#4,#6,#8,#9,#10,#11,#12,#13 and #14).

For 8 subjects in case A, we can select sub-data set using FSM_2 values as shown in Table I.

Thus, we can make additional 8 FNN-Cs using selected sub-data set in Table I. Finally, we apply the proposed strategy for 8 subjects and get the results as Table II.

The cross-validation is used with the same number of training vectors and test vectors. As it is shown in the Table II, for 8 subjects in case A, the classification rate is remarkably enhanced from 62.5% to 83.5% using proposed strategy.

Finally, in view of whole 22 subjects, classification rate is enhanced from 82.8% to 90.4% using proposed strategy.

The main advantages of proposed strategy can be summarized as follows:

[5]Each feature vector contains five image features defined in [7].

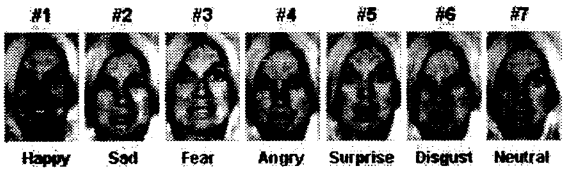


Fig. 5. An example set of facial images with seven facial expressions [1]

- A reasonable numerical result is obtained for 22 subjects which cannot be achieved using conventional learning scheme.
- A priori knowledge (in the pool of FNN-Cs) is preserved.
- Additional FNN-C is easily trained using only selected elements DSS_x by FSM_2 values.
- Reusability of FNN-C in the pool of FNN-Cs is increased.
- On-line learning is possible in human's thinking way.

IV. CONCLUDING REMARKS

In this paper, a design problem of personalized classifier is mentioned using fuzzy similarity measure(FSM)-based strategy. A preliminary result for proposed scheme is given with the facial expression recognition. As shown in Section III, proposed method is very effective which shows a reusability, on-line learning capability, easy-to-expand and so on. As a further work, the ensemble structure might be refined with another learning techniques.

REFERENCES

[1] P. Ekman et al. "Facial Action Coding System", 1978.

[2] K. Chung et al., "Performance Comparison of Several Feature Selection Methods Based on Node Pruning in Handwritten Character Recognition," *Proceedings of ICDAR'99*, pp. 11–15, 1997.

[3] V. Onnia et al., "Feature Selection method using neural networks," *Proceedings of ICIP2001*, pp. 513–516, 2001.

[4] D. Chakraborty et al., "Designing rule-based classifiers with on-line feature selection: a neuro-fuzzy approach," *Proceedings of AFSS2002*, pp. 251–259, 2002.

[5] J.-S. Han, W.-C. Bang and Z. Zenn Bien, "New Feature Set Selection Algorithm based on Soft Computing Techniques and its Application to EMG Pattern Classification," Journal of Fuzzy Optimization and Decision Making, Vol. 8, No. 3, pp. 1–30, 2002.

[6] M.-J. Kim, J.-S. Han, K.-H. Park, W.-C. Bang and Z. Zenn Bien, "Classification of Arrhythmia Based on Discrete Wavelet Transform and Rough Set Theory," ICCAS 2001, October 17-21, 2001.

[7] Kim, D.-J., Bien, Z. 2003. A Personalized Facial Expression Recognition System using Fuzzy Neural Networks with A Novel Feature Selection Scheme. Proceedings of FUZZ-IEEE 2003, 908–913.

[8] C.T. Lin "Neural Fuzzy Control Systems with Structure and Parameter Learning", *World Scientific*, 1994.

[9] Collins Cobuild English-English Dictionary, 2002.

[10] R. Polikar et al., "Learn++: An Incremental Learning Algorithm for Supervised Neural Networks," *IEEE Tr. on SMC: Part C*, vol. 31, no. 4, 2001.

TABLE II

PERFORMANCE ENHANCEMENT OF 8 SUBJECTS WITH PROPOSED STRATEGY

Subject ID	Classification rate (%)	
	w/o learning	w/ learning
15	85.7	100.0
16	83.7	96.9
17	78.0	84.6
18	65.1	75.1
19	54.0	84.9
20	58.0	84.9
21	66.3	80.9
22	14.6	60.6
Total	62.5	83.5