Facial Emotional Expression Recognition with Soft Computing Techniques

Zeungnam Bien, Dae-Jin Kim and Sang-Wan Lee
Dept. of EECS, KAIST
373-1 Guseong-dong, Yuseong-gu, Daejeon 305-701, Korea
Tel./Fax.: +82-42-869-3419/8410, E-mail: zbien@ee.kaist.ac.kr

Abstract— **The facial expression recognition (FER) is one of the biosignal-based recognition techniques which attract a lot of attention recently. To deal with its complex characteristics effectively, we adopt the soft computing techniques (SCT) such as fuzzy logic, neural networks, genetic algorithm and/or rough set technique. In this paper, we overview the state-of-the-art reports on FER in view of SCT, and introduce some interesting works done by our group on the SCT-based facial emotional expression recognition. Specifically, 1) Fuzzy observer-based approach, 2) personalized FER system based on fuzzy neural networks, and 3) Gabor wavelet neural networks are briefly discussed.**

Keywords— **facial expression recognition, soft computing techniques, human-robot interaction.**

I. Introduction

Recognition and understanding of human facial expressions can be useful in human-friendly man-machine interface of the modern human-in-the-loop systems [1][2]. As in inter-human communication and nonverbal communication for the hearing-impaired [3], the human facial expression recognition becomes essential in advanced automation systems such as service robots for which man-machine interaction is an integral part of the system. Typically, facial 'emotional' expressions mean the psychological work mentioned by Dr. Ekman et al.'s group (6 universal emotions; happy, sad, fear, angry, surprise and disgust).

In general, the problem of recognizing emotion from face is known to be very complex and difficult because individuality may not be ignored in expressing and observing emotions. In other words, various visual features, caused directly by anatomical structure such as bone structure, muscle structure and the tissue system [4], are so interrelated, and also the facial expressions are affected by many nonvisual factors such as experience, familiarity and attention [5].

The problem of recognizing facial expressions has recently been handled by various intelligent techniques [6][7][8], including artificial neural networks [2], fuzzy modeling approach using manually extracted distance-based features [8], Action Unit (AU) based approach using computer graphics model [6], and motion based approach using optical flow [7]. But there remain a number of difficulties or limitations in realizing a practical system that is capable of extracting reliable features in real time for easy analysis. Furthermore, because of the subjectivity characteristic of facial expressions, the features of faces can be very vague and the recognition results can be poor. In fact, many analysis-based approaches have been criticized in view of reliability. To enhance success rate, the approaches using SCT have become an alternative.

This paper is organized as follows. In Section II, some representative approaches for FER with SCT are revisited. Next, our group's research works on FER with SCT are shown with the detailed processes and numerical results in Section III. Finally, summary and future prospects are mentioned in Section IV.

II. Facial Expression Recognition with Soft Computing Techniques: A Survey

Among many SCT, artificial neural networks are widely used due to its learning capability. In [16], 2D image itself is used as the input of neural network with 40 hidden neurons. Because the whole image information is used to classify, it needs lots of hidden neurons and does not show satisfactory performance. In [17], the system extracts two kinds of features, geometric positions and Gabor wavelet coefficients. The combination of these two features becomes the input of the two-layer neural network. The performance of the system records success rate of 90.1% for 7 facial expressions. In [18][19], a neural network-based system was used to classify facial action units. Here, separate network structures are used for upper part's recognition [18] and lower part's recognition [19], respectively. With more exquisitely-defined features than the system in [16][17], each network shows performance of 93.7% and 96.71% for CMU DB, respectively. Because they deal with only facial action units, however, the satisfactory performance in FER may not be guaranteed.

For recognizing more complex facial emotional expressions, a modular-type expert system was proposed in [20] in which the system is organized by independent classifiers. Action units extracted from all modules are mixed to describe various facial emotional expressions. The success rate is 86.3% for recognitoin of six facial expressions.

In [21], both Multi-Layer Perceptron(MLP) and Radial Basis Function Network(RBFN) are used for FER. The recognition rate is 73% for MLP, and 65% for RBFN. In this case, the inputs of each network are 395 feature vectors, consisting of geometric positions, euclidean distance of contour points, the differences in inter-feature distances and Hu-moment invariants. It shows that too many features can degrade the performance of recognition.

In [22], a more progressive neural network structure, called Convolutional Neural Network (CNN), was used to classify facial expressions. The learned weights of convolutional layers enable the system to extract features relevant to a given facial analysis task, whereas the sub-sampling layers provide robustness against facial location changes and scale variations. Therefore, the employed CNN structure recognizes facial expressions in the presence of pose

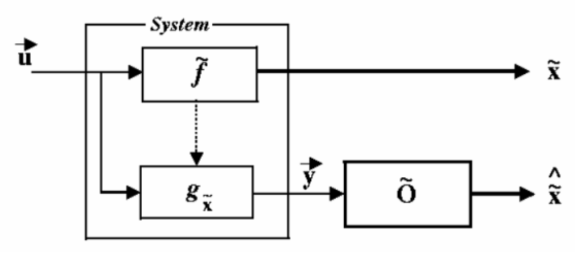


Fig. 1. System model for the fuzzy observer

variations without requiring extensive pose normalization or feature tracking initialization procedure. It shows performance of 91.7% for JAFFE DB. In the other paper [23], CNN is only used for face detection task, and then rule-based processing is performed with local features detected by CNN. Even though classification result is as high as 97.6%, only 'smile' expression is considered.

Support Vector Machine (SVM) is also used to classify facial expressions. In [24], the pupil is detected by infrared camera using red-eye effect. Next, PCA is used to extract features. Finally, SVM classify action units in terms of facial components. 81.22% recognition rate for CMU DB is recorded in this fully automated system.

Comparing with other non-SCT, SCT enable us to deal with various kinds of uncertainty and complex aspects of facial expression recognition. However, we found that the previous works shorts of following aspects;

1. The previous works didn't effectively use the psychological results given by Ekman et al. [4].

2. The previous works can't deal with the individual characteristics for each subject.

In our group, we have been tried to overcome above shortcomings of the previous works. According to our experiences, SCT plays a key role to handle these shortcomings.

III. Examples of Facial Expression Recognition with Soft Computing Techniques

A. Fuzzy Observer-based Facial Expression Recognition

In this work, SCT are used for recognizing a positive expression of happiness. Specifically, wrinkle in the nasolabial fold (NLF) region is adopted as feature, which is observed to be one of major factors in human determination of the intensity of happiness [4]. To handle the recognition system by employing a traditional fuzzy logic framework, a novel concept termed as "fuzzy observer" is proposed to indirectly estimate a linguistic variable from conventionally measured data, along with image features based on slice DFT (Discrete Fourier Transform) [9] to automatically extract wrinkle features in the NLF region.

Fig. 1 shows a fuzzy observer which are intended to replace the human decision maker and human observer. Suppose that a face with some emotion is presented as an input u. Then, a human observer gives the linguistic value such as *high* for a linguistic variable $\widetilde{\mathbf{x}}$ such as wrinkledness of a face, while the designed fuzzy observer provides a corresponding approximated fuzzy set. For example, the fuzzy set is described by some triangular-shaped membership function. Humans make judgments on the facial expressions using human knowledge described as linguistic values; for example, when the wrinkles in the NLF region or crow's feet wrinkles in a human face are intensified, human beings recognize the face as a happy one [4]. In the facial analysis, the deepened wrinkles in the nasolabial fold area are used as a visual cue, though many people may have permanent wrinkles. The fuzzy subsystem performs fuzzy approximate reasoning based on a fuzzy rule set for a happy face.

A.1 Design of fuzzy observer

Specifically, let $\mathcal{R}$ denote the set of all real numbers. Let $\widetilde{\mathcal{P}}(X)$ be the set of all fuzzy sets in $X \subset \mathcal{R}$. Let $\widetilde{\mathbf{f}}$ be a fuzzifying function from $U \subset \mathcal{R}^m$ to $\widetilde{\mathcal{P}}(X)$ such that for any $\vec{\mathbf{u}} \in U$, $\widetilde{\mathbf{x}} = \widetilde{\mathbf{f}}(\vec{\mathbf{u}})$ is a fuzzy number[10]. Let $\widetilde{\mathbf{x}}(x)$ denote a membership function value of $\widetilde{\mathbf{x}}$ for $x \in X$. For a given $\widetilde{\mathbf{x}}$, let $\mathbf{g}_{\widetilde{\mathbf{x}}}$ denote a crisp vector-valued function from U to Y, where $Y \subset \mathcal{R}^n$ and write $\vec{\mathbf{y}} = \mathbf{g}_{\widetilde{\mathbf{x}}}(\vec{\mathbf{u}})$ for $\vec{\mathbf{y}} \in Y$. Here, $\vec{\mathbf{u}}$ and $\vec{\mathbf{y}}$ are a system input vector and a measurement system output vector, respectively. The function $\mathbf{g}_{\widetilde{\mathbf{x}}}$ is introduced to represent an indirect quantity measurement system such as a feature extractor. Let a fuzzy observer $\widetilde{O}$ be defined as a fuzzifying function from Y to $\widetilde{\mathcal{P}}(X)$, and denote the relation by $\widehat{\widetilde{\mathbf{x}}} = \widetilde{O}(\vec{\mathbf{y}})$, where $\widehat{\widetilde{\mathbf{x}}}(\in \widetilde{\mathcal{P}}(X))$ is the estimated fuzzy number of $\widetilde{\mathbf{x}}$ (see Fig. 1).

We assume that $\widetilde{\mathbf{x}}$ is *observable* from $\vec{\mathbf{y}}$ in the sense that the output $\vec{\mathbf{y}}$ contains information about the linguistic values of the variable $\widetilde{\mathbf{x}}$. Then the problem is to design a fuzzy observer so that the output $\widehat{\widetilde{\mathbf{x}}}$ of the fuzzy observer can approximate $\widetilde{\mathbf{x}}$ for any $\vec{\mathbf{u}} \in U$. Our approach for the problem is to incorporate some learning capability in the observer system so that its performance of estimation is reasonably satisfactory for wide range of input space U.

In order to make the proposed fuzzy observer trainable, the fuzzy observer is represented as a fuzzifying function in which a triangular-shape fuzzy number is generated with two parameters, i.e., its maximum possibility center m and its uncertainty width $\sigma(\sigma > 0,\ m \in \mathcal{R})$, as shown in Fig. 2. More specifically, let

$$\widehat{\widetilde{\mathbf{x}}}(x) = \Lambda(m, \sigma; x), \tag{1}$$

$$\Lambda(m, \sigma; x) \equiv \begin{cases} 1 - \frac{|m-x|}{\sigma} & \text{if } |m-x| \le \sigma \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Suppose that we are given N input–output pairs ($\vec{\mathbf{y}}^p$, $\widetilde{\mathbf{x}}^p(x)$), $p = 1, \cdots$, N. Here, $\vec{\mathbf{y}}^p$ $(\in Y)$ is a crisp measurement vector and $\widetilde{\mathbf{x}}^p(x)(\in \widetilde{\mathcal{P}}(X))$ is a fuzzy number. Then, the problem for training is stated as follows :
Given a sample set $\{(\vec{\mathbf{y}}^p, \widetilde{\mathbf{x}}^p(x)), p = 1, \cdots, \mathrm{N}\}$, determine the fuzzy observer so that the cost function $J^p(W)$ defined by

$$J^p(W) = \alpha_1 J_1^p(W) + \alpha_2 J_2^p(W) \tag{3}$$

is minimized in the weight space W. Here α_1, $\alpha_2 > 0$, and

$$J_1^p(W) = \frac{1}{2}\int_{x \in X} |\widehat{\widetilde{\mathbf{x}}}(x; W) - \widetilde{\mathbf{x}}^p(x)|^2 dx \tag{4}$$

$$J_2^p(W) = \frac{1}{2}(m(W) - m^p)^2 \tag{5}$$

$$(m^p \quad : \quad \text{center of area of } \widetilde{\mathbf{x}}^p)$$

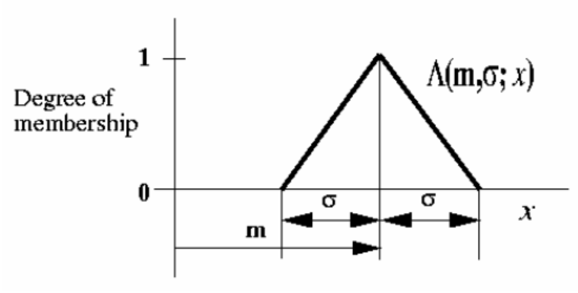


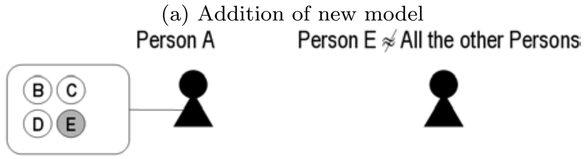
Fig. 2. Definition of triangular-shape membership function

(a) Conventional Facial Expression Recognition

(b) 'Personalized' Facial Expression Recognition

Fig. 3. Conventional Approach vs. 'Personalized' Approach

(a) Addition of new model



(b) Modification of old model#1

(c) Modification of old model#2

(d) Feature Selection via Personality

Fig. 4. Personalized facial expression recognition process in human's mind

In [1], the detail explanation of the learning can be found.

A.2 Experiment with facial images

Recall that humans can identify and distinguish various facial expressions of emotions including surprise, fear, anger, disgust, sadness and happiness, which are termed as six universal expressions[5]. In this work, we concentrate on the facial expression of emotion "happiness." The experiment is performed for a happy face image sequence, which constitutes a neutral face into happy ones and then very happy ones and again neutral to happy ones.

Finally, the 100 % (=15/15) success rate is obtained in detecting automatic facial parts and NLF regions.

B. *Personalized Facial Expression Recognition: Fuzzy Similarity Measure and Novel Feature Selection*

Simply speaking, 'personalized' classification is considered as 'customized' (or 'custom-tailored') classification. For example, Fig. 3(a),(b) shows a conceptual difference between the conventional approach and the personalized approach. On the contrary to the conventional approach(In Fig. 3(a), ∀ means that each expression for whole people is modelled in advance.), the 'personalized' approach starts with many mental models for each individual(That is, in Fig. 3(b), B and C mean that each expression for each person is personally modelled according to the individualities). Using a pool of mental models, person A can perform FER for other people.

In Fig. 4, all possible cases of FER in a human's way of thinking is depicted. Basically, the whole cases can be categorized into three steps such as: 1) Addition of new mental model(Fig. 4(a)), 2) Modification of old mental model(Fig. 4(b),(c)), and 3) Feature selection according to each person's individuality(Fig. 4(d)). When person A meet new person E(Fig. 4(a)), after checking the similarity between the facial expressions of person E and other persons in the pool of mental model, new mental model for person E is added due to its big difference with other persons. Modification of old mental model is caused by two cases; 1) Even though the person B is already in mental model of person A, his/her facial expressions are slightly changed due to time, life situation and so on. 2) In some cases, the facial expressions of person F may resemble those of a person in the pool (In Fig. 4(c), person D is the most similar person with person F.) Finally, a feature selection process is applied for each person in the pool to enhance his/her individualities(As shown in Fig. 4(d), person B and person C have very different facial characteristics).

Among many possible classification methods, especially, we use Lin's fuzzy neural networks(FNN)-based classifier [12]. The FNN is used due to its capability for effective implementation of human expertise-based decision making and learning by well-known techniques. In fact, we found that FNN can easily handle the psychological expertise for facial expressions as shown in [4]. Furthermore, due to its direct inference structure, FNN can easily locate causes of errors, input/output relationship and so on.

Before explaining about our procedure for making a personalized classifier, the symbols used and their meanings are denoted in Table I. [1]

[1]Here, Γ is the set for seven facial expressions; happy, sad, fear, angry, surprise, disgust and neutral.

TABLE I
NOMENCLATURE

Symbol	Meaning
$FNNC_i$	FNN-based classifier for person i
$\vec{x}_{ij}^{fe}$	the j^{th} feature vector of person i for specific facial expression $fe \in \Gamma$ $\vec{x}_{ij}^{fe} = [x_{ij1}^{fe}, \ldots, x_{ijI}^{fe}]^T \in \mathcal{R}^I$
FE_i^{fe}	a sub-data set of person i for specific facial expression $fe \in \Gamma$ $FE_i^{fe} = \{\vec{x}_{i1}^{fe}, \ldots, \vec{x}_{iS}^{fe}\}$
DS_i	a data set for whole facial expressions of person i $DS_i = \{FE_i^{fe}\|_{fe\in\Gamma}\}$
$A_i^{fe}(m_{ik}^{fe}, \sigma_{ik}^{fe})$	a bell-shaped fuzzy set used for modelling FE_i^{fe} $m_{ik}^{fe} = \frac{1}{S}\sum_{j=1}^{S} x_{ijk}^{fe}$ $\sigma_{ik}^{fe} = \sqrt{\frac{1}{S-1}\sum_{j=1}^{S}\left(x_{ijk}^{fe} - m_{ik}^{fe}\right)^2}$

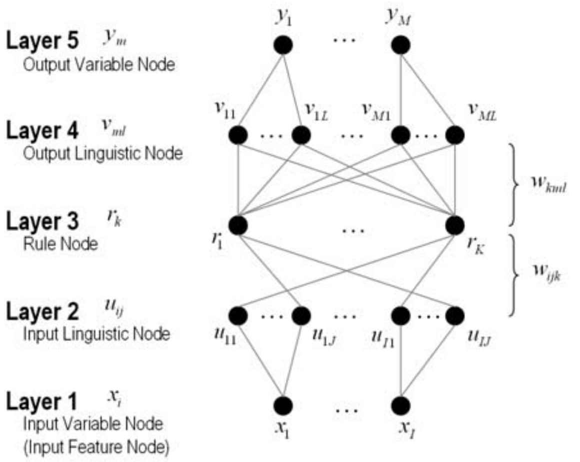


Fig. 5. A Lin's 5-layered FNN-structure [12]

B.1 Personalized Classifier: Model Selection

At first, according to the observations in Fig. 4, we have to select appropriate models for further process. To discriminate 'addition of new model (ANM)' and 'modification of old model (MOM)', we need a similarity measure. Using a defined similarity measure, a decision maker can be made. Especially, in case of MOM, a special method should be used to preserve a priori knowledge of old model.

Using the equation $E(A, B)$ for calculating the fuzzy similarity measure between the fuzzy sets A and B [12], we defined two fuzzy similarity measures as follows;

$$FSM_1(DS_i, DS_j) = \frac{1}{|\Gamma|}\sum_{fe\in\Gamma} FSM_2(FE_i^{fe}, FE_j^{fe}) \quad (6)$$

$$FSM_2(FE_i^{fe}, FE_j^{fe}) = \frac{1}{I}\sum_{k=1}^{I} E(A_i^{fe}, B_j^{fe}) \quad (7)$$

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.

Using FSM_1, we can categorize whole persons into two cases(Model MOM and/or Model ANM). For the Model ANM, we have to train new $FNNC_j$ with given data set DS_j. However, for the Model MOM, we can minimize the learning process with FSM_2 values not only using a priori knowledge but also reducing required data set $DSS_j \subseteq DS_j$. To deal with the Model MOM, for each person j, we can pick up a set of facial expressions which has FSM_2 values lower than 0.5. That is, like discriminating Model MOM and Model ANM, we can select sub-data set DSS_j from the original data set DS_j for person k. Learning with sub-data set $DSS_j \subseteq DS_j$ (Model MOM) enables a very fast and efficient learning than learning with DS_j itself (Model ANM).

It is well known that 'ensemble' of classifiers is an effective solution for incremental learning[13]. That is, when a new(unlearned) data set is available, a small-sized classifier is made and used for classification of the additional data set. Inspired by above properties, we use ensemble of $FNNC$ to implement an incremental learning.

B.2 Personalized Classifier: Feature Selection

The last decade witnesses various methods on feature selection to handle complex pattern recognition problems (e.g. see [14], [15]). One of the recent versions is related to the FNN-based approach, which is capable of handling the human expertise [14]. In this paper, we rely on the fact that FNN can represent some human expertise in the forms input/output relationships of the networks.

A feature selection problem can be formulated as follows;

Given $\vec{p}$**,** $X = \{\vec{x}_s^m \in \mathcal{R}^I | m = 1, \ldots, M, s = 1, \ldots, S\}$ **and** $Y = \{\vec{y}_s^m \in \mathcal{R}^M | m = 1, \ldots, M, s = 1, \ldots, S\}$**,**
Find $T(\vec{p}) \in \mathcal{R}^{O\times I}$ **s.t.** $E' < E$**.**
where, $E = \frac{1}{MS}\sum_{s=1}^{S}\sum_{m=1}^{M} ||\vec{y}_s^m - C(\vec{x}_s^m, \vec{p})||$, $E' = \frac{1}{MS}\sum_{s=1}^{S}\sum_{m=1}^{M} ||\vec{y}_s^m - C(\vec{x'}_s^m, \vec{p}')||$.

Here, $\vec{x'}_s^m (= T(\vec{p})\cdot\vec{x}_s^m \in \mathcal{R}^O, O < I)$ is a reduced feature vector of $\vec{x}_s^m$.

To find appropriate $T(\vec{p})$ in the feature selection problem, a connection degree and a histogram of a connection degree are devised for denoting the input/output relationships. Fig. 5 shows a Lin's 5-layered FNN structure [12].

A connection degree between an input linguistic node u_{ij} and an output linguistic node v_{ml} is defined as follows;

$$F(u_{ij}, v_{ml}) = \frac{1}{J^{I-1}}\sum_{k=1}^{K} h(u_{ij}, v_{ml}, r_k) \quad (8)$$

where, $h(u_{ij}, v_{ml}, r_k) = w_{ijm} \times w_{mkl}$ is used and the value of w_{ijk} and w_{kml} is a binary value (0 or 1) according to the existence of the link.

Now, for calculating the histogram value with respect to each input linguistic node, two functions are defined as follows;

$$G(i, j) = \frac{1}{M}\left|\sum_{m=1}^{M}\sum_{l=1}^{L} g(F(u_{ij}, v_{ml}))F(u_{ij}, v_{ml})\right| \quad (9)$$

$$H(i,j,m) = \left| \sum_{l=1}^{L} g(F(u_{ij}, v_{ml}))F(u_{ij}, v_{ml}) \right| \quad (10)$$

where, $g(F(u_{ij}, v_{ml})) = 2\frac{l-1}{L-1} - 1$.

In Eqn. (10), $H(i,j,m)$ is used to represent the effect of an input linguistic node u_{ij} on an output node y_k. That is, if $H(i,j,m)$ is high, the node u_{ij} has a higher importance than other nodes with low $H(i,j,m)$ values. Similarly, as it is shown in Eqn. (9), $G(i,j)$ is used to represent the effect of an input linguistic node u_{ij} on the whole output.

In [25], we setup a procedure to select important features based on two histogram values $G(i,j)$ and $H(i,j,m)$.

B.3 Application to Personalized Facial Expression Recognition

With predefined five image features[2] 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 [4] plus one neutral facial expression.

For the training/test of a FNN-based classifier, we make a data set from Ekman's Facial Expression DB which consists of 94 facial photos from 14 different persons [4] plus 56 facial photos from 8 additional persons.

According to the proposed 'Model Selection' process, EXP-DS#2 is divided into two categories: Model MOM(person 15-22) and Model ANM(person 1-14). Thus, we can make 8 additional $FNNC$s using selected sub-data set. In fact, this is more efficient than making a new $FNNC$ as Model ANM does.

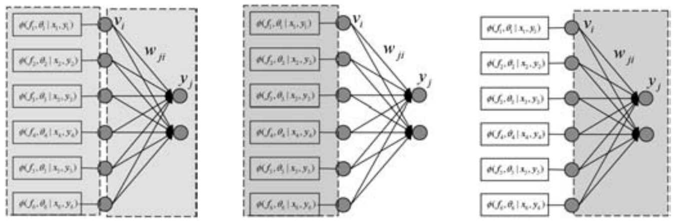
In view of whole 22 persons, classification rate is enhanced from 81.0% to 90.4% using proposed strategy.

C. Adaptation: Unsupervised Fuzzy Neural Network in Facial Expression Recognition using Gabor-Wavelet Neural Network

In this work, we found that the combination of supervised learning, reinforcement learning and unsupervised learning can enable us to model the human's learning way. In every cases, SCT also play a key role.

Fig. 6 shows overall process integrating the learning of Gabor-Wavelet Neural Network (GWNN) and adaptation phases. In the first stage (see Fig. 6(a)), initial learning is carried out based on back-propagation algorithm. The system learns both Gabor filters' parameter and weights. This stage is for finding global solution and assembling network itself, and basically performs supervised learning in off-line way. An adaptation process is activated only when a new (unlearned) user starts to use the system. In this stage (see Fig. 6(b)), the system tunes Gabor nodes' parameters to extract reliable features for the new user's facial expressions. After tuning of Gabor nodes, as shown in Fig. 6(c), the system's weights are modified by using fuzzy neural network model [27]. During organizing clusters, it uses the same learning data already used in Fig. 6(b). This adaptation process is completed when the error converges into some error bound. In this process, the on-line adaptation is performed.

[2]degree of eye openness, degree of nasolabial root, degree of nasolabial fold, degree of mouth openness and vertical distance between eyebrows and eyes.



(a) Initial Learning (b) Adaptation (I) (c) Adaptation (II)

Fig. 6. Overall process of Gabor Wavelet Neural Network

C.1 Initial Learning: Gabor Wavelet Neural Network

Proposed system uses a reduced Gabor-Wavelet Neural Network (GWNN) which contains both Gabor filters and a two layer neural network. The first stage is a feature extracting layer which has the set of Gabor filters, and the second stage is a recognizer which decides a decision boundary in feature space. Fig. 6 shows the structure of the system. When the input image is denoted by I, I_i can be a partial input image which is the same size as i-th Gabor Wavelet. Each Gabor wavelet nodes are applied to 6-facial points $(cx_i, cy_i),\ i = 1, 2, \ldots, 6$, defined as facial definition parameters (FDP) in MPEG4 standard. Node i's outputs in first stage and node j's outputs in second stage can be calculated as follows.

$$G_i^{even/odd} = I_i \otimes g_i^{even/odd} = \sum_{k=1}^{K_i}\sum_{l=1}^{L_i} I_i(K_i - k, L_i - l) g_i^{even/odd}(k,l) \quad (11)$$

$$y_j = \bar{w_j} \cdot \bar{v} = \sum_{i=1}^{6} w_{ij} y_i = \sum_{i=1}^{6} w_{ij} \frac{\sqrt{|G_i^{even}|^2 + |G_i^{odd}|^2}}{G_i^{MAX}} \quad (12)$$

In above equation, Gabor filters' outputs are divided by estimated maximum values of each filter, so each node's output is normalized in fuzzy manner. Therefore, we can say that the each output of the first stage is a fuzzification of Gabor filter.

The detailed part of learning method for GWNN is mentioned in [26].

C.2 Adaptation Process: Q-Learning and Unsupervised Fuzzy Neural Network

After finishing initial learning stage, we use Q-learning [11] for adaptation when a new user starts to use the system (see Fig. 6(b)). This stage is activated only when the new user's data are presented. Whenever user's data are presented, state transition happens. Here, each state is a parameter set (f, θ), and state transition can be done by total 9 directions including 'hold' transition that means no parameter change.

Besides of Q-Learning, Unsupervised Fuzzy Neural Network (UFNN) is used for another adaptation process as follows (see Fig. 6(c)). At first, Gabor nodes' outputs are applied to unsupervised neural network. After input pattern is applied, as the second process, competition among output neurons occurs in a winner-take-all fashion.

$$winning\ neuron = \arg\min_j ||v - w_j|| \qquad (13)$$

where v is the output vector of Gabor nodes which is identical to the input vector of this algorithm, w_j is weight vector related to the j-th output node. As the third process, this algorithm, performs the vigilance test according to the vigilance criterion defined as fuzzy manner.

$$e^{-\gamma u_i} ||x - v_i|| \leq T \qquad (14)$$

where, $u_i = \left(\frac{1}{||x-v_i||^2} \right)^{\frac{1}{m-1}} / \sum_{j=1}^{n} \left(\frac{1}{||x-v_i||^2} \right)^{\frac{1}{m-1}}$

If the winning cluster satisfies the vigilance criterion, the centroids of all clusters are updated regardless of winning or losing. However, if it fails the vigilance test, the input data is registered as the weight of a new cluster.

The test result shows the superiority of adaptation process. Test data are cropped from two individuals with PC-CAM. User#1 DB contains 53 happy images and 61 sad images. User#2 DB contains 60 happy images and 60 sad images. For User#1 DB, we got 100% with initial learned system. Thus, adaptation process is not necessary in this case. For User#2 DB, proposed method enhances the recognition result from 87,5% to 93.35%. By using adaptation process with UFNN, the recognition capability is increased. After this process, unsupervised network is organized, and makes the network adapting the new user continuously.

IV. Concluding Remarks

Due to its complexity, FER is not an easy task for the machine. In the last decade, many FER methods with SCT emerged to show the effectiveness of SCT in FER task.

In this paper, we discussed about our group's three works on FER with SCT. From the numerical results, these works show its excellence well. Our works include the psychological works in effective way and adapt to individual according to each person's characteristics.

In these days, we study the evolutionary adaption process of facial expressions in human's life time and try to implement it in portable electronics such as PDA, tablet PC and so on. We hope that this work can be widely usable in many intelligent/human-friendly human-robot interaction system.

Acknowledgments

This work is supported by 21C Frontier Project, "Development of Intelligent Robot Technology for Assisting Human Functions" program and Human-friendly Welfare Robot System Engineering Research Center, KAIST.

References

[1] G.-T. Park and Z. Bien, "Neural network-based fuzzy observer with application to facial analysis," *Pattern Recognition Letters*, vol. 21, pp. 93–105, Feb. 2000.

[2] F. Hara and H. Kobayashi, "Use of face robot for human-computer communication," *Proc. of IEEE Conf. on SMC'95*, pp. 1515–1520, 1995.

[3] T.H. Haynes, *An Investigation of The Recognition of Facial Expression of Emotion After Training of Deaf Students at a Residential School,* PhD thesis, The Florida State University, 1980.

[4] P. Ekman and W. Friesen, *Unmasking the Face,* Prentice-Hall, Inc., 1975.

[5] V. Bruce, *Recognising Faces,* Lawrence Erlbaum Associates, Publishers, 1988.

[6] I. A. Essa and A. P. Pentland, "Coding analysis, interpretation, and recognition of facial expressions," *IEEE Tr. on Pattern Analysis Machine Intelligence*, vol. 19, no. 7, Jul. 1997.

[7] M. Rosenblum, Y. Yacoob and L.S. Davis, "Human expression recognition from motion using a radial basis function network architecture," *IEEE Transactions on Neural Network* 7(5):1121–1138., 1996.

[8] A. Ralescu and R. Hartani, "Fuzzy modeling based approach to facial expressions understanding," *Journal of Advanced Computational Intelligence*, vol. 1, no. 1, pp. 45–61, Oct. 1997.

[9] D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing,* Prentice-Hall, Inc., 1984.

[10] D. Dubois and H. Prade, *Fuzzy Sets and Systems:Theory and Applications,* Academic Press, Inc., 1980.

[11] S. Haykin, *Neural Network–A Comprehensive Foundation,* Macmillan College Publishing Company, Inc., 1994.

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

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

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

[15] J.-S. Han et al., "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.

[16] C.L. Lisetti and D.E. Rumelhart, "Facial expression recognition using a neural network," *Proc. of FLAIRS Conference*, 1998.

[17] Z. Zhang, M. Lyons, M. Schuster and S. Akamatsu, "Comparison Between Geometry-Based and Gabor-Wavelets-Based Facial Expression Recognition Using Multi-Layer Perceptron," *Proceedings of FG'98*, pp. 454–459, 1998.

[18] Y. Tian, T. Kanade and J.F.Cohn, "Recognizing Upper Face Action Units for Facial Expression Analysis," *Proceedings of CVPR'00*, 2000.

[19] Y. Tian, T. Kanade and J.F.Cohn, "Recognizing Lower Face Action Units for Facial Expression Analysis," *Proceedings of FG'00*, 2000.

[20] M. Pantic, L.J.M. Rothkrantz, "Expert system for automatic analysis of facial expressions," *Image and Vision Computing 18*, pp. 881–905, 2000.

[21] M.Gargesha and P.Kuchi, "Facial expression recognition using a neural network," *Artificail Neural Computation Systems*, 2002.

[22] B. Fasel, "Robust Facial Analysis using Convolutional Neural Networks," *Proc. of ICPR'02*, 2002.

[23] M. Matsugu, K. Mori, Y. Mitari and Y. Keneda., "Facial Expression Recognition Combined with Robust Face Detection in A Convolutional Neural Network," *Proc. of IJCNN'03*, pp. 2243–2246, 2003.

[24] A. Kapoor, Y. Qi and R.W. Picard, "Fully Automatic Upper Facial Action Recognition," *Proc. of IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, 2003.

[25] D.-J. Kim and Z. Bien, "A Novel Feature Selection for Fuzzy Neural Networks for Personalized Facial Expression Recognition," *IEICE Tran. on Fundamentals of Electronics, Communications and COmputer Science*, vol.E87-A, no.6, pp. 1386–1392, June 2004.

[26] S.-W. Lee, D.-J. Kim, K.-H. Park and Z. Bien, "Gabor Wavelet Neural Network-Based Facial Expression Recognition," *Proc. of SCI2004*, 2004.

[27] Y.-S. Kim, "An Unsupervised Neural Network Using a Fuzzy Learning Rule," *Proc. of FUZZ-IEEE'99*, 1999.