An Emotion Recognition Based on Advanced Clonal Selection and Extreme Learning Method

Jae-Hoon Cho[1], Dong-Hwa Kim[2], Dae-Jong Lee[3], Keun-Chang Kwak[4], Myung-Geun Chun[1] and Yeong Dal Kim[5]

[1]*Dept. of Electrical and Computer Eng., Chungbuk National University, Cheongju, Korea*
[2]*Dept. of Instrumentation and Control Eng., Hanbat National University, Daejeon, Korea*
[3]*BK21 Chungbuk Information Technology Center, Chungbuk National University, Cheongju, Korea*
[4]*Dept. of Control, Instrumentation, and Robotic Eng., Chosun University, Gwangju, Korea*
[5]*Dept. of Electrical Eng., Hanbat National University, Daejeon, Korea*

(Tel : 82-42-821-1170; Fax : 82-42-821-1164)
(mmi8988@lycos.co.kr, kimdh@hanbat.ac.kr)

Abstract: In this paper, we propose a method for recognizing the human's emotion using extreme learning machine (ELM) and advanced clonal selection of immune algorithm. The most conventional method is to use neural networks that classify several emotions by using speech data and face image data in human's emotion recognition field. However, this approach shows low-speed learning of neural networks due to high-dimension and large size. The main idea of the proposed method is to learn neural networks by high-speed learning algorithm called Extreme Learning Machine (ELM) and to adjust the parameters of ELM by an Advanced Clonal Selection (ACS) of immune algorithm. The experimental results reveal that the proposed method shows a better performance in comparison to the conventional neural networks for human's emotion recognition.

Keywords: Extreme learning machine (ELM), Human's emotion Recognition, Advanced Clonal Selection(ACS).

I. INTRODUCTION

In recent years there has been a growing interest in improving all aspects of the interaction between humans and computers. Human-computer intelligent interaction is a need for the computer to be able to interact naturally with the user, similar to the way human-human interaction takes place. For achieving this, emotion recognition methods have been studied on language, voice, gesture, sight, and facial expressions. Particularly, one of the important ways to recognize the human's emotions is performed by facial expressions. These studies are usually based on the optical flow analysis [1], PCA (principal component analysis)[2], SVM (support vector machines)[3], NN (Neural Networks)[4], and so on.

On the other hand, neural networks have been applied to various facial expression recognition methods [1][4]. However, it has several shortcomings such as difficult setting of learning parameters, slow convergence, and training failures due to local minima and repetitive learning to improve performance of multi-layer neural networks. Also, it is noted that gradient descent-based learning methods are generally very slow, since many iterative learning steps are required by such learning algorithms to obtain better learning performance.

To solve these problems, Huang [5] proposed a novel learning method called Extreme Learning Machine (ELM) to easily achieve good generalization performance at extremely fast learning speed. In the ELM, the input weights and the hidden layer biases are randomly chosen and the output weights (linking the hidden layer to the output layer) are analytically determined by using Moore-Penrose (MP) generalized inverse. Thus, ELM not only learns much faster with higher generalization performance than the traditional gradient-based learning algorithms, but also avoids many difficulties caused by gradient-based learning methods such as stopping criteria, learning rate, learning epochs, and local minima. However, ELM usually needs higher number of hidden neurons due to the random determination of the input weights and hidden biases.

On the other hand, Artificial Immune System (AIS) is one of the developed evolutionary techniques, inspired by the theory of immune system (IS). Immunology is the scientific discipline that studies the response of IS, when an antigenic pattern is recognized by antibodies [6]. The general characteristics of IS are immune memory, hypermutation, and receptor editing. Furthermore, the learning and optimization schemes using the clonal selection principle of IS are proposed in [6]. It has the ability of getting out local minima, operates on a population of points in search space simultaneously, not on just one point, and does not use the derivative or any other information.

In this paper, we proposed a hybrid emotion recognition system using clonal selection to search the optimal input weights and hidden biases of ELM. The proposed method consists of two parts: a preprocessing and optimization part. In the preprocessing, we use PCA to reduce the face image dimension. And then the optimization scheme to improve the performance of ELM is carried out by clonal selection.

This paper is organized as follows. In Section II, the fundamental concept of ELM algorithm is described. In Section III, the emotion recognition based on CS and ELM is proposed. The experimental results and discussion are presented in Section IV. Finally, conclusions are given in Section V.

II. Extreme Learning Machine Algorithm

All the parameters of multi-layer neural networks based on gradient descent-based learning methods need to be learned and usually many iterative learning steps are required to obtain better learning performance. So, gradient descent based learning methods are apt to be slow due to improper learning steps. Moreover, these approaches may easily converge to local minimums. In the ELM, the output weights are analytically computed by using the MP generalized inverse instead of iterative learning scheme. The significant features of ELM can be summarized as follows [6]:

- The learning speed of ELM is extremely fast. It can train SLFNs much faster than classical learning methods.
- The ELM tends to reach not only the smallest training error but also the smallest norm of weights. Thus, the ELM tends to have good performance for neural networks.
- The ELM learning algorithm can be used to train SLFNs with non-differentiable activation functions.
- The ELM tends to reach the solutions straightforward without such trivial issues.

Assuming that we are training SLFNs with K hidden neurons to learn N distinct samples (x_i, t_i), where $\mathrm{x}_i = [x_{i1}, x_{i2}, \cdots, x_{in}]^T \in \Re^n$ and $t_i = [t_{i1}, t_{i2}, \cdots, t_{im}]^T \in \Re^m$, SLFNs with $\tilde{N}$ hidden neurons and activation function $g(x)$ are mathematically modeled as:

$$\sum_{i=1}^{\tilde{N}} \beta_i g(v_i \cdot \mathbf{x}_j + b_i) = o_j, \qquad j = 1, \cdots, N \tag{1}$$

where, $\mathrm{v}_i = [v_{i1}, v_{i2}, \cdots, v_{in}]^T$ is the weight vector connecting the i th hidden neuron and the input neurons, and $w_i = [w_{i1}, w_{i2}, \cdots, w_{im}]^T$ is the weight vector connecting the i th hidden neuron and output neuron. $\mathrm{v}_i \cdot \mathrm{x}_j$ denotes the inner product of v_i and x_j. The output neurons are chosen linear.

That standard SLFNs with $\tilde{N}$ hidden neurons with activation function $g(x)$ can approximate these N samples with zero error means that $\sum_{j=1}^{\tilde{N}} \left\| o_j - t_j \right\| = 0$, i.e., there exist w_i, v_i and b_i such that

$$\sum_{i=1}^{\tilde{N}} w_i g(\mathrm{v}_i \cdot \mathrm{x}_j + b_i) = t_j, \qquad j = 1, \cdots, N \tag{2}$$

The above N equations can be written concisely as:

$$\mathbf{H} w = \mathbf{T} \tag{3}$$

where

$$\mathbf{H}(\mathrm{v}_1, \cdots, \mathrm{v}_{\tilde{N}}, b_1, \cdots b_{\tilde{N}}, \mathrm{x}_1, \cdots \mathrm{x}_{\tilde{N}}) = \begin{bmatrix} g(\mathrm{v}_1 \cdot \mathrm{x}_1 + b_1) & \cdots & g(\mathrm{v}_{\tilde{N}} \cdot \mathrm{x}_1 + b_{\tilde{N}}) \\ \vdots & \vdots & \vdots \\ g(\mathrm{v}_1 \cdot \mathrm{x}_N + b_1) & \cdots & g(\mathrm{v}_{\tilde{N}} \cdot \mathrm{x}_N + b_{\tilde{N}}) \end{bmatrix}$$
$$w = \begin{bmatrix} w_1^T \\ \vdots \\ w_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m} \qquad \mathbf{T} = \begin{bmatrix} \mathrm{t}_1^T \\ \vdots \\ \mathrm{t}_N^T \end{bmatrix}_{N \times m} \tag{4}$$

where $\mathbf{H}$ is the hidden layer output matrix of the neural network; the i th column of $\mathbf{H}$ is the i th hidden neuron's output vector with respect to inputs $x_1, x_2, \cdots, x_N$.

III. Emotion Recognition using Hybrid System of CS and ELM

1. Traditional Clonal Selection

De Castro and Von Zuben [6] presented a clonal selection algorithm to solve complex problems such as learning and multi-modal optimization. The clonal selection is used to describe the basic features of an immune response to an antigenic stimulus. When an antigen encounters the immune system, its epitopes eventually will react with a B-lymphocyte with B-cell receptors on its surface that more or less fit and this activates that B-lymphocyte. This process is known as clonal selection. That is, a B-lymphocyte with an appropriately fitting B-cell receptor can now react with epitopes of an antigen having a corresponding shape. This activates the B-lymphocyte.

Cytokines produced by activated T4-helper lymphocytes enable the now activated B- lymphocyte to rapidly proliferate to produce large clones of thousands of identical B-lymphocytes. In this way, even though only a few B-lymphocytes in the body may have an antibody molecule able to fit a particular epitope, eventually many thousands of cells are produced with the right specificity. This is referred to as clonal expansion. During this time, as the B-lymphocytes proliferate, they undergo affinity maturation as a result of somatic hypermutations. This allows the B-lymphocytes to "fine-tune" the shape of the antibody for better fit with the original epitope. B-lymphocytes having better fitting B-cell receptor on their surface bind epitope longer and more tightly allowing these cells to selectively replicate. The main features of the clonal selection theory are as follows [6]:

1) generation of new random genetic changes, subsequently expressed as diverse antibody patterns by a form of accelerated somatic mutation.
2) phenotypic restriction and retention of one pattern to one differentiated cell (clone).
3) proliferation and differentiation on contact of cells with antigens.

The overall procedure of clonal selection is schematically shown in Fig. 1.

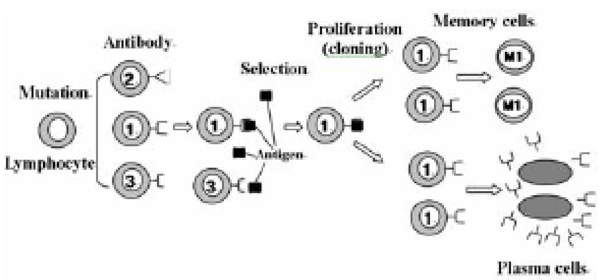

Fig. 1. Procedure of clonal selection

2. Hybrid System of Clonal Selection and Extreme Learning Machine

The proposed method is composed of preprocessing for facial image data and emotion recognition by the proposed hybrid system. In the preprocessing step, we use the PCA to reduce the dimension of the facial image data. The cropped facial image data are applied as the input of the recognition step by the proposed method. After the preprocessing step, the recognition step is performed as follows.

[Step 1] Initialization and recognition of antigen: set the clonal selection algorithm parameters, such as antibody population size, memory cell size, the number of clone, and the value of mutation probability. In the optimization problems, the antigen is a problem to be solved. Here we use the following fitness function as follows

$$f(PI, T\_PI) = \theta \cdot PI + (1-\theta) \cdot T\_PI \tag{5}$$

where f is the fitness function, PI is the training accuracy of ELM, T_PI is the testing accuracy of ELM, and θ is a fitness weights determined according to the preference of individual objectives, respectively.

[Step 2] Product of initial antibody: each clone represents the input weights and the hidden bias of the ELM. Fig. 2 shows the structure of the antibody in the clonal selection.

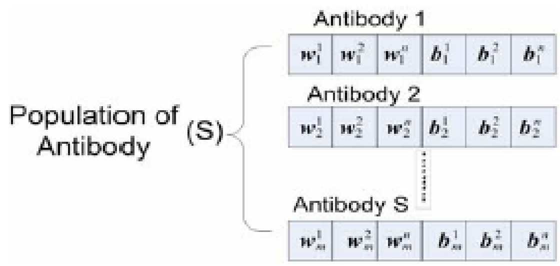

Fig. 2. Structure of the antibody in the clonal selection

[Step 3] Calculation for searching a optimal solution: calculate the fitness values for each antibody by using Eq. (5).

[Step 4] Sort the antibody: we sort the antibody according to their fitness value in an ascending order.

[Step 4-1] Differentiation: differentiate the clones from the memory cell; we select the antibody with the best fitness (the best accuracy of ELM) and differentiate the clones form the memory cell.

[Step 4-2] Mutation: Mutate the population of clones.

[Step 4-3] Calculate the fitness values for each clone.

[Step 4-4] Sort the clones and send the best clone to the memory cell.

[Step 5] Mutation: mutate the antibody population of memory cell.

[Step 6] Calculation the fitness values for the mutated memory cell.

[Step 7] Stop if the specified maximum generation is met, otherwise return to [Step 4].

IV. Experimental results

In the experiment, we use 640×80 facial images containing 6 basic emotions (happiness, sadness, anger, surprise, fear, dislike) for 10 people [7]. Figure 3 shows sample training images of ten persons used in our experiment. In training data containing 30 images of the data for each class (each basic emotion), chosen randomly, were created from the total 360 images and the remaining images were selected as the testing data. After applying PCA, we use the first 50 eigenfaces. In all the simulations, BP, ELM, and the proposed method algorithms are carried out under MATLAB 7.1 environment running in a Pentium 4, 3.2 GHZ CPU. Table 1 shows the initial parameters of the proposed method.

Fig. 3. Training images for ten persons(Happiness)

Table 2 summarizes the results obtained using NN, ELM, and the proposed method. The 50 trials have been conducted for NN and ELM. The average results are listed in Table 2. The results of the NN are obtained by the NN toolbox in MATLAB. As listed in Table 2, the proposed method showed not only higher performance for both the accuracy of training and testing, but also faster learning speed in comparison to others. Table 3 listed the emotion recognition results according to the varying fitness weights. As you can see in Table 3, one can select the fitness weights for choosing the accuracy with respect to training data or testing data.

Table 1. Initial parameters of the proposed method

parameter	value
The number of antibody	50
Differentiation rate	0.1
Generation	20
Mutation rate	0.7

Table 2. Emotion recognition results using facial expression

Method	No. of Hidden Neuron	Training Data (Accuracy)	Test Data (Accuracy)	Training Time (second)
NN	100	97.7	77.0	479
ELM	100	97.0	78.1	199
Proposed method	100	98.8	85.0	198

Table 3. Emotion recognition results by variable fitness weights

Fitness weights		Training Data (Accuracy)	Test Data (Accuracy)
Training	Testing		
0.5	0.5	98.8	85.0
0.4	0.6	98.3	86.7
0.3	0.7	97.2	86.7
0.2	0.8	97.2	87.8

V. CONCLUSION

In this paper, we have proposed a hybrid emotion recognition system using clonal selection to search for the optimal input weights and hidden biases of ELM. To demonstrate the performances of the proposed method, we use facial expression images containing six basic emotions (happiness, sadness, anger, surprise, fear, dislike) for 10 individuals. Finally, the experimental results reveal that the proposed method showed a better generalization performance in comparison with the conventional neural network methods.

REFERENCES

[1] I. Essa and A. Pentland(1995), Facial expression recognition using a dynamic model and motion energy. Int. Conf. on Computer Vision, pp. 360-367.

[2] D. Lin and J. Chen(1999), Facial expressions classification with hierarchical radial basis function networks. Int. Conf. on Neural Information Processing, pp. 1202-1207.

[3] J.M. Susskind, G. Littlewort, M.S. Bartlett, J. Movellan and A.K. Anderson(2007), Human and computer recognition of facial expressions of emotion. Neuropsychologia, Vol. 45, Issue 1, pp. 152-162.

[4] Ioannou, S., Kessous, L., Caridakis, G., Karpouzis, K., Aharonson, V., Kollias, S., & Kollias, S. (2006). Adaptive on-line neural network retraining for real life multimodal emotion recognition. Int. Conf. on Artificial Neural Networks (ICANN'06), Athens, Greece. Lecture Notes in Computer Science, Volume 4131, pp 81-92.

[5] G.-B. Huang, Q.-Y. Zhu, C.-K. Siew(2004), Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks. Int. Joint Conference on Neural Networks (IJCNN2004), pp. 25-29.

[6] L.N. De Castro, F.J. Von Zuben(2002), Learning and optimization using the clonal selection principle. IEEE Trans. Evol. Comput, Vol. 6, pp. 239- 251.

[7] H. J. Go, Y. T. Kim and M. G. Chun(2005), A Multimodal Emotion Recognition Using the Facial Image and Speech Signal. Int. Journal of Fuzzy Logic and Intelligent Systems, Vol 5, No. 1, pp. 1-6.