IMPROVEMENT OF THREE MIXTURE FRAGRANCE RECOGNITION USING FUZZY SIMILARITY BASED SELF-ORGANIZED NETWORK INSPIRED BY IMMUNE ALGORITHM

M.R. Widyanto and K. Hirota

Department of Computational Intelligence & Systems Science, Tokyo Institute of Technology
4259 Nagatsuta, Midori-ku, Yokohama 226-8502, Japan
e-mail: {widyanto,hirota}@hrt.dis.titech.ac.jp

ABSTRACT

To improve the recognition accuracy of a developed artificial odor discrimination system for three mixture fragrance recognition, Fuzzy Similarity based Self-Organized Network inspired by Immune Algorithm (F-SONIA) is proposed. Minimum, average, and maximum values of fragrance data acquisitions are used to form triangular fuzzy numbers. Then, the fuzzy similarity measure is used to define the relationship between fragrance inputs and connection strengths of hidden units. The fuzzy similarity is defined as the maximum value of the intersection region between triangular fuzzy set of input vectors and the connection strengths of hidden units. In experiments, performances of the proposed method is compared with the conventional Self-Organized Network inspired by Immune Algorithm (SONIA), and the Fuzzy Learning Vector Quantization (FLVQ). Experiments show that F-SONIA improves recognition accuracy of SONIA by 3-9%. Comparing to the previously developed artificial odor discrimination system that used FLVQ as pattern classifier, the recognition accuracy is increased by 14-25%.

Keywords: Artificial odor recognition system, Fuzzy Neural Networks, Self Organized Network, Immune algorithm

Manuscript received [15 Juli 2003]. Revised [20 Agustus 2003].

1. INTRODUCTION

An artificial odor discrimination system has been developed to detect and classify odors and gas [1]. The system has been designed to replace the human sensory system, that is used to control the qualities in a variety of industrial fields, e.g., food & beverage industries, cosmetics & perfume industries. The system uses 8 chemical arrayed quartz resonator sensors and Fuzzy Learning Vector Quantization (FLVQ) [2] as pattern classifier.

The artificial system shows a high recognition accuracy to classify pure odors as well as 2 mixture fragrances, but the recognition accuracy for 3 mixture fragrances is a subject to improve. To improve the recognition accuracy for 3 mixture fragrances, the Back-Propagation (BP) based Self-Organized Network inspired by Immune Algorithm (SONIA) [3] is used as pattern classifier. In addition, to further improve the capability of SONIA, the use of a fuzzy similarity [4] instead of Euclidean distance is proposed. The proposed method is called Fuzzy similarity based Self-Organized Network inspired by Immune Algorithm (F-SONIA). In Sec. 2, the scheme of the artificial odor discrimination system is described. The F-SONIA is proposed in Sec. 3. Experimental results on the 3 mixture fragrance problem are presented in Sec. 4.

2. ARTIFICIAL ODOR DISCRIMINATION SYSTEM

The artificial odor discrimination system consists of 3 subsystems, i.e., a sensory system, a frequency counter system, and a neural network as a pattern classifier system. The sensory system and the frequency counter system are used to measure frequency changes during data acquisition, and the pattern classifier system is used to discriminate odor characteristics obtained by the other systems. Figure 1 shows a diagram of the artificial odor discrimination system.

The sensory system used is quartz-resonator crystals that were constructed by sensitive thin chemical membranes. When odorant molecules are absorbed onto the membranes, the resonance frequency of the crystals will decrease significantly and return to the normal resonance frequency after de-absorption process. The change of the frequency is proportional to the total mass of absorbed odorant molecules [4]. These changes of frequency is tranferred by oscillators to the frequency counter. The frequency counter calculates the total frequency changes and through interfacing, data are transfered to computer for input to the pattern classifier.

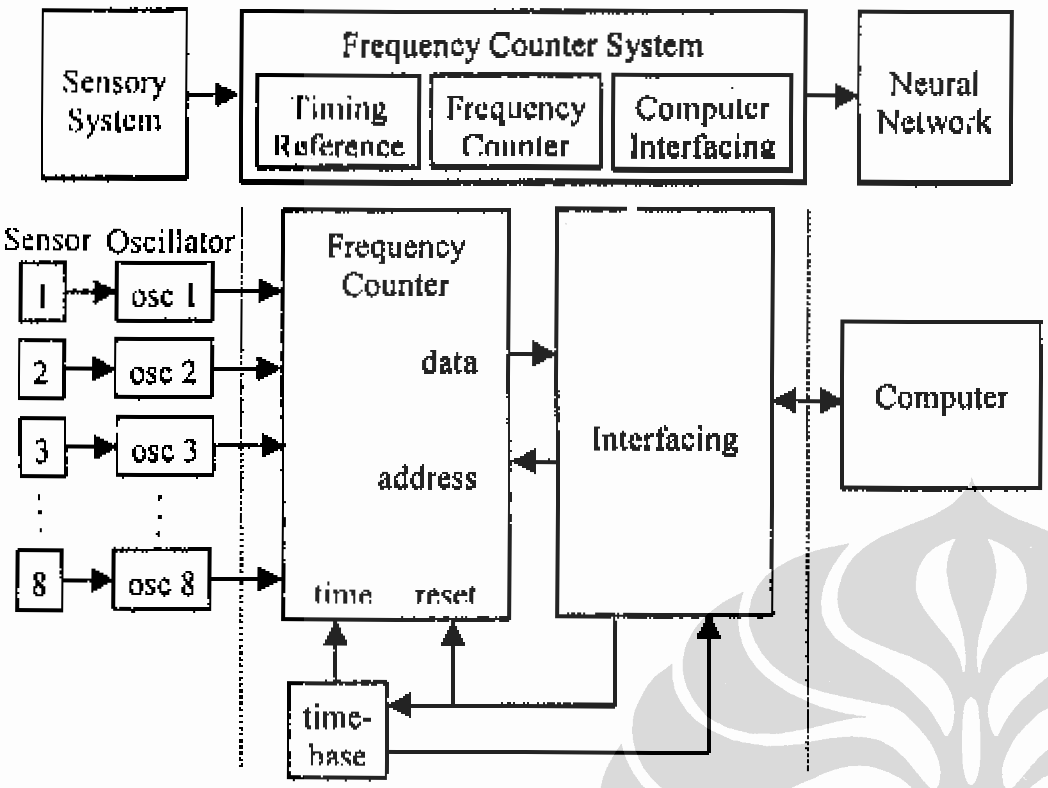

Figure 1. Diagram of Artificial Odor Discrimination System

For each data acquition, samplings are done for several times and an average of these samplings becomes the value for the corresponding data acquition. These values are inputs for the pattern classifier system where a neural network is used. Therefore, input data for the neural network is 8 dimensional data originating from changes of frequency from the 8 sensors used. The data are normalized between 0 and 1.

3. FUZZY SIMILARITY BASED SELF-ORGANIZED NETWORK INSPIRED BY IMMUNE ALGORITHM (F-SONIA)

Fuzzy Similarity based Self-Organized Network inspired by Immune Algorithm (F-SONIA) employs fuzzy similarity measure, in the form of triangular fuzzy sets, between the input vectors and connection strengths of hidden units instead of Euclidean distance, that is used in SONIA [3]. The reason is that the data acquisition of the artificial odor recognition system is calculated based on statistical processes, that is an averaging process. Therefore the use of triangular fuzzy sets is more representative than the use of crisp numbers to represent the statistical properties of fragrance characteristics.

As described in II, for each data acquition, samplings are done for several times and the average of these samplings become the value for the corresponding data acquition. Here, the minimum, the maximum, and the average of samplings are used to form triangular fuzzy sets. The triangular fuzzy sets taken from data acquisition become inputs for the pattern classifier system, where neural network is used. Then the fuzzy similarity measure between an input vector and connection strengths of hidden units is calculated.

F-SONIA consists of three layers, i.e., fuzzy input layer, fuzzy hidden layer, and crisp output layer. Architecture of F-SONIA is shown in Fig. 2. Each layer consists of units i.e., fuzzy input units $\{1,\ldots,N_I\}$, fuzzy hidden units $\{1,\ldots,N_H\}$, and crisp output units $\{1,\ldots,N_O\}$ where $N_I, N_H, N_O \in$ ¥ .

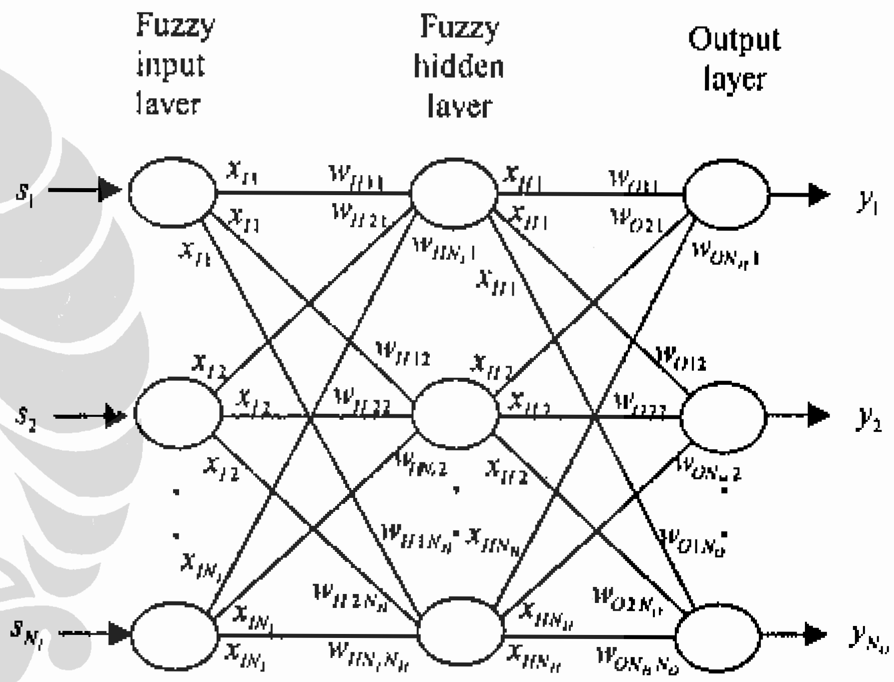

Figure 2. Architecture of F-SONIA (Fuzzy similarity based Self-Organized Network inspired by Immune Algorithm)

Let $T = [0,1]$, the set of normalized inputs, be a universal set, and let $F(T)$ be the family of all fuzzy sets on T :

$$F(T) = \{A \mid T \to [0,1]\}. \quad (1)$$

$$s_i \in F(T), \quad (i = 1,\ldots,N_I),$$
$$x_i \in F(T), \quad (i = 1,\ldots,N_I), \quad (2)$$
$$w_{Hij} \in F(T), \quad (i = 1,\ldots,N_I, j = 1,\ldots,N_H).$$

The *i-th* element of fuzzy input s_i and the strength of the connection between the *i-th* input neuron and the *j-th* hidden neuron w_{Hij} are elements of $F(T)$. Fuzzy input s_i is defined by

$$s_i(t;a_i,b_i,c_i)=\begin{cases} 0, & t\le a_i \\ \dfrac{t-a_i}{b_i-a_i}, & a_i\le t\le b_i \\ \dfrac{c_i-t}{c_i-b_i}, & b_i\le t\le c_i \\ 0, & c_i\le t \end{cases}, \qquad (3)$$

where a_i, b_i, c_i are minimum, average, and maximum values of samplings for the *i-th* sensor respectively. Connection strength w_{Hij} is defined by $w_{Hij}(t;a_{Hij},b_{Hij},c_{Hij})$, where $a_{Hij}, b_{Hij}, c_{Hij}$ are the mean of minimum, average, and maximum values of input vectors belong to the *j-th* hidden neuron constructed through the hidden neuron construction procedure [3] for the *i-th* sensor respectively.

The *i-th* input unit receives fuzzy input s_i $(i = 1,\ldots, N_I)$ to the network and outputs it as inputs to fuzzy hidden units. Output $x_{Ii} \in F(T)$ of the *i-th* input unit is given by

$$x_{Ii} = s_i \qquad (i=1,\ldots,N_I). \qquad (4)$$

For the description of the relationship between x_{Ii} and w_{Hij}, the fuzzy similarity [4] of the two fuzzy set is used through

$$\mu_{ij}(x_{Ii}, w_{Hij}) = \max_{t\in T}\left\{x_{Ii}(t)\wedge w_{Hij}(t)\right\}. \qquad (5)$$

Output of the fuzzy hidden units is calculated by an average of μ_{ij}, therefore the output of the *j-th* hidden unit is given by

$$x_{Hj} = g_l\left(\sum_{i}^{N_I}\mu_{ij} \Big/ N_I\right) \in ([0,1]), \qquad (j=1,\ldots,N_H), \qquad (6)$$

where $g_l(.)$ is a linear function [8]. This use of fuzzy similarity helps the neural network to easily discriminate the fragrance characteristics rather than the use of Euclidean distance. This is because the use of triangular fuzzy set is more representative than crisp numbers to represent statistical properties of fragrance characteristics.

The outputs of hidden units become inputs of output units, these output units provide the network's output

$$y_k = g_u\left(\sum_{j=1}^{N_H} w_{Ojk}x_{Hj} + \theta_{Ok}\right) \in ([0,1]), \qquad (k=1,\ldots,N_O), \qquad (7)$$

where $w_{Ojk} \in$ ¡ represents the strength of the connection from the *j-th* hidden unit to the *k-th* output unit, $\theta_{Ok} \in$ ¡ is a bias associated with the *k-th* output unit, and $g_u(.)$ is a log sigmoid function [8].

The over-all input and output mapping of F-SONIA can be written as

$$y = f(W,\theta_O,s), \qquad (8)$$

where $y=[y_1,\ldots,y_{N_O}]^T$, $W=[w_{Hij}; w_{Ojk}]$ $(i=1,\ldots,N_I,\ j=1,\ \ldots, N_H,\ k=1,\ldots,N_O)$, $\theta_O=[\theta_{O1},\ldots,\theta_{ON_O}]^T$, $s=[s_1,\ldots,s_{N_I}]^T$, and $f=[f_1,\ldots,f_{N_O}]^T$.

Let (s_m, y_m) $(m=1,\ldots,M,\ M\in$ ¥ $)$ be a given set of pairs of input and output to be learned. In order to obtain a network that produces output y_m with respect to input s_m, the values of W and θ_O should be determined such that they will minimize the following error function :

$$Q(W,\theta_O)=\sum_{m=1}^{M}\left\|y_m - f(s_m)\right\|^2, \qquad (9)$$

$$=\sum_{m=1}^{M}\sum_{k=1}^{N_O}(y_{mk}-f_k(s_m))^2, \qquad (10)$$

where y_{mk} is the *k-th* component of vector y_m, $f_k(s_m)$ is the *k-th* component of vector $f(s_m)$, and $f(s_m)$ is a simplification of $f(W,\theta_O,s)$ since the values of W and θ_O should be determined. To minimize Q, the initial values of w_{Ojk} are set to zero and the initial values of θ_O are given randomly. Then the values of w_{Ojk} and θ_O are modified iteratively, i.e., the modification of w_{Ojk} and θ_O should be

$$\nabla w_{Ojk} = -\eta \partial Q / \partial w_{Ojk}, \quad (j = 1, \ldots, N_H, k = 1, \ldots, N_O), \tag{11}$$

$$\nabla \theta_{Ok} = -\eta \partial Q / \partial \theta_{Ok} \;\; (k = 1, \ldots, N_O), \tag{12}$$

where $\eta \in (0,1)$ is a positive constant. The calculations of $\partial Q / \partial w_{Ojk}$ and $\partial Q / \partial \theta_{Ok}$ are performed by the back-propagation algorithm [6].

4. EXPERIMENTS ON 3 MIXTURE FRAGRANCE PROBLEM

The experiments used three data sets of 3 mixture fragrances namely Citrus-Canangga-Ethanol, Citrus-Rose-Ethanol, and Rose-Canangga-Ethanol. Each data set is a mixture of 2 vegetal aromas and 1 ethanol concentration with composition 1:1:1. There are 6 ethanol concentrations (0%, 15%, 25%, 35%, 45% and 70%) in those formed 6 fragrance categories for each data set. Each data set consists of 120 entries for training and 120 entries for testing. The performance of the proposed F-SONIA is compared with conventional SONIA [3], FLVQ [2], LVQ [9] and BP [6] neural network. For every data set, performances of those methods were observed through learning phase and testing phase. In learning phase, the learning time required is recorded, meanwhile in testing phase, the recognition accuracy is measured by using training and testing data as inputs. The parameters of SONIA and F-SONIA are set therefore 30 neurons are formed. For a balance comparison, the hidden neuron number of FLVQ, LVQ, and BP are set to be 30, other parameters are set as advised by the Matlab neural network toolbox [8]. For each experiment, the learning & testing phases are carried out for 10 times, and the recognition accuracy is the average of these trials. The experiments have been done using Matlab 6.1 C/C++ Compiler under Microsoft Windows 2000 operating system on Pentium 4 2.0 GHz processor with 128 Megabytes memory. For LVQ and BP, Matlab neural network toolbox is used [8].

For all data sets, F-SONIA improves the recognition accuracy of the conventional SONIA by 3-9% (Table 1). Meanwhile the average learning time of F-SONIA is 0.3 times compared to the conventional SONIA (Fig. 3). This shows that the use of fuzzy similarity, that contains statistical properties of mixture fragrances, helps the neural network to easily discriminate the fragrance characteristics therefore the recognition accuracy is improved and the learning time is reduced. Compared to the previously developed artificial odor recognition system that uses FLVQ as pattern classifier, the recognition accuracy is increased by 14-25%.

5. CONCLUSIONS

A Fuzzy Similarity based Self-Organized Network inspired by Immune Algorithm (F-SONIA) is proposed. F-SONIA employs fuzzy similarity measure between inputs and the connection strength of hidden units instead of Euclidean distance, (that is used in SONIA). The use of fuzzy similarity measure instead of Euclidean distance enables F-SONIA to easily discriminate statistical characteristics of mixture fragrance.

From the experiments, the proposed method improves the recognition accuracy of the conventional SONIA by 3-9% for 3 mixture fragrance problem. Meanwhile the average learning time of F-SONIA is 0.3 times comparing to the conventional SONIA. This shows that the use of fuzzy similarity, that contains statistical properties of mixture fragrances, helps the neural network to easily discriminate the fragrance characteristics therefore the recognition accuracy is improved.

Compared to the previously developed artificial odor recognition system which used FLVQ as pattern classifier, the recognition accuracy is increased by 14-25%.The use of fuzzy similarity helps the neural network to easily discriminate the fragrance characteristics. In the near future, we will investigate other types of fuzzy similarity measures on F-SONIA and observe how these kinds of fuzzy similarity affect the recognition accuracy.

Method	Citrus-Canangga-Ethanol (%)		Citrus-Rose-Ethanol (%)		Rose-Canangga-Ethanol (%)		Average Learning Time (seconds)
	Training Data	Testing Data	Training Data	Testing Data	Training Data	Testing Data	
F-SONIA [proposed]	100	100	95.83	98.33	90.83	94.17	99.77
SONIA [3]	87.50	90.83	89.17	89.17	89.17	90.83	332.56
FLVQ [2]	70.00	75.00	77.50	74.17	74.17	75.00	271.77
LVQ [9]	33.33	33.33	35.00	34.17	37.50	40.83	289.15
BP [6]	34.17	32.50	32.50	33.33	31.67	37.50	9.82

Table 1. Recognition Accuracy and Learning Time for 3 Mixture Fragrances

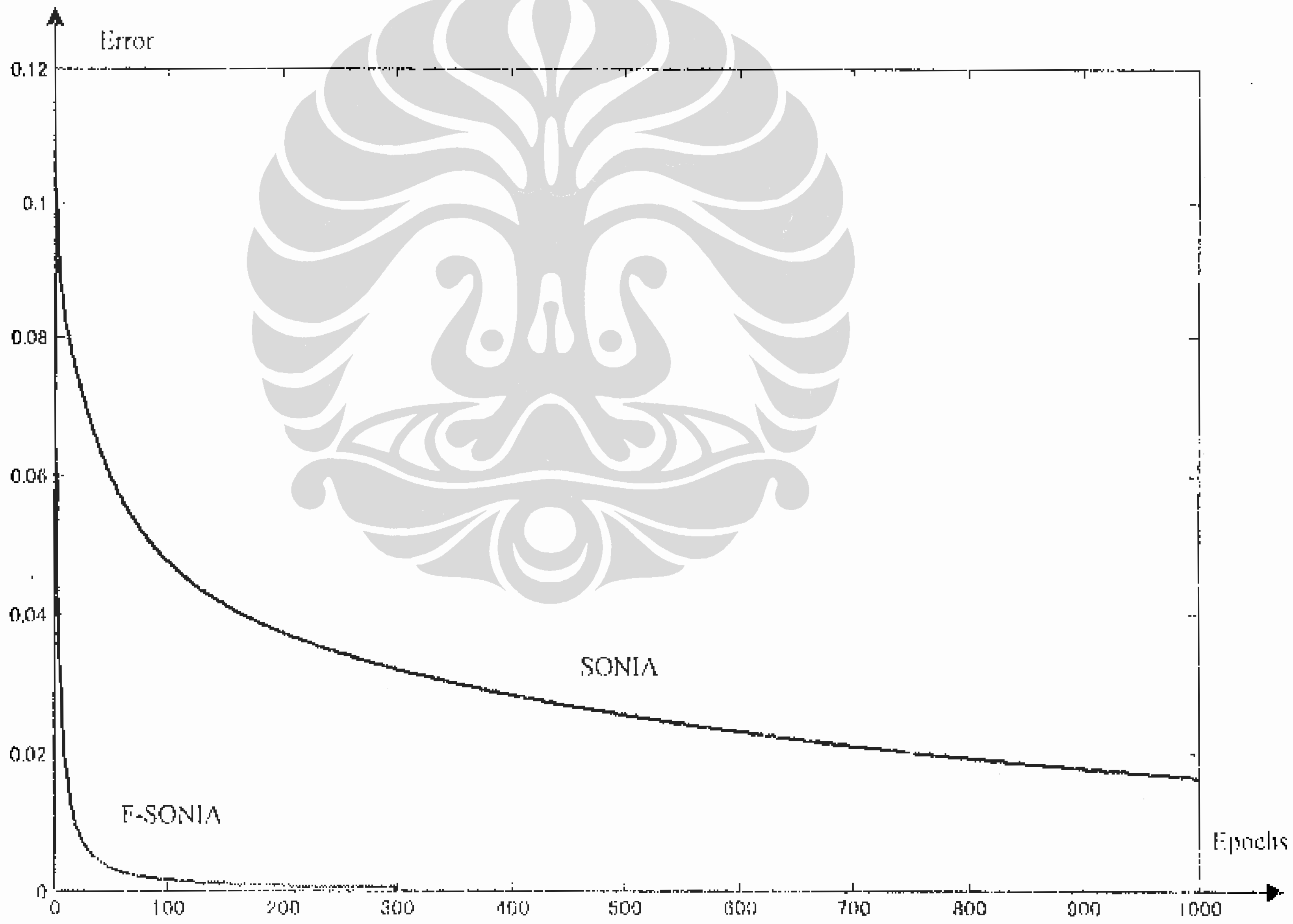

Figure 3. F-SONIA is faster convergence than SONIA

REFERENCES

[1] B. Kusumoputro, M.R. Widyanto, M.I. Fanany, and H. Budiarto, "Improvement of artificial odor discrimination system using FLVQ neural network", *Proceedings of the 3rd International Conference on Computational Intelligence and Multimedia Application*, New Delhi, India, pp. 474-478, 1999.

[2] Y. Sakuraba, T. Nakamoto, and T. Moriizumi, "New method of learning vector quantization", *Systems and Computer in Japan*, (22) 13, pp. 93-102, 1991.

[3] M. R. Widyanto, Megawati, Y. Takama, and K. Hirota, "A time-temperature based food quality prediction using a self-organized network inspired by immune algorithm", *Proceedings of the 1st International Conference on Soft Computing and Intelligent Systems*, Tsukuba, Japan, 2002.

[4] L. A Zadeh, "Similarity relations and fuzzy orderings", *Information Sciences*, (3), pp. 177-200, 1971.

[5] G. Sauerbrey, "Verwedung von schwingquaren zur wagung dunner shichten und zur mikrowagung", *Z. Physics*, (155), pp. 206-222, 1959.

[6] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by error propagation", *Nature*,(323), pp. 533-536, 1986.

[7] J. I. Timmis, *Artificial immune systems : a novel data analysis technique inspired by the immune network theory*, PhD. Dissertation, University of Wales, Aberystwyth, 2001.

[8] H. Demuth, and M. Beale, "Neural network toolbox", *Matlab 6.1 User Guide*, Mathworks, Chapter 5, 2001.

[9] T. Kohonen, "Learning vector quantization for pattern recognition", *Technical Report TKK-F-A601*, Helsinki University of Technology, Finland, 1986.