2008 IEEE Workshop on
Nonlinear Circuit Networks, Tokushima, Dec 5-6, 2008

Improved Fuzzy Adaptive Resonance Theory Combining Overlapped Category in Consideration of Connections

Haruka ISAWA, Haruna MATSUSHITA and Yoshifumi NISHIO
Dept. Electrical and Electronic Eng.,
Tokushima University
Email: {isawa, haruna, nishio}@ee.tokushima-u.ac.jp

***Abstract*— The Adaptive Resonance Theory (ART) is an unsupervised neural network. Fuzzy ART is a variation of ART, allows both binary and continuous input patterns. However, Fuzzy ART has the category proliferation problem. To solve this problem, we have proposed Fuzzy ART Combining Overlapped Category in consideration of connections (C-FART) in previous study. In this study, we propose a improved C-FART. The important features of the improved C-FART are that the vigilance parameters are arranged for every category and they are varied according to the size of respective categories with learning. We investigate the behavior of the improved C-FART, and compare the improved C-FART with the simple Fuzzy ART and the conventional C-FART.**

I. Introduction

Self-organized clustering is a powerful tool whenever huge sets of data have to be divided into separate categories. In the field of neural network, the Adaptive Resonance Theory (ART) [1], introduced and developed G.A.Carpenter and S.Grossberg, is a popular representative for self-organized clustering. Some outstanding features of ART, besides its clustering capabilities, have attracted the attention from application engineers. This theory has evolved as a series of real-time neural network models that perform unsupervised and supervised learning, pattern recognition, and prediction. These models are capable of learning stable recognition categories in response to arbitrary input sequences. However, the fusion of a computational efficiency of neural network and a capability of fuzzy logic to represent complex class boundaries, has created a lot of interest in neurofuzzy pattern recognition systems. Then, we pay attentions Fuzzy ART [2], which is the merger of fuzzy logic and ART neural network. Fuzzy ART is applied to association, clustering and memory of input pattern. Fuzzy ART classified input data into each appropriate category by creating rectangles. However, Fuzzy ART has category proliferation problem. For this reason, Fuzzy ART performance is highly dependant on a vigilance parameter, which controls category size.

Then, we have noted overlapped category and have proposed Fuzzy Adaptive Resonance Theory Combining Overlapped Category in Consideration of Connections (C-FART) [5] in previous study. C-FART has two important features. One is to make connections between similar categories, and the other is to combine overlapped categories into one category with connections. With this method, we can reduce category proliferation. However, C-FART also has a disadvantage. Categories get larger by combining overlapped categories as Fig. 1. Therefore, if input data is inputted in the shaded area in Fig. 1(b), it is recognized as a new category. In other words, a new category is created inside a category. This is because the vigilance parameter is constant value. For all of these reasons, a lot of overlapped categories are created.

In this study, we propose improved C-FART. The important features of the improved C-FART are that the vigilance parameters are arranged for every category and they are varied according to the size of respective categories with learning. We investigate the behavior of the improved C-FART and compare the improved C-FART with the simple Fuzzy ART and the conventional C-FART.

II. Proposed Fuzzy ART (The improved C-FART)

A. Structure of the improved C-FART

The structure of the improved C-FART is same as the conventional Fuzzy ART.

Input vector: Each input $\boldsymbol{I}$ is an m-dimensional vector $\boldsymbol{I} = (I_1, \cdots, I_m)$, where $I_i \in [0,1]^m$.

Weight vector: Each category j corresponds to a vector $\boldsymbol{w}_j = (w_{j1}, \cdots, w_{jm})$, $(j = 1, \cdots, n)$ of an adaptive weight. The number of potential categories n is arbitrary. Initially $w_{j1} = \cdots = w_{jm} = 1$.

Parameters: The improved C-FART dynamics are determined by a *choice parameter* $\alpha > 0$; a *learning parameter* $\beta \in [0,1]$; and the *vigilance parameter* $\rho_j \in [0,1]$. The vigilance parameter is a fixed value for all categories in the simple Fuzzy ART and the conventional C-FART. The important features of the improved C-FART are that the vigilance parameters are arranged for every category and they are varied according to the size of respective categories with learning. Initially $\rho_1 = \cdots = \rho_n = \rho_0$.

Connection: The improved C-FART has a connection C and the age of the connections age. Both C and age are $n \times n$ matrices. The initial values of C and age are set to zero. If the categories J and j are connected with learning, $C_{J,j}$ changes from zero to one.

B. Learning algorithm of the improved C-FART

The learning algorithm of the improved C-FART consists of mainly four processes: 1) Learning, 2) Update Connections, 3) Combining Categories and 4) Varying the vigilance parameter. These four steps are repeated for all input data set. Therefore, the improved C-FART makes or releases connections at each step, and overlapped categories are combined with considering their connections. Furthermore, the vigilance parameter varies according to the size of the combined category.

Process I : Learning

(StepI-1) An input vector $\boldsymbol{I}$ is inputted to the category layer F_2 from the input layer F_1.

(StepI-2) For the input vector $\boldsymbol{I}$ and the category j, *choice function* T_j is defined by

$$T_j(\boldsymbol{I}) = \frac{|\ \boldsymbol{I} \wedge \boldsymbol{w}_j\ |}{(\alpha + |\ \boldsymbol{w}_j\ |)}, \tag{1}$$

where the fuzzy AND [4] operator $\wedge$ and the norm $|\cdot|$ are defined by

$$(\boldsymbol{p} \wedge \boldsymbol{q})_i \equiv \min(p_i, q_i), \quad |\ \boldsymbol{P}\ | \equiv \sum_{i=1}^{m} |\ p_i\ |. \tag{2}$$

The winning category J, whose T_j is maximum, is found;

$$J = \arg\max_j \{T_j\}. \tag{3}$$

If more than one T_j is maximal, the category j with the smallest index is chosen as the winning category J. Furthermore, a second-winning category J_2, whose T_{J_2} is the second largest next to T_J, is found for updating connections if J_2 exists.

(StepI-3) The similarity of $\boldsymbol{I}$ and $\boldsymbol{w}_J$ is measured by the vigilance criterion according to

$$\frac{|\ \boldsymbol{I} \wedge \boldsymbol{w}_J\ |}{|\ \boldsymbol{I}\ |} \geq \rho_J. \tag{4}$$

If Eq. (4) is not satisfied, a new index J is chosen by Eq. (3) and the search process continues until the chosen J satisfies Eq. (4).

(StepI-4) If any J satisfies Eq. (4), $\boldsymbol{w}_J$ is updated by

$$\boldsymbol{w}_J^{\text{new}} = \beta(\boldsymbol{I} \wedge \boldsymbol{w}_J^{\text{old}}) + (1-\beta)\boldsymbol{w}_J^{\text{old}}, \tag{5}$$

and we perform (StepII-1). On the contrary, if all categories do not satisfy Eq. (4), a new category is established;

$$\boldsymbol{w}_{n+1} = \boldsymbol{I}, \tag{6}$$

and we proceed to next input data.

Process II : Update Connections

(StepII-1) If J_2 does not exist, we skip this step and perform (StepII-2). The similarity of $\boldsymbol{I}$ and $\boldsymbol{w}_{J_2}$ is measured as Eq. (4). If its similarity satisfies ρ_{J_2}, a connection between the winning category J and the second-winning category J_2 is created;

$$C_{J,J_2} = 1. \tag{7}$$

The age of the connection between J and J_2 is set to zero ("refresh" the age);

$$age_{J,J_2} = 0. \tag{8}$$

On the contrary, if Eq. (4) is not satisfied, the connection is not updated.

(StepII-2) The age of all categories, which directly connect with the winning category J, are increased one;

$$age_{J,j}^{\text{new}} = age_{J,j}^{\text{old}} + 1, \qquad j \in N_J, \tag{9}$$

where N_J is the set of categories which directly connect with J, namely $C_{J,j} = 1$.

(StepII-3) The connections are removed, if their age exceeds a threshold value $AT(t)$;

$$C_{J,j} = 0, \quad if \ \ age_{J,j} \geq AT(t), \tag{10}$$

where

$$AT(t) = AT_i \left(\frac{AT_f}{AT_i}\right)^{\frac{t}{t_{\max}}}, \tag{11}$$

where t is the learning step, $t_{\max}$ is the learning length, AT_i and AT_f is the initial value and the final value of AT, respectively.

Process III : Combining Categories

(StepIII-1) We check whether the winning category J is overlapping with other category and combine these categories. The inputs belonging to the category k are classified to the category J, and the category k is removed, therefore, $n^{\text{new}} = n^{\text{old}} - 1$.

Without loss of generality, we assume the weight vector $\boldsymbol{w}_j$ can be written in complement cording form [4]: $\boldsymbol{w}_j = (\boldsymbol{u}_j, \boldsymbol{v}_j^c)$, where $\boldsymbol{u}_j$ and $\boldsymbol{v}_j^c$ are m-dimensional vectors. We find the category k which are overlapping with J as Fig. 1(a), in other words, Fig. 1(a) is satisfied for any dimension i.

Let vector $\boldsymbol{u}_J$ and $\boldsymbol{v}_k$ define one corner of a rectangle, and let $\boldsymbol{v}_J$ and $\boldsymbol{v}_k$ define another corner of rectangle as Fig. 1(a). If k exists and directly connects with J, namely $C_{J,k} = 1$, J and k are combined as Fig. 1(b) by

$$\boldsymbol{w}_J^{\text{new}} = \{(\boldsymbol{u}_J \wedge \boldsymbol{u}_k), (\boldsymbol{v}_J^c \wedge \boldsymbol{v}_k^c)\}. \tag{12}$$

Process IV : Varying the vigilance parameter

(StepIV-1) The vigilance parameter ρ_J is varied by

$$\rho_J = \frac{1 - |\ \boldsymbol{w}_J\ |}{m}, \tag{13}$$

In the conventional Fuzzy ART, the vigilance parameter is invariant. However, the vigilance parameter is varied according to the size of combined category in the improved C-FART.

III. Simulation Results

We apply the improved C-FART to 2-dimensional input data, and compare the improved C-FART with the simple Fuzzy ART and the conventional C-FART.

We consider 2-dimensional input data as Fig. 2(a), consisting of 2-clusters. Total number of the input data is 450 points,

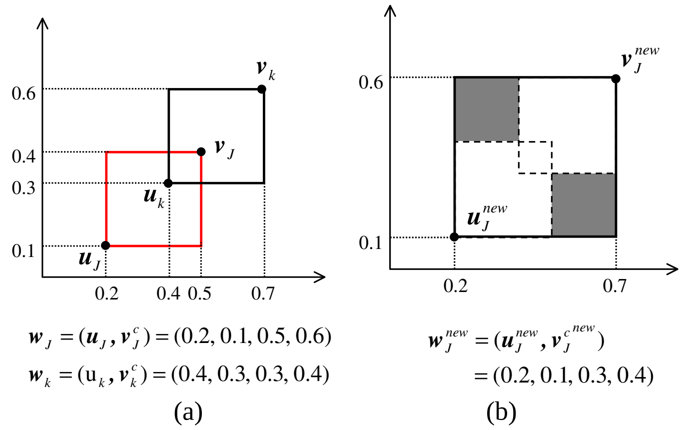

Fig. 1. Overlapped categories combining process in the improved C-FART. (a) Category J and category k are overlapping. (b) If $C_{J,k} = 1$, $\boldsymbol{w}_J$ are combined according to Eq. 12, and k is removed. The vigilance parameter is constant in the conventional Fuzzy ART, however the vigilance parameter is varied at (StepIV-1) in the improved C-FART.

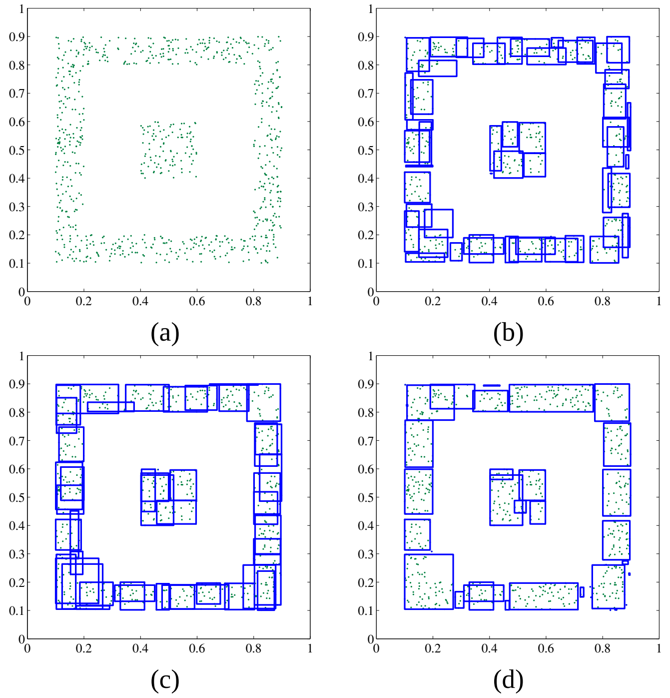

Fig. 2. Simulation results. (a) Input data. (b) Simulation result of the simple Fuzzy ART. (c) Simulation result of the conventional C-FART. (d) Simulation result of the improved C-FART.

TABLE I

NUMBER OF CATEGORIES

Algorithm	Number of categories		
	Minimum	Maximum	Average
Fuzzy ART	47	59	52.88
C-FART	39	54	45.66
Improved CFART	17	46	29.77

the inside cluster has 100 points and the outside cluster has 350 points. The input data are sorted at random. When learning starts, there is no category, namely $n = 0$. The parameters for the learning are chosen as follows;

(For Fuzzy ART)

$$\alpha = 0.1, \;\; \beta = 1.0, \;\; \rho = 0.9,$$

(For the conventional C-FART)

$$\alpha = 0.1, \;\; \beta = 1.0, \;\; \rho = 0.9, \;\; AT_i = 2, AT_f = 7,$$

(For the improved C-FART)

$$\alpha = 0.1, \;\; \beta = 1.0, \;\; \rho_0 = 0.9, \;\; AT_i = 2, AT_f = 7,$$

The learning results of the simple Fuzzy ART and the conventional C-FART are shown in Figs. 2(b) and (c), respectively. We can see that category proliferation occurs and a lot of overlapped categories are created. The other side, the result of the improved C-FART is shown in Fig. 2(d). From this figure, we can see that category proliferation and overlapped categories are reduced. This is because the vigilance parameter is varied according to the size of category with learning. In addition, the improved C-FART can classify input data into more appropriate category because the vigilance parameter is arranged for all categories.

We perform the learning simulation 100 times to the input data created in the same way as Fig. 2(a). The input data are created each time and are sorted by random to remove the dependence on the order of the input data.

Table I summarizes minimum, maximum and average values of the number of categories. In other words, the minimum and the maximum values mean best and worst results in 100 simulations, respectively. All the number of categories of the improved C-FART are smaller than the simple Fuzzy ART and the conventional C-FAT. From the average values, the improved C-FART has reduced category proliferation and overlapped categories by 43% as compared to the simple Fuzzy ART. Furthermore, the improved C-FART has improved 34% from using the conventional C-FART. We have obtained more effective results than the simple Fuzzy ART and the conventional C-FART by using the improved C-FART.

IV. CONCLUSIONS

In this study, we have proposed the improved C-FART to solve the category proliferation and overlapped category problem. The improved C-FART varies the vigilance parameter in accordance with the size of category. We have applied the improved C-FART into 2-dimensional input data, and the learning behaviors of the improved C-FART have been investigated. We have confirmed that the improved C-FART can obtain more effective result than the simple Fuzzy ART and the conventional C-FART, and have reduced category proliferation and overlapped categories.

REFERENCES

[1] C.A. Carpenter, "Distributed learning, recognition, and prediction by ART and ARTMAP neural networks," *Neural Networks*, vol. 10, pp. 1473–1494, 1997.

[2] G.A. Carpenter, S. Grossberg, D.B. Rosen, "Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system," *Neural Networks*, vol. 4, pp. 759–771, 1991.

[3] T. Frank, K.F. Kraiss and T. Kuhlen, “Competitive analysis of Fuzzy ART and ART-2A network clustering performance,” *IEEE Trans. Neural Networks*, vol. 9, no. 3, pp. 544–559, 1998.
[4] L. Zadeh, “Fuzzy sets,” *Information and Control*, vol. 8, pp. 338–353, 1965.
[5] H. Isawa, H. Matsushita and Y. Nishio, “Fuzzy Adaptive Resonance Theory Combining Overlapped Category in Consideration of Connections” *Proc. of International Joint Conference on Neural Networks*, pp. 3594–3599, 2008.