2009 International Workshop on Nonlinear Circuits and Signal Processing
NCSP'09, Waikiki, Hawaii, March 1-3, 2009

Fuzzy ART Combining Overlapped Categories Using Variable Vigilance Parameters

Haruka Isawa, Haruna Matsushita and Yoshifumi Nishio

Tokushima University
2-1 Minami-Josanjima, Tokushima, Japan
Phone:+81-88-656-7470, Fax:+81-88-656-7471
Email: {isawa, haruna, nishio}@ee.tokushima-u.ac.jp

Abstract

The Fuzzy Adaptive Resonance Theory (Fuzzy ART) is an unsupervised neural network and allows both binary and continuous input patterns. In this study, we propose a Fuzzy ART Combining Overlapped Categories Using Variable Vigilance Parameters. The vigilance parameters of the proposed method are arranged for every category, and they are varied according to the size of respective categories with learning. We confirm that the proposed Fuzzy ART can classify input data more flexible than the conventional Fuzzy ART.

1. 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] is a popular representative for self-organized clustering. 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. The Fuzzy ART [2]-[3] is the merger of fuzzy logic and ART neural network. Fuzzy ART is applied to association, clustering and memory of input pattern, and it classifies the input data into each appropriate category by creating rectangles. However, because Fuzzy ART frequently classifies the input data of the common categories into several categories, a category proliferation problem occurs. For this reason, Fuzzy ART performance is highly dependent on a vigilance parameter which controls category size.

To solve these problems, we have proposed two types of Fuzzy ART in our past studies. The first is Fuzzy ART with Group Learning (FART-GL) [4]. FART-GL makes connections between categories or releases connection as human relationships which keep changing with time in the real world. The connection is created between similar categories, and the connected categories are learned as "group" of category. By using this method, the input data are classified into each appropriate group. However, FART-GL can not solve the category proliferation problem.

We have also proposed a Fuzzy ART Combining Overlapped Category in Consideration of Connections (C-FART) [5]. C-FART makes connections between similar categories or releases connections at each step as FART-GL and combines overlapping categories by using created connections. C-FART can reduce category proliferation by combining the categories with due consideration of their similarity. However, because the vigilance parameter is constant value although the categories become larger by combining overlapped categories, a new category is created inside existing categories. This creates a lot of overlapped categories.

In this study, we improve the conventional C-FART and propose a Fuzzy ART combining overlapped categories using variable vigilance parameters. The proposed method has both FART-GL and C-FART abilities. Moreover, the vigilance parameters are arranged for every category, and they are varied according to the size of respective categories with learning. We investigate the behaviors of the proposed method by applying to various input data.

2. Proposed Fuzzy ART

The structure of the proposed Fuzzy ART is same as the standard 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 proposed Fuzzy ART dynamics are determined by a *choice parameter* $\alpha > 0$; a *learning parameter* $\beta \in [0, 1]$; and the *vigilance parameter* $\rho_j \in [0, 1]$. In contrast to the standard Fuzzy ART and the conventional C-FART whose vigilance parameter is a fixed value for all categories, the vigilance parameters of the proposed method are arranged for every category and are varied according to the size of respective categories with learning. Initially $\rho_1 = \cdots = \rho_n = \rho_0$.

Connection: The proposed method 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.

2.1. Learning Algorithm of Proposed Method

The learning algorithm of the proposed method 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 the input data set. Therefore, the proposed method 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.

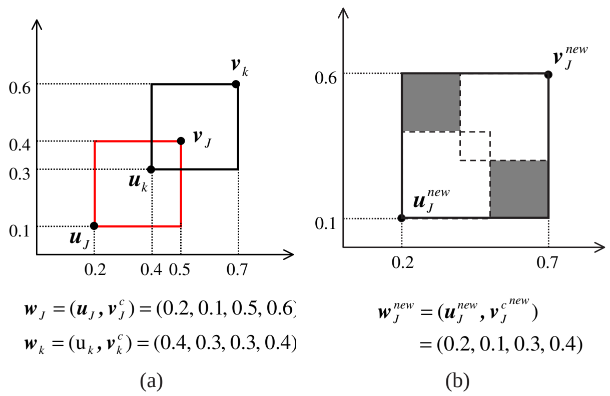

Figure 1: Combining process of overlapped categories. (a) Category J and category k are overlapping. (b) If $C_{J,k} = 1$, $\boldsymbol{w}_J$ are combined according to Eq. (9), and k is removed. The vigilance parameter of the proposed method is varied according to (StepIV-1).

Process I : Learning

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

(StepI-2) We calculate *choice function* T_j of each category j;

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

where the fuzzy AND [6] 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.

A second-winning category J_2 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 $\boldsymbol{w}_{n+1} = \boldsymbol{I}$, 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 as $C_{J,J_2} = 1$. The *age* of the connection between J and J_2 is set to zero ("refresh" the age) $age_{J,J_2} = 0$. 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;

$$age_{J,j}^{\text{new}} = age_{J,j}^{\text{old}} + 1, \qquad j \in N_J, \tag{6}$$

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 $AT(t)$;

$$C_{J,j} = 0, \quad if \;\; age_{J,j} \geq AT(t), \tag{7}$$

where

$$AT(t) = AT_i \left(AT_f / AT_i\right)^{\frac{t}{t_{\max}}}, \tag{8}$$

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

Process III : Combining Categories

(StepIII-1) We check whether the winning category J is overlapping with other category k and combine these categories. The input data belonging to the category k overlapping with the category J are classified to 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 [6]: $\boldsymbol{w}_j = (\boldsymbol{u}_j, \boldsymbol{v}_j^c)$, where $\boldsymbol{u}_j$ and $\boldsymbol{v}_j^c$ are m-dimensional vectors. We find the

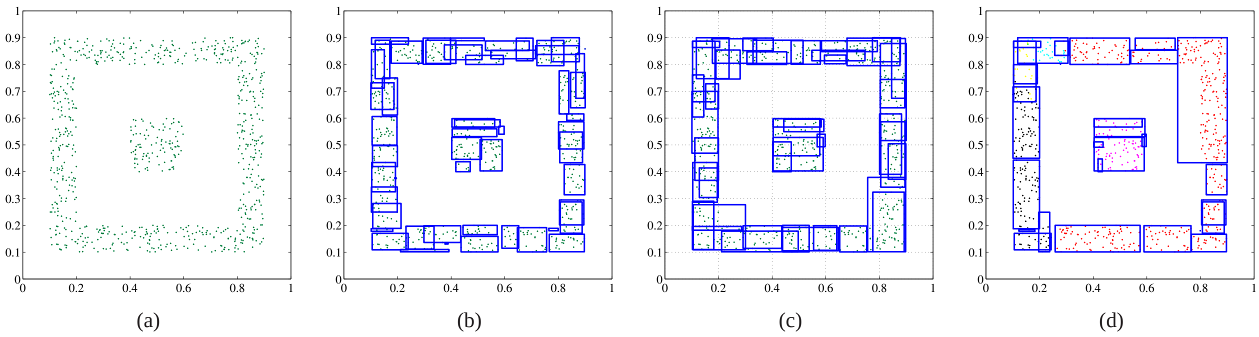

Figure 2: Simulation 1 for rectangular-shaped data. (a) Input data. (b) Simulation result of Fuzzy ART. (c) Simulation result of C-FART. (d) Simulation result of the proposed method.

category k 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

$$\boldsymbol{w}_J^{\text{new}} = \left\{ (\boldsymbol{u}_J \wedge \boldsymbol{u}_k), \left(\boldsymbol{v}_J^c \wedge \boldsymbol{v}_k^c\right) \right\}. \quad (9)$$

Process IV : Varying the vigilance parameter

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

$$\rho_J = \frac{1 - |\boldsymbol{w}_J|}{m}, \quad (10)$$

In the standard Fuzzy ART, the vigilance parameter is invariant. However, the vigilance parameter of the proposed method is varied according to the size of combined category.

3. Computer Simulations

We apply the proposed method to 2-dimensional input data, and compare the proposed method with the standard Fuzzy ART and conventional C-FART. Preferable features of the obtained results are listed below;

1. High classification accuracy.
2. The number of categories is small.
3. Overlapped area are small.

Without high classification accuracy, the input data belonging to different cluster into same category. If there are many categories, information of categorization is too much. With high classification accuracy, it is preferable that overlapped area is small.

3.1. Simulation 1

First, we consider 2-dimensional rectangular-shaped input data as Fig. 2(a), consisting of 2-clusters. Total number of the input data is 450 points, the inside cluster has 100 points and the outside cluster has 350 points. The input data are sorted at random. When $t = 0$, 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 C-FART) $\alpha = 0.1$, $\beta = 1.0$, $\rho = 0.9$, $AT_i = 2, AT_f = 25$
(For proposed method) $\alpha = 0.1$, $\beta = 1.0$, $\rho_0 = 0.9$, $AT_i = 2, AT_f = 25$.

The learning results of the Fuzzy ART and C-FART are shown in Figs. 2(b) and (c), respectively. We can see that the category proliferation occurs, and there are a lot of overlapped area. On the other hand, the result of the proposed method is shown in Fig. 2(d). From this figure, we can see that the category proliferation and the overlapped area are reduced. This is because that the vigilance parameter is varied according to the size of a category with learning. Therefore, the proposed method can categorize the input data more flexible than other two methods. Furthermore, the proposed method can recognize the group of the categories by whether there are connections between categories, and its results are shown in different colors. Therefore, we can obtain not only the detailed informations by the categorization results, but also the relationships between categories by the groups.

Table 1: Number of categories of Simulation 1.

Algorithm	Number of categories		
	Minimum	Maximum	Average
Fuzzy ART	47	59	52.9
C-FART	39	54	45.7
Proposed method	17	46	29.8

Furthermore, we carry out the learning simulations repeated 100 times, and Table 1 summarizes performances with the minimum, maximum and average values of the number of categories over 100 independent runs. The minimum and the maximum values mean best and worst results in 100 sim-

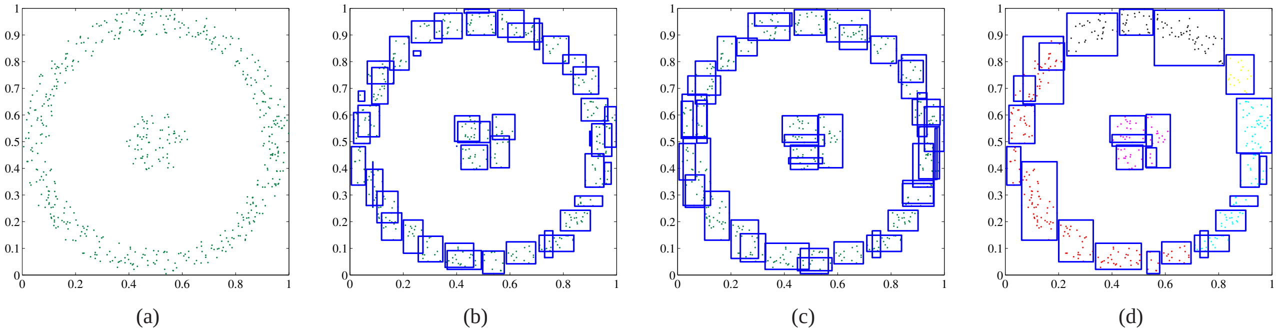

Figure 3: Simulation 2 for circular-shaped data. (a) Input data. (b) Simulation result of Fuzzy ART. (c) Simulation result of C-FART. (d) Simulation result of the proposed method.

ulations, respectively. All the number of categories of the proposed method are smaller than Fuzzy ART and C-FART. From the average values, the proposed method has reduced category proliferation 43% and 34% from Fuzzy ART and C-FART, respectively. We can obtain more effective information of categorization than Fuzzy ART and C-FART because the proposed method can reduce the number of categories.

3.2. Simulation 2

Next, we apply the proposed method to circular-shaped input data, which have 2-clusters as Fig. 3(a). The inside cluster has 70 points and the outside cluster has 500 points. It is difficult to classify the input data such as Fig. 3(a) into appropriate categories because categories are created by rectangles. The learning conditions are the same used in Simulation 1.

Figures 3(b) and (c) show the simulation results of Fuzzy ART and C-FART, respectively. We can see that category proliferation occurs, and there are a lot of overlapped area. On the other side, in the results of the proposed method shown in Fig. 3(d), the category proliferation and the overlapped area are reduced. Furthermore, we can recognize the group of the categories from its color from the result of categorization.

The performances with the minimum, maximum and average number of categories over 100 independent runs are listed in Table 2. We can see that all the number of categories of the proposed method are smaller than Fuzzy ART and C-FART. From the average values, the proposed method has reduced category proliferation 45% and 35% from Fuzzy ART and C-FART, respectively. Therefore, we have obtained more effective information of categorization than Fuzzy ART and C-FART by using the proposed method.

Table 2: Number of categories of Simulation 2.

Algorithm	Number of categories		
	Minimum	Maximum	Average
Fuzzy ART	43	55	48.4
C-FART	32	51	40.4
Proposed method	14	41	26.4

4. Conclusions

In this study, we have proposed the new type of Fuzzy ART. As important features of the proposed method, the vigilance parameter are arranged for every category, and they are varied in accordance with the size of category with learning. We have applied the proposed method into 2-dimensional input data, and the learning behaviors of the proposed method have been investigated. We have confirmed that the proposed method can categorize the input data more flexible than the standard Fuzzy ART and the conventional C-FART, and have reduced category the proliferation and the overlapped area.

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] H. Isawa, M. Tomita, H. Matsushita and Y. Nishio, "Fuzzy Adaptive Resonance Theory with Group Learning and its Applications," *Proc. of International Symposium on Nonlinear Theory and its Applications*, pp. 292-295, 2007.

[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.

[6] L. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338–353, 1965.