2007 International Symposium on Nonlinear Theory and its Applications
NOLTA'07, Vancouver, Canada, September 16-19, 2007

Fuzzy Adaptive Resonance Theory with Group Learning and its Applications

Haruka Isawa†, Masato Tomita†, Haruna Matsushita† and Yoshifumi Nishio†

†Department of Electrical and Electronic Engineering, Tokushima University,
2-1 Minami-Josanjima, Tokushima 770–8506, JAPAN
Telephone: +81–88–656–7470, Fax: +81–88–656–7471
Email:{isawa, tomita, haruna, nishio}@ee.tokushima-u.ac.jp

Abstract—Adaptive Resonance Theory (ART) is an unsupervised neural network based on competitive learning which is capable of automatically finding categories and creating new ones. Fuzzy ART is a variation of ART, allows both binary and continuous input pattern. In this study, we propose an additional step, called "Group Learning", for the Fuzzy ART in order to obtain more effective categorization. This algorithm is called Fuzzy ART with Group Learning (Fuzzy ART-GL). The important feature of the group learning is that creating connections between similar categories. In other words, the Fuzzy ART-GL learns not only categories but also its connections, namely, groups of the categories. We investigate the behavior of Fuzzy ART-GL with application to the recognition problems.

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), introduced and developed by 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. Then, we pay our attentions the Fuzzy ART of some models. Fuzzy ART is a variation of ART, incorporates the basic features of all ART systems [2]-[4]. The difference between the conventional ART and the Fuzzy ART is that the Fuzzy ART implements fuzzy logic into pattern recognition and can learn stable recognition categories in response to either analog or binary input vectors. Furthermore, input vectors are classified in each appropriate category. However, the conventional Fuzzy ART often makes input data of the common categories classify several categories. Therefore, the Fuzzy ART has the category proliferation problem.

In this study, we propose an additional step, called "Group Learning", for the Fuzzy ART in order to obtain more effective categorization. This algorithm is called Fuzzy ART with Group Learning (Fuzzy ART-GL). The group learning does not change the weight vectors of the Fuzzy ART, however, this additional step creates connections between categories. In other words, the Fuzzy ART-GL learns not only categories but also its connections, namely, groups of the categories at each step, Therefore, even if a wrong category is selected, the Fuzzy ART can modify it. This idea takes some sort of reference to Competitive Hebbian Learning proposed by Martinetz and Schulten [5] [6]. In other words, we can see that this method is the fusion of the Fuzzy ART and the Competitive Hebbian Learning. We can confirm that the Fuzzy ART-GL reduces the category proliferation problem of the conventional Fuzzy ART and increases performance.

In the Section 2, the algorithm of the conventional Fuzzy ART is introduced. In the Section 3, we explain the learning algorithm of the proposed Fuzzy ART-GL algorithm. In the Section 4, the behavior of the Fuzzy ART-GL is explained with some simulation results.

2. Fuzzy Adaptive Resonance Theory (Fuzzy ART)

2.1. Fuzzy ART

Fuzzy ART incorporates the basic features of all ART systems and implements fuzzy logic into pattern recognition.

2.2. Structure of Fuzzy ART

Fuzzy ART is composed of F_1 (input layer) and F_2 (category layer). F_1 and F_2 are connected by the bottom-up-weight vector $\boldsymbol{w}_{ij}$ and top-down-weight vector $\boldsymbol{w}_{ji}$. m neurons of the input layer F_1 correspond to the an input vector I.

Input vector: Each input $\boldsymbol{I}$ is an m-dimensional vector $\boldsymbol{I} = (i_1, i_2, \cdots, i_m)$, where each component i_i $(i = 1, \cdots, m)$ is in the interval [0, 1].

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

$$w_{j1} = \cdots = w_{jm} = 1. \quad (1)$$

Parameters: Fuzzy ART dynamics are determined by *choice parameter* $\alpha > 0$; *learning parameter* $\beta \in [0, 1]$; and *vigilance parameter* $\rho \in [0, 1]$.

2.3. Learning Algorithm of Fuzzy ART

We explain the learning algorithm of the conventional Fuzzy ART.

(FART1) An input vector $\boldsymbol{I}$ is inputted to the category layer F_2 from the input layer F_1.

(FART2) A winning category is chosen. For the input vector $\boldsymbol{I}$ and category j, *choice function* T_j can be seen as the degree of prototype $\boldsymbol{w}_j$, being fuzzy subset of $\boldsymbol{I}$.

$$T_j(\boldsymbol{I}) = \frac{|\ \boldsymbol{I} \wedge \boldsymbol{w}_j\ |}{(\alpha + |\ \boldsymbol{w}_j\ |)}, \tag{2}$$

where the fuzzy AND operator and tne norm $|\cdot|$ are defined by

$$(\boldsymbol{p} \wedge \boldsymbol{q})_i \equiv \min(p_i, q_i), \tag{3}$$

$$|\ \boldsymbol{P}\ | \equiv \sum_{i=1}^{m} |\ p_i\ |. \tag{4}$$

The winning category J whose maximal T_j is found;

$$T_J = \max\{T_j : j = 1 \cdots n\}. \tag{5}$$

If more than one T_j is maximal, the category j with the smallest index is chosen.

(FART3) The similarity of $\boldsymbol{I}$ and the current winning prototype $\boldsymbol{w}_J$ is measured by the vigilance criterion, that is, if

$$\frac{|\ \boldsymbol{I} \wedge \boldsymbol{w}_J\ |}{|\ \boldsymbol{I}\ |} \geq \rho, \tag{6}$$

w_J is updated by

$$\boldsymbol{w}_J(t+1) = \beta(\boldsymbol{I} \wedge \boldsymbol{w}_J(t)) + (1-\beta)\boldsymbol{w}_J(t), \tag{7}$$

where t is the learning step. On the contrary, Eq. (6) is not satisfied, that is, if

$$\frac{|\ \boldsymbol{I} \wedge \boldsymbol{w}_J\ |}{|\ \boldsymbol{I}\ |} < \rho, \tag{8}$$

a new index J is chosen by Eq. (5), and w_J is updated by Eq. (7). The search process continues until the chosen J satisfies Eq. (6). If all available F_2 nodes reset, new categories are established in F_2.

$$\boldsymbol{w}_{n+1} = \boldsymbol{I}. \tag{9}$$

(FART4) The steps from (FART1) to (FART3) are repeated for all the input data.

2.4. Complement Coding

Because vector element of prototype can only become smaller by adaptation, a Fuzzy ART network tends to create more and more prototype over time. This behavior is avoided by normalizing input to constant vector length, this method is called *complement coding*. The complement coded input $\boldsymbol{I}$ to the recognition system is the $2m$-dimensional vector. In the case of 2-dimensional vector $\boldsymbol{a}$,

$$\boldsymbol{I} = (\boldsymbol{a}, \boldsymbol{a}_i^c) \equiv (a_1, a_2, 1-a_1, 1-a_2). \tag{10}$$

3. Fuzzy ART with Group Learning (Fuzzy ART-GL)

3.1. Group Learning

In this study, we propose the fuzzy ART with Group Learning (Fuzzy ART-GL) which has the additional step, Group Learning (GL), for the Fuzzy ART. The group learning does not change the weight vectors of the Fuzzy ART, however, this additional step creates connections between categories. In other words, the Fuzzy ART-GL learns not only categories but also its connections, namely, groups of the categories in parallel. As important features, the Fuzzy ART-GL has a connection matrix denoted by C and the age of the connections denoted by *age*.

Connection: If category J and j are connected, $C_{J,j}$ is set from zero to one.

3.2. Learning Algorithm of Fuzzy ART-GL

Basic learning algorithm is the same as the general Fuzzy ART.

(FART-GL1) An input vector $\boldsymbol{I}$ is inputted to the category layer F_2 from the input layer F_1.

(FART-GL2) As the step (FART2), a winning category J is chosen. Furthermore, the second-winning category, denoted by J_2, is found for the group learning, namely, T_{J_2} is the second largest.

(FART-GL3) As the step (FART3), the similarity of the input $\boldsymbol{I}$ and the current winning prototype $\boldsymbol{w}_J$ is measured by the vigilance criterion by Eqs. (6) and (8). Learning ensues by Eqs. (7) and (9).

(FART-GL4) In this step, it is decided whether if a connection is formed. The similarity of the input $\boldsymbol{I}$ and the second-winning category $\boldsymbol{w}_{J_2}$ is measured by

$$\frac{|\ \boldsymbol{I} \wedge \boldsymbol{w}_{J_2}\ |}{|\ \boldsymbol{I}\ |} \geq \rho. \tag{11}$$

If Eq. (11) is satisfied, a connection between the winning category J and the second-winning category J_2 is created;

$$C_{J,J_2} = 1. \tag{12}$$

The *age* of the connection between the winning category J and the second-winning category J_2 is set to zero ("refresh" the age);

$$age_{J,J_2} = 0. \tag{13}$$

On the contrary, if Eq. (11) is not satisfied, the connection is not formed.

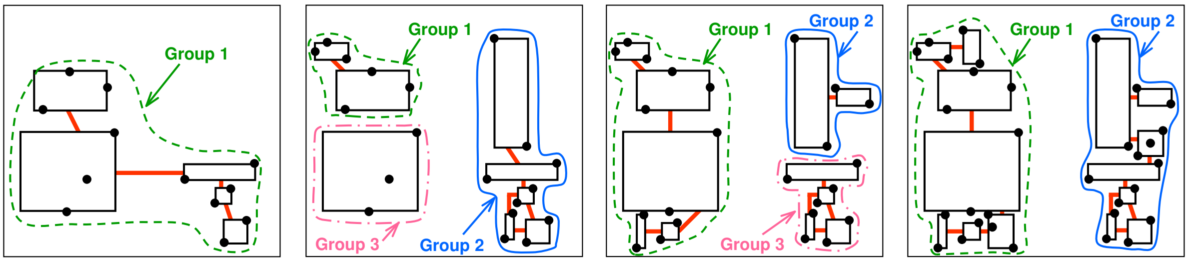
(FART-GL5) The *age* of all categories which directly connect with the winning category J are increased one;

$$age_{J,j}^{new} = age_{J,j}^{old} + 1, \qquad j \in N_J, \tag{14}$$

where N_J is the set of categories which directly connect with J.

(FART-GL6) The connections are removed, if their *age* exceeds a threshold value $AT(t)$;

$$C_{J,j} = 0, \qquad age_{(J,j)} \geq AT(t), \tag{15}$$



(a) (b) (c) (d)

Figure 1: Example of connection process. Fuzzy ART-GL learns not only the categories but also its connections, namely, groups of the categories in parallel. (a) First stage of learning. (b) Middle stage 1. (c) Middle stage 2. (d) Last stage.

$$AT(t) = AT_i(AT_f/AT_i)^{t/t_{max}}, \tag{16}$$

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

(FART-GL7) The steps from (FART-GL1) to (FART-GL6) are repeated for all the input data. Therefore, the Fuzzy ART-GL makes connections or releases connections at each step (as Fig. 1). In other words, even if a wrong category is selected, the Fuzzy ART-GL can modify it.

(FART-GL8) Finally, groups of categories are defined by categories which have connected directly or indirectly.

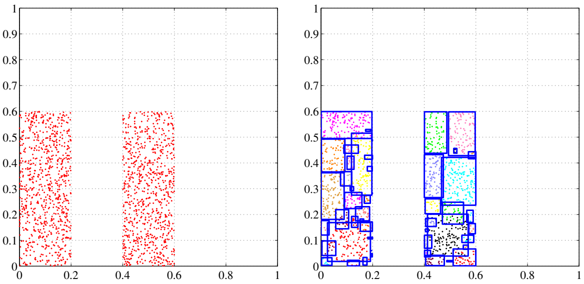
4. Learning Simulation

4.1. Simulation 1

We consider 2-dimensional input data of 1600 points, which have 2-clusters, whose distribution is non-uniform as Fig. 2(a). 800 points are distributed within a rectangular range from 0.0 to 0.2 horizontally and from 0.0 to 0.6 vertically. The remaining 800 points are distributed within a rectangular range from 0.4 to 0.6 horizontally and from 0.0 to 0.6 vertically. The parameters for the learning of the conventional Fuzzy ART and the proposed Fuzzy ART are chosen as follows;

$$\alpha = 0.1,\ \beta = 1.0,\ \rho = 0.8.$$

The learning result of Fuzzy ART is shown in Fig. 2(b). The learning result of Fuzzy ART-GL is shown in Fig. 3. Like we described in the section 2, rectangles represent categories. From these results, we can see that the category proliferation occurs with the conventional Fuzzy ART. Furthermore, Fuzzy ART has a lot of categories in one cluster, namely, Fuzzy ART makes a input data of the common category classify several categories as shown in Fig. 2(b). In contrast, the categorization result of Fuzzy ART-GL shown in Fig. 3(a) is identical with the result of the conventional Fuzzy ART, however, we can see that the proposed Fuzzy ART-GL can recognize the input data as two groups as shown in Figs. 3(b) and (c). Therefore, the input data of Fuzzy ART-GL are effectively classified in each appropriate group. Consequently, we can confirm that Fuzzy ART-GL reduces the category proliferation problem of the conventional Fuzzy ART and improves the performance.



(a) (b)

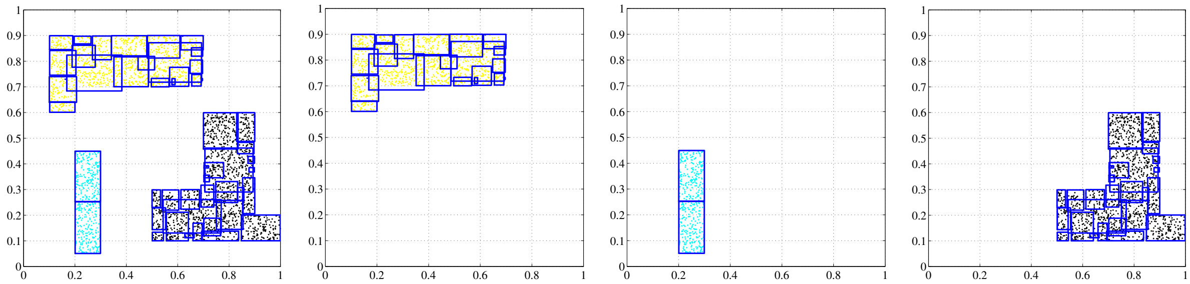
Figure 2: Simulation of Fuzzy ART for 2-clusters input data. (a) Input data. (b) Simulation result of conventional Fuzzy ART.

4.2. Simulation 2

We consider 2-dimensional input data of 3200 points, which have 3-clusters, whose distribution is non-uniform as Fig. 4(a); the top-left cluster has 1300 points, the bottom-left cluster has 400 points, and the bottom-right cluster has 1500 points. It is difficult to classify the input data such as Fig. 4(a) into appropriate categories very well. The parameters for the learning of Fuzzy ART and Fuzzy ART-GL are chosen as follows;

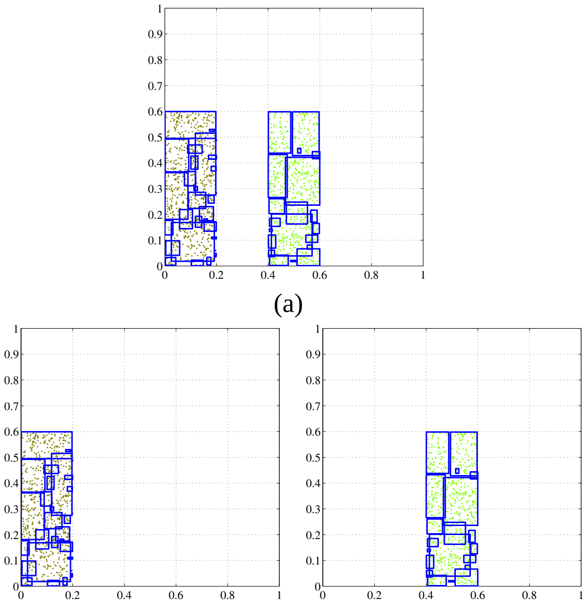
$$\alpha = 0.1,\ \beta = 1.0,\ \rho = 0.8.$$

The learning result of Fuzzy ART is shown in Fig. 4(b). The learning result of Fuzzy ART-GL is shown in Fig. 5. From these results, we can see the difference between Fuzzy ART and Fuzzy ART-GL in the Simulation 2 clearer than the Simulation 1. The conventional Fuzzy ART makes the input data of a common category classify several categories because the shape of the input data is complicated like L-shape and T-shape as shown in Fig. 4(a) In contrast, the categorization result of Fuzzy ART-GL shown in Fig. 5(a) is identical with the result of the conventional Fuzzy ART, however, the proposed Fuzzy ART effectively classify three clusters into three groups as shown in Figs. 5(b)-(d). We consider that this effective behavior is caused by making connections or releasing connections, which can modify wrong learning. Therefore, we can see



(a) (b) (c) (d)

Figure 5: Simulation of Fuzzy ART-GL for 3-clusters input data. (a) Simulation result of Fuzzy ART-GL. (b) Extracted cluster by Group 1. (c) Extracted cluster by Group 2. (d) Extracted cluster by Group 3.



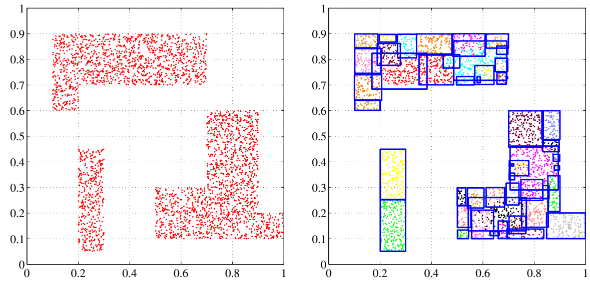
(a) (b) (c)

Figure 3: Simulation of Fuzzy ART-GL for 2-clusters input data. (a) Simulation result of Fuzzy ART-GL. (b) Extracted cluster by Group 1. (c) Extracted cluster by Group 2.

that Fuzzy ART-GL has a beneficial effect on connecting similar categories in order to obtain more effective categorization from the results of Simulation 1 and Simulation 2. Furthermore, we can confirm that Fuzzy ART-GL reduces the category proliferation problem of the conventional Fuzzy ART and improves the performance.

5. Conclusions

In this study, we have proposed the Fuzzy ART adaptation of "Group Learning". This additional step creates a connection between similar categories. Furthermore, we have investigated its behaviors with application to categorization of 2-dimensional input data and we have confirmed the efficiency of Fuzzy ART-GL. In the future, we will classify higher-dimensional input data and moving input data.



(a) (b)

Figure 4: Simulation of Fuzzy ART for 3-clusters input data. (a) Input data. (b) Simulation result of conventional Fuzzy ART.

Acknowledgments
This work was partly supported by Yamaha Music Foundation.

References

[1] 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.

[2] G. A. Carpenter, S. Grossberg, and D. B. Rosen, N. Markuzon, J. H. Reynolds, D. B. Rosen, "Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of Analog Multidimensional Maps," *Neural Networks*, vol. 3,pp. 698-713, 1991. Tech. Rep. CAS/CNS-TR-91-016. Boston, MA: Boston University.

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

[4] 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, pp. 544-559, 1998.

[5] T. M. Martinetz and K. J. Schulten, "A "neural-gas" network learns topologies," *Artificial Neural Networks*, pp. 397-402, 1991.

[6] T. M. Martinetz and K. J. Schulten, "Topology representing networks," *Neural Networks*, vol. 7, no. 3, pp. 507-522, 1994.