8-2013

Vigilance adaptation in adaptive resonance theory

Lei MENG

Ah-hwee TAN
Singapore Management University, ahtan@smu.edu.sg

Donald C. WINSCH

Published in 2013 International Joint Conference on Neural Networks IJCNN: August, 4-9, Dallas, TX: Proceedings.
DOI: 10.1109/IJCNN.2013.6706857

Vigilance Adaptation in Adaptive Resonance Theory

Lei Meng, Ah-Hwee Tan and Donald C. Wunsch II

***Abstract*— Despite the advantages of fast and stable learning, Adaptive Resonance Theory (ART) still relies on an empirically fixed vigilance parameter value to determine the vigilance regions of all of the clusters in the category field (F_2), causing its performance to depend on the vigilance value. It would be desirable to use different values of vigilance for different category field nodes, in order to fit the data with a smaller number of categories. We therefore introduce two methods, the Activation Maximization Rule (AMR) and the Confliction Minimization Rule (CMR). Despite their differences, both ART with AMR (AM-ART) and with CMR (CM-ART) allow different vigilance levels for different clusters, which are incrementally adapted during the clustering process. Specifically, AMR works by increasing the vigilance value of the winner cluster when a resonance occurs and decreasing it when a reset occurs, which aims to maximize the participation of clusters for activation. On the other hand, after receiving an input pattern, CMR first identifies all of the winner candidates that satisfy the vigilance criteria and then tunes their vigilance values to minimize conflicts in the vigilance regions. In this paper, we chose Fuzzy ART to demonstrate these concepts, but they will clearly carry over to other ART architectures. Our comparative experiments show that both AM-ART and CM-ART improve the robust performance of Fuzzy ART to the vigilance parameter and usually produce better cluster quality.**

I. Introduction

Adaptive Resonance Theory (ART) [10] is a neural theory of cognitive information processing, which states that fast learning is a resonant phenomenon in neural circuits. This has led to the development of the ART 1 neural network model [1] for unsupervised learning. ART 1 performs unsupervised learning by modeling clusters as memory prototypes and encoding binary input patterns incrementally through a two-way similarity measure for searching suitable clusters, which simulates how a human brain captures, recognizes and memorizes information regarding objects and events. As long as the difference between the input pattern and the selected winner cluster from the category field does not exceed a certain threshold, called the vigilance parameter, the input pattern is considered a member of the winner cluster. Fuzzy ART [2] replaces the intersection operator ($\cap$) in ART 1 by the min operator ($\wedge$) of fuzzy set theory so as to learn both binary and analog patterns. Fuzzy ART inherits the advantages of ART 1 including fast and stable learning and the incremental clustering.

Although Fuzzy ART and its variants are useful for unsupervised learning in many areas [3], [4], [5], they require manual value selection for the vigilance parameter. Specifically, Fuzzy ART still relies on an empirically fixed vigilance value to scale the cluster size, which makes its performance highly dependent on the vigilance parameter value. For example, with a small vigilance value, Fuzzy ART permits high generalization, which may lead to the generation of several big clusters mixed with patterns from multiple classes. On the other hand, a large vigilance value may incur the over-generation of clusters such that one class may be represented by multiple small clusters. Therefore, similar to the selection of the number of clusters in the K-means clustering algorithm [9], selecting a suitable vigilance value for Fuzzy ART poses a great challenge.

For adapting the vigilance parameter, a match tracking rule has been proposed in ARTMAP [6], which may adapt the vigilance value during the cluster selection process for the input pattern. However, the class label for each input pattern should be used to identify an incorrect classification, which is not available for unsupervised learning. He et al. has proposed one approach [11], [12], called ART under Constraint (ART-C), for tuning the vigilance parameter in ART 1, ART 2 [13], ART 2A [14] and Fuzzy ART. However, ART-C requires a user-predefined number of clusters so that the selection of vigilance parameter is transformed to that of the number of clusters. Therefore, under the original ART clustering paradigm, there is still no work on the vigilance adaptation task.

In view of this issue, this paper presents two methods, the Activation Maximization Rule (AMR) and the Confliction Minimization Rule (CMR), for adapting the vigilance parameter to alleviate its effect to the performance of Fuzzy ART. Despite the differences between the two models, both Fuzzy ART with AMR (AM-ART) and with CMR (CM-ART) allow different vigilance levels for different clusters, which are incrementally adapted during the clustering process. However, they work in different ways. AMR maximizes the cluster activation to promote the participation of clusters in encoding patterns. Specifically, AMR increases the vigilance value of the winner cluster when a resonance occurs and decreases it when a reset occurs. This rule helps to alleviate cases in which the initial vigilance value is too small, causing patterns belonging to different classes always to be encoded by the same cluster, or too large, leading to the generation of many small clusters. On the other hand, CMR minimizes the conflicts between the vigilance regions of clusters to produce better cluster boundaries. Different from traditional winner selection procedures in Fuzzy ART, the resulting CM-ART first identifies all of the winner candidate clusters that satisfy the vigilance criteria and then increases the vigilance values of all of the candidates, except for the winner cluster, to

Lei Meng and Ah-Hwee Tan are with the School of Computer Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore, 639798. Email: {meng0027, asahtan}@ntu.edu.sg. Donald C. Wunsch II is with the Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, Mo65409, USA. Email: dwunsch@mst.edu.

violate the vigilance criteria so as the reduce the overlap between these clusters.

The performances of AM-ART and CM-ART have been evaluated on a subset of the public image data set NUS-WIDE [7] in terms of the weighted average precision, network complexity and convergence analysis. Our comparative experiments show that both AM-ART and CM-ART improve the robustness of Fuzzy ART to the change in vigilance parameter and usually produce better cluster quality. We also find that AM-ART achieves convergence performance comparable to Fuzzy ART and that CM-ART converges faster than Fuzzy ART.

The remainder of the paper is organized as follows. Section 2 reviews related works on the tuning of vigilance parameter. Section 3 presents a summary of the Fuzzy ART algorithm. Section 4 and Section 5 present the detailed algorithm of Fuzzy ART with AMR and CMR, respectively. The experimental results are presented in Section 5. The last section concludes our work.

II. Related Work

ARTMAP is two-channel neural network, with one input field for the input patterns and the other one for the class labels, for supervised learning. The match track rule was firstly used in the search and prediction process of ARTMAP to maximize code compression which fits the data with a minimum number of cluster nodes. Specifically, at the beginning of each input pattern presentation, the vigilance parameter is set to a baseline. When the template matching procedure, which evaluates the degree of similarity between an input pattern and the selected winner cluster, causes a resonance for the input pattern channel but causes a reset for that of its class label, a change in the vigilance value is triggered such that the vigilance value is increased to just larger than the match function value of the input pattern. The search process then selects another winner from the rest of the F_2 clusters under the revised vigilance criteria. In this way, the vigilance value is self-adapted to help the winner to reject the patterns from other classes.

ART under constraint (ART-C) was firstly proposed in [11] to allow a series of ART models, including ART 1, ART 2 and Fuzzy ART, to be able to produce a fixed number of clusters in the category field F_2. Specifically, when the number of clusters exceeds the limit, a pair of clusters in the category field will be selected according to a match score. Then, these two clusters are merged and the vigilance parameter value is set to the match score. The ART-C model is further applied for ART 2A in [12]. The key difference between these two work is that in [12], the similarity between a pair of F_2 clusters is defined by the dot product instead of the function for match score in [11].

Based on the above discussion, we find that ARTMAP requires the class labels for the patterns to tune the vigilance parameter and ART-C should introduce a pre-defined number of clusters to adapt it. Therefore, under the original ART clustering paradigm, there is no work on the adaptation of the vigilance parameter.

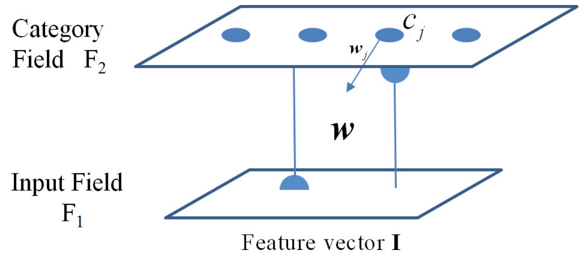


Fig. 1: The Fuzzy ART architecture.

III. Fuzzy ART

The architecture of Fuzzy ART (Fig. 1) comprises one input field F_1 for receiving the input patterns and one category field F_2 for the clusters. The generic network dynamics of Fuzzy ART are described as follows.

Input vectors: Let $\mathbf{I}$ denote the input pattern. With complement coding [2], $\mathbf{I}$ is further augmented with a complement vector $\bar{\mathbf{I}}$ such that $\bar{\mathbf{I}}_i = 1 - \mathbf{I}_i$ in the input field F_1.

Weight vectors: Let $\mathbf{w}_j$ denote the weight vector associated with the j-th cluster c_j $(j = 1, ..., J)$ in the category field F_2.

Parameters: The Fuzzy ART's dynamics is determined by choice parameter $\alpha > 0$, learning parameter $\beta \in [0, 1]$ and vigilance parameter $\rho \in [0, 1]$.

The clustering process of Fuzzy ART comprises three key steps:

1) Category choice: For each input pattern $\mathbf{I}$, fuzzy ART calculates the choice function for all of the clusters in the category field F_2 and select the most suitable cluster (winner cluster) c_{j^*}, which has the largest value. The choice function for the j-th cluster c_j is defined by

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

where the fuzzy AND operation $\wedge$ is defined by $(\mathbf{p} \wedge \mathbf{q})_i \equiv \min(\mathbf{p}_i, \mathbf{q}_i)$, and the norm $|.|$ is defined by $|\mathbf{p}| \equiv \sum_i \mathbf{p}_i$.

2) Template matching: The similarity between the input pattern $\mathbf{I}$ and the winner c_{j^*} is evaluated using a match function M_{j^*}, which is defined by

$$M_{j^*} = \frac{|\mathbf{I} \wedge \mathbf{w}_{j^*}|}{|\mathbf{I}|}. \tag{2}$$

If the winner satisfies the vigilance criteria such that $M_{j^*} > \rho$, a resonance occurs which leads to the learning step. Otherwise, a new winner is selected from the rest of the clusters in the category field. If no winner satisfies the vigilance criteria, a new cluster is generated to encode the input pattern.

3) Prototype learning: If c_{j^*} satisfies the vigilance criteria, its corresponding weight vector $\mathbf{w}_{j^*}$ is updated through a learning function, defined by

$$\mathbf{w}_{j^*}^{(new)} = \beta(\mathbf{I} \wedge \mathbf{w}_{j^*}) + (1 - \beta)\mathbf{w}_{j^*}. \tag{3}$$

IV. FUZZY ART WITH AMR

Through the summary of the Fuzzy ART algorithm in the previous section, we find that the clustering strategy of Fuzzy ART can be seen as a selection of the most "similar" cluster, called a winner cluster, from the category field to the input pattern defined by the choice function (1), which evaluates to which degree the input pattern matches the winner cluster. The winner is subsequently evaluated for its degree of similarity with the input pattern through the match function (2) and the vigilance parameter, which determines the similarity threshold for the vigilance criteria. Therefore, the vigilance parameter affects the classification of patterns by thresholding the vigilance regions of clusters. With a small vigilance value, patterns, even from different classes, are likely to incur a resonance for the same cluster. On the other hand, a large vigilance value may lead to the reset of input patterns for all clusters in the category field so that a new cluster should be created to encode the input pattern, which may result in the generation of an overly complicated network.

To alleviate such problems, the activation maximization rule (AMR) is proposed, which, different from Fuzzy ART, allows different vigilance levels for different clusters in the category field and incrementally tunes the vigilance value for the winner after a resonance or reset occurs. Specifically, when a resonance occurs, AMR increases the vigilance value of the winner cluster and, in contrast, decreases it when a reset occurs. In this way, AMR may help to improve the clustering performance when the initial vigilance value is not suitable. For example, given a data set, when the initial vigilance value is small, the vigilance value of the winner cluster will be increased gradually to restrain its activation to the input pattern. Similarly, when the vigilance value is large, the vigilance value of the winner cluster will be gradually decreased to promote its activation. Note that AMR restrains the continuous activation of the same cluster and promotes the activation of clusters with a large vigilance value. Therefore, AMR also helps to even out the size of clusters representing the same class which helps to prevent the generation of small clusters, which contains limited number of patterns.

The complete algorithm of Fuzzy ART with AMR (AM-ART) is summarized as follows. Similar to the original Fuzzy ART, the vigilance parameter of the uncommitted cluster is set to ρ_0, which is the initial vigilance value. The key difference between AM-ART and Fuzzy ART appears in step 6) and step 8), wherein AM-ART allows each cluster to have its own vigilance value, which is tuned by a restraint parameter $\sigma \in [0,1]$. A small σ leads to a small change in the vigilance value of the winner cluster, which has a small influence on the effect of the initial vigilance value on the performance of AM-ART. On the other hand, a large σ may help to make AM-ART more robust to the change in the initial vigilance value, but may also result in unstable vigilance regions of clusters, which may decrease the clustering performance.

Clustering algorithm of AM-ART

1) Create an uncommitted cluster c_1 with all weight vectors containing 1's in the category field F_2 and set $\rho_1 = \rho_0$.
2) Receive an input pattern $\mathbf{I}$, normalize it with complement coding such that $\hat{\mathbf{I}} = [\mathbf{I}, \bar{\mathbf{I}}]^\top$, and then present it into the input field F_1.
3) For each cluster c_j in the category field F_2, calculate the choice function defined in (1).
4) Identify the winner c_{j^*} with the largest value of the choice function such that $j^* = \arg\max_{j:c_j \in F_2} T_j$.
5) Calculate the match function M_{j^*} defined in (2).
6) If $M_{j^*} < \rho_{j^*}$, a reset occurs. Set $\rho_{j^*}^{(new)} = (1-\sigma)\rho_{j^*}$. Set $T_{j^*} = 0$ and go to 4; Otherwise, a resonance occurs, so go to 7.
7) If c_{j^*} is uncommitted, set the cluster weight to the input pattern such that $\mathbf{w}_{j^*} = \mathbf{I}$ and set $\rho_{j^*} = \rho_0$. Create a new uncommitted node and go to 9; otherwise, go to 8.
8) Update $\mathbf{w}_{j^*}$ according to (3) and set $\rho_{j^*}^{(new)} = (1+\sigma)\rho_{j^*}$. Go to 9.
9) If no input pattern exists, the algorithm stops. Otherwise, go to 2.

V. FUZZY ART WITH CMR

As described in the previous section, AMR tunes the vigilance parameters of clusters by preventing cases caused by an inappropriate vigilance value. From another perspective, in Fuzzy ART, the incorrect recognition of patterns from different classes is usually caused by a small vigilance value. For example, given two classes A and B, which lie near each other in the feature space, a new cluster of class A in the margin between class A and B, with a small vigilance value, may move to class B by encoding and learning the patterns of class B. As thus, it subsequently becomes competitive with the clusters of class B. Therefore, the overlap of vigilance boundaries of clusters in the feature space increases the risk of pattern misclassification.

To address this problem, we propose a confliction minimization rule (CMR) to reduce the overlap of vigilance regions of clusters in the feature space. Different from traditional winner search procedures of Fuzzy ART, the resulting Fuzzy ART with CMR (CM-ART) first identifies all of the winner candidates to the input pattern through the match function and subsequently reduces the overlap between the vigilance regions of these candidate clusters so as to achieve a local minimization of their competitive conflicts.

The complete algorithm of CM-ART is summarized as follows. Different from AM-ART, which first uses the choice function to identify a winner, CM-ART identifies all clusters that satisfy the vigilance criteria, and then uses the choice function to identify the winner. In step 6), the vigilance values for all of the winner candidates, except the winner itself, are updated to minimize the overlap between vigilance regions. Recall that $\frac{|\mathbf{I} \wedge \mathbf{w}_j|}{|\mathbf{I}|}$ is the value of the match function, and $\Delta > 0$ is a very small number. Therefore, the purpose of this updating equation, similar to the match tracking rule proposed in [6], is to shrink the vigilance regions of the winner candidates to violate the vigilance criteria for the input pattern.

Clustering algorithm of CM-ART

1) Create an uncommitted cluster c_1 with all weight vectors containing 1's in the category field F_2, and set $\rho_1 = \rho_0$.
2) Receive an input pattern $\mathbf{I}$, normalize it with complement coding such that $\hat{\mathbf{I}} = [\mathbf{I}, \bar{\mathbf{I}}]^\top$ and then present it into the input field F_1.
3) For all the clusters c_j ($j = 1, ..., J$) in the category field F_2, calculate the match function defined in (2). Select the clusters that satisfy the vigilance criteria as winner candidates such that $M_j \geq \rho_j$.
4) For all of the candidates, calculate the choice function defined in (1) and identify the winner c_{j^*} with the largest value such that $j^* = \arg\max_{j:c_j \in F_2} T_j$.
5) If c_{j^*} is uncommitted, set the cluster weight to the input pattern such that $\mathbf{w}_{j^*} = \mathbf{I}$ and set $\rho_{j^*} = \rho_0$. Create a new uncommitted node and go to 7; otherwise, a resonance occurs, so go to 6.
6) Update $\mathbf{w}_{j^*}$ according to (3). Set the corresponding vigilance values of the rest of the winner candidates using $\rho_j^{(new)} = \frac{|\mathbf{I} \wedge \mathbf{w}_j|}{|\mathbf{I}|} + \Delta$. Go to 7.
7) If no input pattern exists, the algorithm stops. Otherwise, go to 2.

VI. Experiments

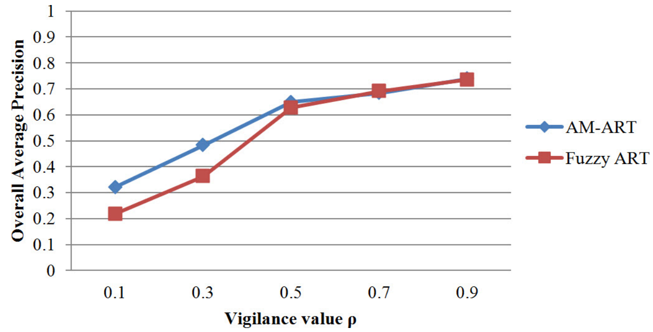
A. NUS-WIDE data set

To evaluate the clustering performance of our proposed methods, we collected a total of 1000 images from a real-world web image set, called the NUS-WIDE data set [7]. The images belonged to the five biggest classes in this data set, including dog, bear, bird, sunset and wedding, each of which contained 200 images.

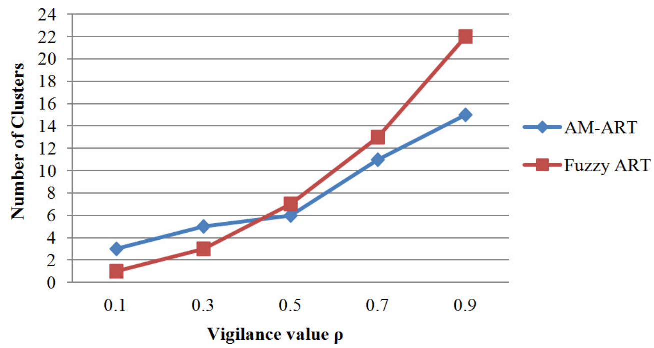
For the feature representation, we used a concatenation of three types of visual features, including Grid Color Moment (255 features), Edge Direction Histogram (73 features) and Wavelet Texture (128features). These global features can be efficiently extracted and have been shown to be effective for image content representation [7]. As thus, each image was represented by a vector of 426 features.

B. Experiments on AM-ART

1) Performance Comparison of AM-ART with Fuzzy ART: In the experiments, we set the following values: choice parameter $\alpha = 0.01$, learning rate $\beta = 0.6$, initial vigilance parameter ρ_0 ranging from 0.1 to 0.9, and restraint parameter $\sigma = 0.1$. A small choice parameter of $\alpha = 0.01$ is commonly used, as it has been shown that the clustering performance is generally robust to this parameter [8]. We empirically used $\beta = 0.6$ to tune the cluster weight towards the geometric center of the cluster. In our experiments, the performance of GHF-ART remained roughly the same when the learning parameter changed from 0.2 to 0.8. Fig. 2 shows the clustering results of AM-ART and Fuzzy ART with changing vigilance values. The overall average precision and the number of generated clusters were used to evaluate



(a)



(b)

Fig. 2: Clustering results of AM-ART and Fuzzy ART with changing vigilance values in terms of (a) overall average precision and (b) number of generated clusters.

performance. Fig. 2(a) shows that the performance of AM-ART is improved over Fuzzy ART by 10% when ρ increases to 0.1 and 0.3. After $\rho = 0.5$, the performance of AM-ART and Fuzzy ART becomes similar. However, Fig. 2(b) indicates that AM-ART greatly improves Fuzzy ART in terms of the number of generated clusters when ρ is larger than 0.5. Moreover, we can see that the increase in both the number of generated clusters of AM-ART and the vigilance value are not as significant as in Fuzzy ART. This demonstrates that, when the vigilance value is small, AM-ART may improve the performance of Fuzzy ART by enhancing its ability to identify more classes, while, when the vigilance value is large, AM-ART may alleviate the over-generation of clusters.

We also investigated the clustering structure to see how AMR helps to improve the cluster quality. Specifically, we studied the performance and the number of generated clusters of AM-ART and Fuzzy ART on specific classes when $\rho = 0.3$ and $\rho = 0.9$, which have the biggest gap in average precision and number of clusters respectively. The results are summarized in Table I(a), which indicates that without AMR, Fuzzy ART fails to identify the patterns from the classes "bear" and "bird" which are buried into the clusters belonging to the other three classes, leading to a poor cluster quality. With AMR, the performance of AM-ART greatly improves on the classes "dog", "sunset" and "wedding". Moreover, the patterns belonging to the class "bear" can be recognized to dominate a cluster. In Table I(b), we observe that, the overall performance of AM-ART in average precision is similar to that of Fuzzy ART while the

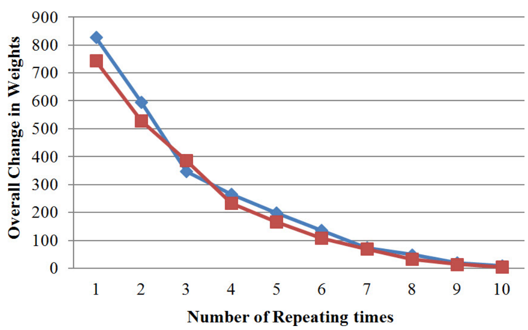
TABLE I: Clustering performance of AM-ART and Fuzzy ART when $\rho = 0.3$ and $\rho = 0.9$ in terms of five classes.

$\rho = 0.3$		dog	bear	bird	sunset	wedding	overall
AM-ART	Average Precision	0.5206	0.3617	-	0.4623	0.5648	0.4824
	Number of Clusters	2	1	0	1	1	5
Fuzzy ART	Average Precision	0.3288	-	-	0.3428	0.4206	0.3629
	Number of Clusters	1	0	0	1	1	3

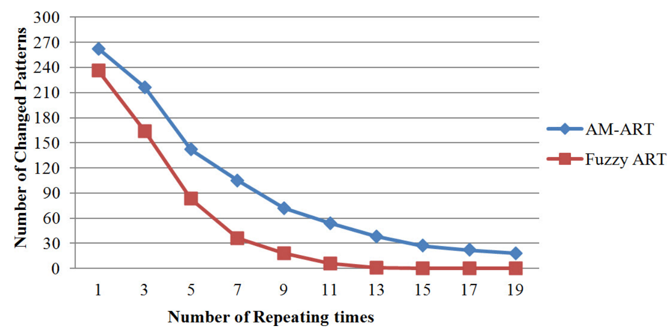
(a) The vigilance parameter $\rho = 0.3$

$\rho = 0.9$		dog	bear	bird	sunset	wedding	overall
AM-ART	Average Precision	0.7848	0.7306	0.6680	0.7244	0.8166	0.7392
	Number of Clusters	4	3	4	2	2	15
	Minimum Recall	0.2466	0.3142	0.1617	0.3774	0.2585	0.2836
Fuzzy ART	Average Precision	0.7626	0.7541	0.7027	0.6817	0.8224	0.7348
	Number of Clusters	6	5	6	3	2	22
	Minimum Recall	0.1446	0.1649	0.0947	0.1667	0.2193	0.1482

(b) The vigilance parameter $\rho = 0.9$



(a)



(b)

Fig. 3: Total change of AM-ART in (a) weight values and (b) number of changed patterns along with the repetition of the data set.

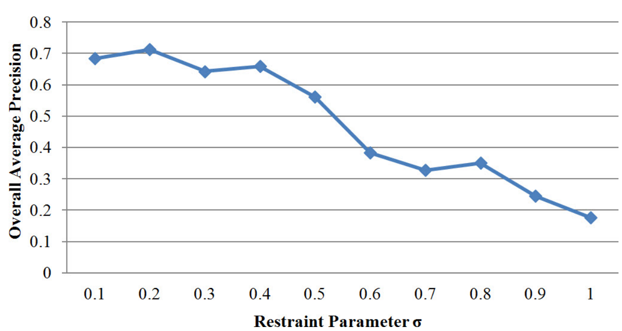


Fig. 4: Clustering Performance of AM-ART with different restraint values.

number of clusters decreases significantly. We also recorded the minimum recall for each class. The results show that AM-ART achieves a much higher performance in minimum recall than Fuzzy ART across all classes. This demonstrates that the proposed AMR can help to prevent the generation of small clusters.

2) Convergence Analysis of AM-ART: We studied the convergence of AM-ART by evaluating the change in cluster weights and patterns during the repetition of the whole data set. We followed the parameter settings as used in the previous section and set the initial vigilance value to 0.7. As shown in Fig. 3(a), AM-ART and Fuzzy ART perform comparably for convergence in cluster weights. Specifically, both the weights of AM-ART and Fuzzy change significantly at the start and remains high before the fifth repetition of the data set. This is due to the newly generated clusters which causes an unstable data structure. After that, the change becomes smooth and is only 7.84 after the tenth run. Finally, after 27 runs of the data set, there is no change in the weights of AM-ART, while fuzzy ART converges after 16 runs.

In Fig. 3(b), we similarly observe that the number of changed patterns of AM-ART remains high and decreases quickly over the first five runs. However, even after 19 repetitions, 22 patterns continue to change clusters. In contrast, Fuzzy ART converges much faster and does not change after 15 runs.

3) Sensitivity of AM-ART to restraint parameter: To evaluate the robustness of AM-ART to the restraint parameter, we varied restraint parameter value from 0.1 to 1 to study how it affects the clustering performance. The results shown in Fig. 4 show that the performance of AM-ART remains relatively constant when the restraint parameter changes from 0.1 to 0.4. Then, the performance decreases significantly

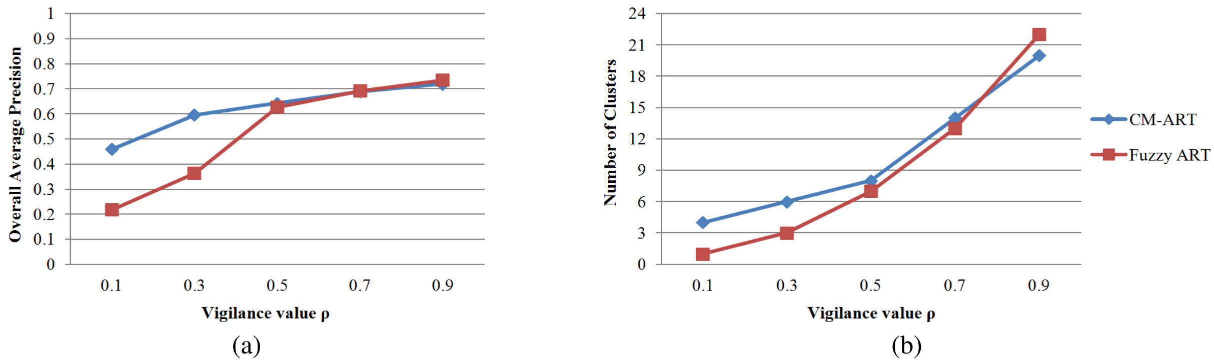


Fig. 5: Clustering results of CM-ART and Fuzzy ART with changing vigilance values in terms of (a) overall average precision and (b) number of generated clusters.

when the restraint value is between 0.4 and 0.6. After that, when the value is 0.9 and 1, the performance remains low and is much worse. This demonstrates that a small restraint value ranging from 0.1 to 0.4 may lead to a satisfactory result.

C. Experiments on CM-ART

1) Performance Comparison of CM-ART with Fuzzy ART: We followed the parameter settings in the experiments on AM-ART. We first compared the performance between CM-ART and Fuzzy ART in terms of the overall performance and the number of clusters, as summarized in Fig. 5. Fig. 5(a) shows that the performance of CM-ART increase greatly when the vigilance value is below 0.5. After that, the performance of CM-ART and Fuzzy ART are similar. More importantly, CM-ART provides a relatively smooth increase along with the increase in the vigilance value, demonstrating that CM-ART may largely enhance the robustness of Fuzzy ART to the changing vigilance values. In Fig. 5(b) shows that CM-ART may identify more clusters than Fuzzy ART when the vigilance value is very small. When the vigilance value is larger than 0.5, the number of clusters generated by CM-ART and Fuzzy ART are similar. However, CM-ART generates fewer clusters than Fuzzy ART when the vigilance value is larger than 0.8. This may due to the fact that CM-ART builds better cluster boundaries which may help to recognize patterns and reduce the chance of creating redundant clusters that are small and that consist of a mix of patterns from different classes.

2) Convergence Analysis of CM-ART: Similar to the convergence evaluation for AM-ART, we studied the convergence of CM-ART by repeating all of the patterns in the data set and recording the changes in cluster weights and the number of changed patterns. The results are shown in Fig. 6. In contrast to AM-ART in Fig. 3, we can observe that, in Fig. 6(a), CM-ART converges faster than Fuzzy ART in terms of the change in weights which approaches convergence after only 8 runs of the data set. Moreover, the total change in weights of CM-ART is much smaller than in Fuzzy ART.

Similarly, Fig. 6(b) indicates that CM-ART triggers a smaller change in the number of changed patterns than Fuzzy ART, and also converges faster than Fuzzy ART to achieve

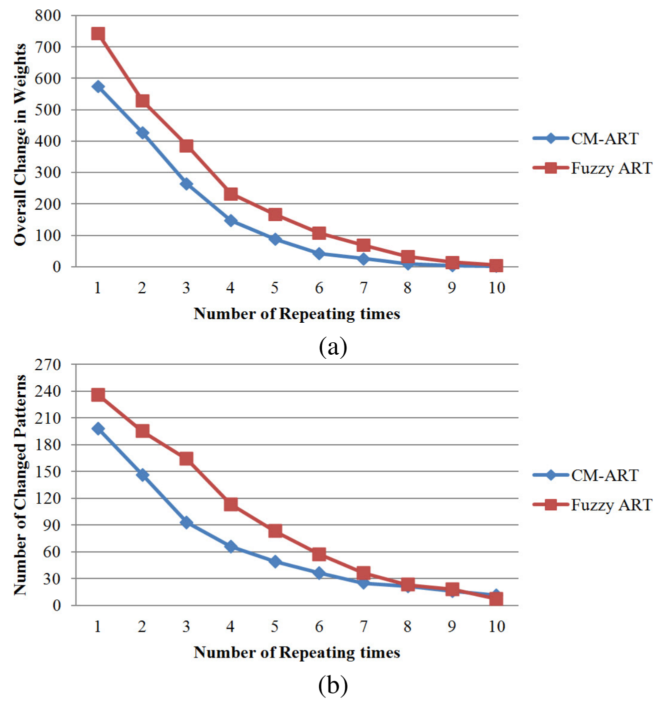


Fig. 6: Total change of CM-ART in (a) weight values and (b) number of changed patterns along with the repetition of the data set.

a stable cluster structure. The number of changed patterns of CM-ART decreases to a small level after the fifth run, while Fuzzy ART reqires two more runs to achieve the same level.

VII. Conclusions

This paper presented two methods, the Activation Maximization Rule (AMR) and Confliction Minimization Rule (CMR), for the incremental adaptation of the vigilance parameter in a Fuzzy ART model during the clustering process. In contrast with Fuzzy ART, both of the resulting clustering models, AM-ART and CM-ART, allow different vigilance levels for different clusters but tune the vigilance parameter from different perspective.

AM-ART employs an activation maximization rule (AMR) to prevent cases in which the input patterns are continuously encoded by the same cluster, which may indicate an excessively small vigilance value, and cases in which the input patterns rarely incur a resonance, which may indicate a vigilance value that is too large. The experimental results show that the proposed AM-ART can help to improve the performance of Fuzzy ART, especially when the vigilance value is far from suitable. Additionally, AMR may help to prevent the generation of small clusters by evening out the size of clusters representing the same class so as to reduce the number of small clusters.

CM-ART minimizes the overlap of vigilance regions of clusters, aiming to improve the clustering performance of Fuzzy ART by producing better cluster boundaries during the clustering process. From the experimental results, we find that the proposed CM-ART may largely alleviate the sensitivity of Fuzzy ART to the change of vigilance parameter so as to improve the robustness of Fuzzy ART. Moreover, CM-ART converges faster than Fuzzy ART to achieve a stable cluster structure.

Comparing the experimental results of CM-ART with that of AM-ART, we find that, under the same parameter settings, CM-ART achieves better performance in terms of average precision and has a better effect on improving the robustness of Fuzzy ART to the change of vigilance parameter. However, AM-ART has a better performance in preventing the generation of complicated network when the vigilance parameter has to be set to a large value.

Acknowledgment

This research is supported in part by Interactive and Digital Media Programme Office (IDMPO), National Research Foundation (NRF) hosted at Media Development Authority (MDA) under Grant No.: MDA/IDM/2012/8/8-2 VOL 01.

References

[1] G. A. Carpenter and S. Grossberg, "A Massively Parallel Architecture for a Self-organizing Neural Pattern Recognition Machine," *Computer Vision, Graphics, and Image Processing*, pp. 54-115, 1987.

[2] G. A. Carpenter, S. Grossberg and 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] A.-H. Tan, H.-L. Ong, H. Pan, J. Ng and Q Li, "Towards Personalized Web Intelligence," *Knowledge and Information Systems*, pp. 595-616, 2004.

[4] T. Jiang and A.-H. Tan, "Learning Image-Text Associations," *IEEE Transctions on Knowledge and Data Engineering*, vol. 21, no. 2, pp. 161-177, 2009.

[5] L. Meng and A.-H. Tan, "Semi-supervised Hierarchical Clustering for Personalized Web Image Organization," *International Joint Conference on Neural Networks*, pp. 2859-2866, 2012.

[6] G. A. Carpenter, S. Grossberg and J. H. Reynolds, "ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network," *Neural Networks*, vol. 4, pp. 565-588, 1991.

[7] T.-S. Chua, J. Tang, R. Hong, H. Li, Z. Luo, Y. Zheng, "NUS-WIDE: A Real-World Web Image Database from National University of Singapore," *In CIVR*, pp. 1-9, 2009.

[8] A.-H. Tan, "Adaptive Resonance Associative Map," *Neural Networks*, pp. 437-446, 1995.

[9] J. MacQueen, "Some methods for classification and analysis of multivariante observations," *In proc. 5th Berkely Symp. Mathematics and Probability*, pp. 281-297, 1967.

[10] S. Grossberg, "How does a brain build a cognitive code?" *Psychological Review*, pp. 1-51, 1980.

[11] J. He, A.-H. Tan, and C.-L. Tan, "ART-C: A neural architecture for self-organization under constraints," *International Joint Conference on Neural Networks*, pp. 2550-2555, 2002.

[12] J. He, A.-H. Tan, and C.-L. Tan, "Modified ART 2A Growing Network Capable of Generating a Fixed Number of Nodes," *IEEE Transections on Neural Networks*, vol. 15, no. 3, pp. 728-737, 2004.

[13] G. A. Carpenter and S. Grossberg, "ART 2: Stable selforganization of pattern recognition codes for analog input patterns," *Applied Optics*, pp. 4919-4930, 1987.

[14] G. A. Carpenter, S. Grossberg and and D. Rosen, "ART 2-A: An adaptive resonance algorithm for rapid category learning and recognition," *Neural Networks*, vol. 4, pp. 493-504, 1991.