Match Adaptive Resonance Theory Neural Network for Arabic Alphabet Recognition

HOMSI MASUN AND SUKKAR FADEL
Department of Computer Science
University of Aleppo
P.O Box 6502
ALEPPO, SYRIA

Abstract: Match Adaptive Resonance Theory (MART2) is developed as a modified version of self-organising Adaptive Resonance Theory (ART2) neural network for Arabic alphabet recognition. The new model does not utilise bi-directional synapses, match-reset loops and vigilance parameter. Novel subsystem is added to select the winning F2 node conserving competitive learning concept applied to reset wave. It relies on different sequence of operations of ART2 algorithm, but the classification of the input patterns remains unchanged. In the new architecture, algorithm execution takes almost equal time for each input pattern to be clustered and it has a new strategy in accessing an appropriate node in F2 without having bottom-up connections, generated from F1 to F2. However, top-down connections play an important role in matching and resonance.

MART2 classifier of Arabic letters signals is implemented. The raw input signals are segmented and preprocessed depending on two criterions, amplitude average and zero crossings rate which determine voiced and unvoiced frames. Fast Fourier Transform (FFT) is used to transform the signals from time domain to frequency domain. The most important features of the letters are extracted to reduce data size. The reduced data are then presented to MART2 for training and classification.

ART2 and MART2 are employed for clustering Arabic letters. Experimental results show that the new algorithm of MART2 generally exhibits faster learning, better clustering performance, lower error level, an improved recognition ability and more accuracy; even without a need of bottom up connections, match-reset loops and a vigilance parameter. That is, a major advantage of a flexible adaptive resonance theory neural network.

Key-Words: Self-oraganising, Adaptive Resonance Theory, Match Adaptive Resonance Theory, Long Term Memory, Short Term Memory, Fast Fourier Transform.

1 Introduction

Artificial self-organising neural networks are simplified models of central nervous system. They are networks of highly interconnected neural computing elements that have the ability to process input stimuli and adapt to the environment by retaining useful facts and information in memory. It also has the capacity to conserve learning new and important information. Adaptive Resonance Theory (ART) is a good example of such neural network [1].

Adaptive Resonance Theory (ART2) developed by Carpenter and Grossberg plays an important role in pattern recognition and signal identification problems [2]. It is capable of fast and stable learning of clustering arbitrary sequence of input patterns which share some similarities [3]. Speech is considered as an important mean for human beings to communicate with each other. Man always tries to develop new technologies imitating such complicated system of speech. In order to mimic the human behaviour in distinguishing and categorising different Arabic letters, a new classifier system is proposed. The classifier consists of three main components. The first component is a preprocessing of the raw speech signal, which captures essential information from noisy data. A pattern transformation, is the second component that takes the processed input data in time domain and transforms it via Fast Fourier Transform (FFT) to alternative representation in frequency domain. The obtained signal is then presented to the third component of the system for higher level processing. This high level processor is implemented as an artificial self-organising neural network using new technique.

Match Adaptive Resonance Theory MART2 is developed as a different version of ART2 architecture. The network entered into resonant state if a category is found with required matching level; otherwise, it generates a new category node that learns the current input.

The elimination of bottom-up adaptive filters, match-reset loops and vigilance parameter are variants introduced to ART2 to make clustering mechanism more flexible. MART2 has new method in accessing an appropriate node in F2 without having bottom-up connections, generated from F1 to F2. Additionally, it conserves competitive learning concept applied to reset waves. The new architecture (MART2) is successfully applied to Arabic letters recognition. Results show the effectiveness of the new technique.

This paper is organised as follows. An overview of basic architecture of ART2 is briefly outlined and a Match Adaptive Resonance Theory MART2 model which introduces new modifications to ART2, are presented in Section 2. In Section 3, dynamic operations of MART2 are discussed. System implementation and practical proof of MART2 are emphasised in Section 4. Finally, Section 5 summarises the comparative study and provides concluding remarks.

2 Match Adaptive Resonance Theory (MART2)

ART2 is a neural network topology with dynamics based on Adaptive Resonance Theory (ART). ART was the result of an attempt to understand how biological system is capable of retaining plasticity throughout life, without compromising the stability of previously learned patterns [4]. In this section, an overview of basic architecture of MART2 is outlined and modifications to self-organising neural network ART2 are described. The elimination of bottom-up adaptive filters, match-reset loops and vigilance parameter are variants introduced to make clustering mechanism more flexible. The objective of deriving MART2 is to make search mechanism and the determination of system parameters shorter, easier and faster.

2.1 Adaptive Resonance Theory

Adaptive Resonance Theory (ART2) neural network is introduced as a theory of human cognitive information processing [5]. It is an unsupervised neural network that based on competitive learning finds categories autonomously and learns new categories if needed. It is developed to overcome the problems of instability of feedforward systems, particularly the stability-plasticity dilemma [6].

The heart of ART2 network consists of two parts; the attentional subsystem and the orienting subsystem. Fig.1 illustrates ART2 architecture.

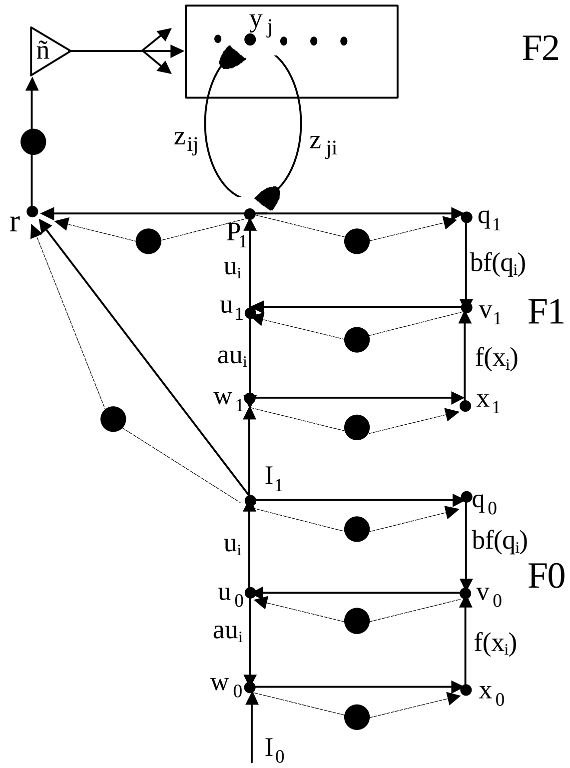


Fig.1, ART2 detailed architecture

Where I is the input vector. p, q, u, v, w and x represent STM activities of F0 and F1 nodes. y is the STM activity of F2 node. Z_{ij} and Z_{ji} denote the bottom-up and top-down LTM adaptive filter respectively and f(x) is a nonlinear function. The attentional subsystem is composed of three fields; two feature representation fields F0 and F1 that include several processing levels and a category representation field F2 where competitive learning takes place. The combination of contrast enhancement, noise suppression, normalisation, and pattern matching is produced in F0 and F1. The two fields F1 and F2 are linked by bottom-up and top-down connections called adaptive filters or Long Term Memory (LTM). The orienting subsystem measures the degree of match between the bottom-up input pattern and top-down template pattern. It also helps guide the attentional subsystem in its search for a new category.

When an input pattern I is presented to F0 field, output activations are produced in F0 and F1 nodes. F1 activities are passed on to F2 field through synaptic weights Z_{ij} to calculate F2 node activation. F2 node receives the maximum activity wins the competition. The winning F2 node produces a signal through top-down adaptive filter that is propagated back to F1 field. All other F2 nodes are inhibited through the competitive learning [7]. The signal fed back to the F1 field is compared with the input signal at upper F1 field.If the match between the two signals is close, the network is in resonance state; otherwise, a reset signal deactivates the winning F2 node and a search mechanism begins to look for a new node in F2 which matches best the input pattern.

Learning occurs in LTM whenever a match is found and resonance happens or whenever a new category node in F2 is chosen.

2.2 MART Basic Architecture

The block diagram and the detailed architecture shown in fig.2 and fig.3 describe the structure of MART2. It shares characteristics of ART2 model, having both two input representation field F0 and F1 and a category representation field F2. MART2 lacks of bottom-up adaptive filters from F1 to F2, in addition to vigilance parameter that determines match criterion, but there are connections from orienting subsystem to F2 which intervene in F2 nodes activation. Also, different winning criterion is used and new match process is obtained.

When an input pattern is presented to F0 field, output activities are produced in F0 and F1 nodes. At the beginning, all nodes in F2 are considered active. Hence, they elicit signals that are gated by top-down LTM traces and summed up at F1. Before any learning occurs, LTM-gated signals have zero values. As soon as arousal reset signals are generated from orienting subsystem to F2, activities of F2 nodes are generated. F2 node that possesses the maximum activity wins the competition and is kept in resonance state while other nodes are reset. Calculating F2 activities depending on arousal reset signal helps the network to bias to the best choice faster without falling into match-reset loops and without being never exhausted. That is, best match can only be achieved using the maximum arousal signal that comes from orienting subsystem. As well as learning occurs only to top-down connections whenever a node in F2 wins the competition. This is a big advantage of MART2 among other neural networks.

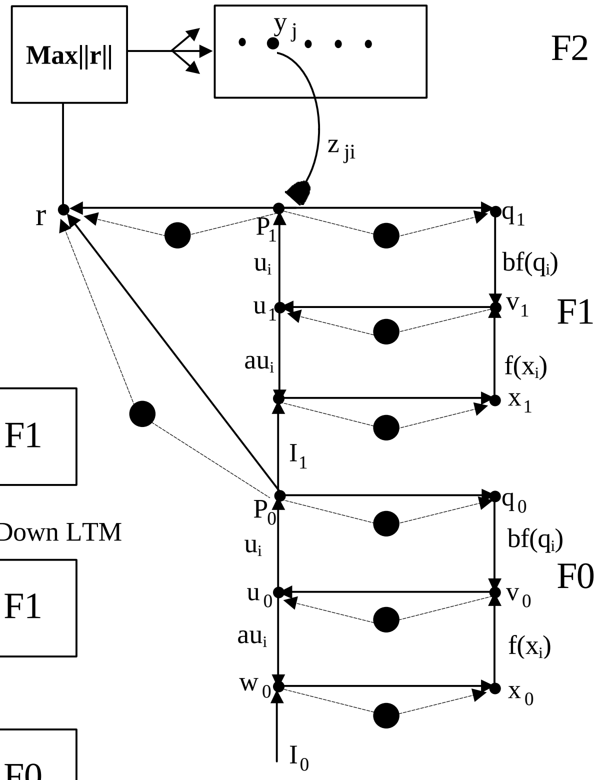


Fig.3, MART2 detailed architecture

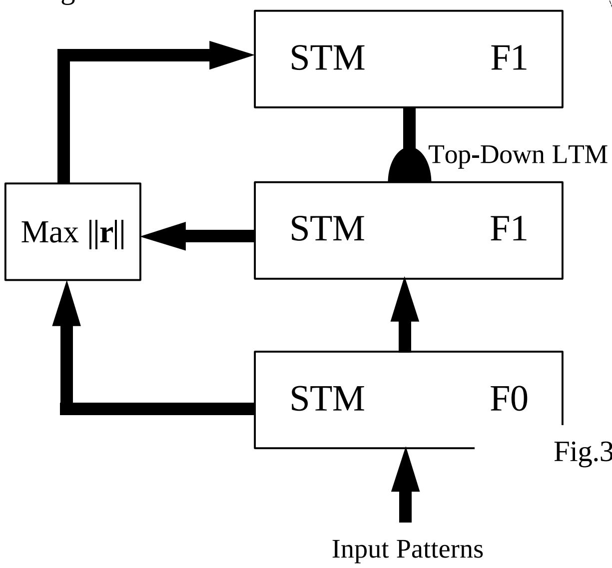


Fig. 2, Block diagram of MART2 architecture

2.3 MART2 Algorithm

1- Initialise STM activities and top-down LTM.
2- Present an input pattern from training set.
3- Calculate Short Term Memory (STM) in F0 field.
4- Calculate Short Term Memory (STM) in F1 field.

5- Send top-down LTM to F1 from all categories in F2 via the top-down adaptive filters according to the following equation:

$$p_{1i} = u_{1i} + dz_{ji} \quad (1)$$

6- Calculate active F2 nodes using new criterion of ||r|| where the activity of F2 is given by:

$$T_j = \|\mathbf{r}\| \qquad j=1,\ldots,N\text{-}M \quad (2)$$

N represents number of total F1 and F2 nodes, M is the number of the nodes in each F0 layer and r_i is given as follow:

$$r_i = \frac{p_{0i} + cp_{1i}}{e + \|\mathbf{p}_0\| + \|\mathbf{p}_1\|} \quad (3)$$

7- Choose the winning node from F2, which has a maximum value of $\|\mathbf{r}\|$.

$$T_J = \max\{T_j : j = 1,\ldots, N - M\} \quad (4)$$

8- Adjust LTM top-down traces associated with the winning node J according to the following equation:

$$\frac{dz_{ji}}{dt} = g(y_j)(p_{1i} - z_{ji}) \quad (5)$$

Where $g(y_j)$ equals a constant d if the winning node is still active; otherwise it is equal to zero.

9- Go to step "2" if there are more input patterns to train.

3 Dynamic Operations of MART2

ART2 and MART2 share some similar characteristics. Both of them perform clustering on noisy input patterns without a supervised learning procedure. They are structures based on the idea of competitive learning among F2 nodes. No bottom-up connections are established between F1 and F2 fields in MART2, but there are STM from orienting subsystem to F2. An obvious difference is that ART2 has a reset loop which controls search mechanism while in MART2, the competition process in F2 requires only one epoch to end the search mechanism. This makes MART2 save time and give better generalisation ability also, less number of data is required for training.

ART2 uses several externally determined system parameters, which directly affect the system dynamics and the learning rate of the network. ART2 parameters satisfy some parameter constraints. MART2 uses the same ART2 parameters, but they do not play any important role in the clustering process. ART2 uses a pre-specified threshold called vigilance parameter that determines how close an input pattern should match Long Term Memory traces [8]. Unlike MART2 does not require this pre-specified parameter, it is eliminated by choosing the maximum arousal signal generated from orienting subsystem to F2 which gives MART2 robust structure and consistent mechanism.

In MART2, considering all F2 nodes active at the beginning may become a fuzzy performance, because it sometimes increases the learning period significantly and other times it exhibits faster clustering of input patterns. On the other hand, in ART2 the complex process of reset loops increases the processing time for learning and clustering, and it may lead the network to be exhausted. MART2 is never exhausted and there is always a winning node in F2 to learn.

In ART2 and MART2 only the LTM traces associated with the winning node are updated. However, in MART2 there is only one set of LTM traces (Top-Down) is modified. This characteristic is particularly advantageous since less storage of memory is needed. MART2 system dynamics enable the possibility of different learning methods, "slow learning", "intermediate learning," and "fast learning". Experimental results show that MART2 requires less epochs than ART2 to reach a stable performance on the same data set. Choosing a winning node in F2, matching the bottom-up input patterns and top-down templates, reset loop and search loop represent the order of operations in ART2. On the other hand, matching the top-down templates and input patterns at F1 and choosing a winning node in F2 are the order of operations in MART2.

From psychological point of view, ART2 considers that using processes of resonance and reset, brain can discover and learn stably only new representations for novel events in an efficient way, without assuming that there are previous internal representations already exist [9]. While MART2 contemplate that brain has some internal representations which may be selected, amplified or used during learning familiar or novel events. This helps MART2 to bias to a stable and consistent clustering of arbitrary input patterns faster without depending upon bottom-up adaptive filters.

4 Speech System Implementation

A classifier system is developed to recognise Arabic letter signals. As mentioned above, It contains three basic components. Fig.4 illustrates a block diagram of the signal flowchart from the input letter to MART2 output.

4.1 Time Domain Processing

4.1.1 Recording

Twenty-eight Arabic letters are recorded three times as monosyllables for an independent-speaker, using a microphone and voice card which converts analogue input signal into digital output signal.
They are recorded with low quality of 8 bits per sampled audio amplitude and at a sampling rate of 22050 Hz.

4.1.2 Noise Suppression

Many factors affect the recording and make it noisy, such as the speaker conditions, environment, microphone variation, voice card, etc; therefore, it is essential to detect pure data from noisy speech signal. Detection of voiced/unvoiced frames is made in time domain, using three parameters of shift ratio, the amplitude average and the zero crossings rate. Fig.5 displays the waveforms of letter ر (Raa: in Arabic) before and after the noise suppression is applied.

4.1.3 Segmentation

A letter spectrum is divided into a number of frames, each frame size is set to 1024 samples.

4.1.4 Noise Suppression and Segmentation Procedures

The noise suppression procedure uses the following steps:

- Recording noisy environment without talking to the microphone.
- Calculate the shift ratio, amplitude average and zero crossings rate of the noise signal.
- Recording the twenty-eight Arabic letters in the same noisy environment.
- Applying segmentation.
- Calculate the shift ratio, amplitude average and zero crossings rate for each frame.

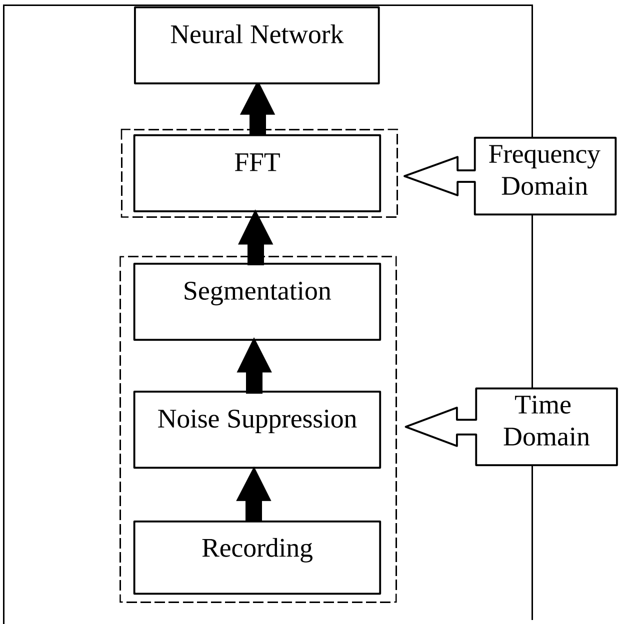
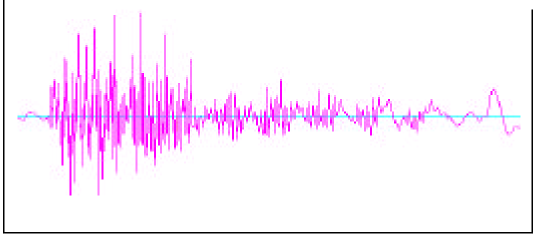
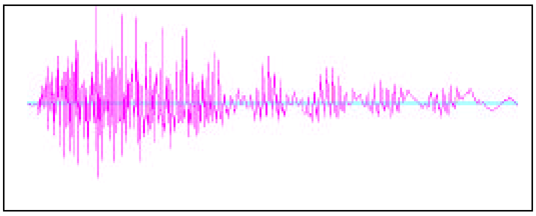


Fig.4, Block diagram of a speech classifier system



a. Original signal of letter Raa before noise suppresion



b. The processed signal of letter Raa after noise suppression.

Fig.5, Arabic Letter Processing

- Compare the three parameters for each frame with the three parameters obtained from the noise signal.
- Suppress all those frames, which have average amplitude and zero crossings less or equal to the average amplitude and zero crossing of noise signal.
- Repeat the calculation for the rest of the frames.

4.1.5 Shift Ratio

Represents the amount of offset of the signal from basic zero level that should be subtracted from all samples to extract the original speech signal for all Arabic letters. Sample values are between 0 and

255, values between 0 and 127 are considered as negative samples; otherwise, they are positive samples, basic zero level is equal to 128.

Shift Ratio = Zero Level Shifted Ratio – Basic Zero Level

$$SR = \frac{1}{N}\sum_{i=1}^{N} X_i - 128 \qquad (6)$$

Where SR, X and N denote the shift ratio, sample value and samples numbers, respectively.

4.1.6 Amplitude Average

Considered a vital parameter in suppression noise procedure, which represents an average absolute value of sample amplitude with respect to zero level. Low values of amplitude average refer to unvoiced frames.

$$Amp\_Avrg = \frac{1}{N}\sum_{i=1}^{N} | X_i | \qquad (7)$$

Amp_Avrg represents the amplitude average.

4.1.7 Zero Crossings rate

Reflexes the number of intersections of a speech signal with respect to zero level shifted ratio. High zero crossings rate indicate signal contains high frequencies and a voiced frame.

4.2 Frequency Domain

In order to reduce samples numbers to 1024 samples for each Arabic letter, FFT is used to transform the speech signal from time domain to frequency domain. This procedure is done utilising butterfly diagram, detail of which can be found in [10].

4.3 Neural Networks

512 samples which obtained for each letter from the FFT stage are presented as input patterns to ART2 and MART2 described above, that have configured to consist of 512 neurons in F0 and F1 for each processing level and 28 neurons in F2 representing Arabic letters. Three groups of 28 Arabic letters are recorded. The first group is used for training the networks, while the others two are used in test performance.

5 Results and Discussion

It has been demonstrated in previous sections, there exist differences on the architectures and learning algorithms between ART2 and MART2. Two comparative experiments are applied where the most favourable parameters have been elected. Fig.6 depicts the category learning of twenty eight Arabic letter signals by ART2 and MART2. The input patterns are presented in an order form during learning procedure for both ART2 and MART2.

First six columns show the first experiment and the last six columns show the second one. Columns 2 and 5 show ART2 clustering results with parameters a=b=8, c=d=0.5, e=0.001 and ñ=0.998, the network can learn with one epoch during 880 sec. Columns 3 and 6 show the MART2 results with the same parameters and without vigilance parameter. The network can learn with one epoch during 889 sec. In both networks, 28 Arabic letters signals are distributed through 28 eight nodes in F2 field. In this case, the convergence of ART2 to a stable clustering depends on the good choice of network parameters, while MART2 can bias the network to get stable and consistent clustering without relying on them.

Columns 8 and 11 show ART2 clustering results with the following parameters a=b=2, c=0.7, d=0.3, e=0.001and ñ=0.999, the network can not learn 28 letters even with many epochs are used and only 25 letters are categorised. Columns 9 and 12 show the MART2 results with the same parameters and without vigilance parameter. The network can learn with one epoch during 908 sec and 28 Arabic letter signals are distributed through 28 eight nodes in F2 field. In this case, MART2 has a better accomplishment than ART2 because it needs fewer epochs to reach a stable performance.

Table 1 illustrates and summarises the clustering results for more experiments. The input patterns are also presented in an order form during the learning procedure for both ART2 and MART2.

On experiments 9 and 10, it is explored the successful achievements of the new search mechanism of MART2 without using bottom-up connections from F1 to F2, additionally the attainment of the absence of vigilance parameter. MART2 has better performance than ART2.

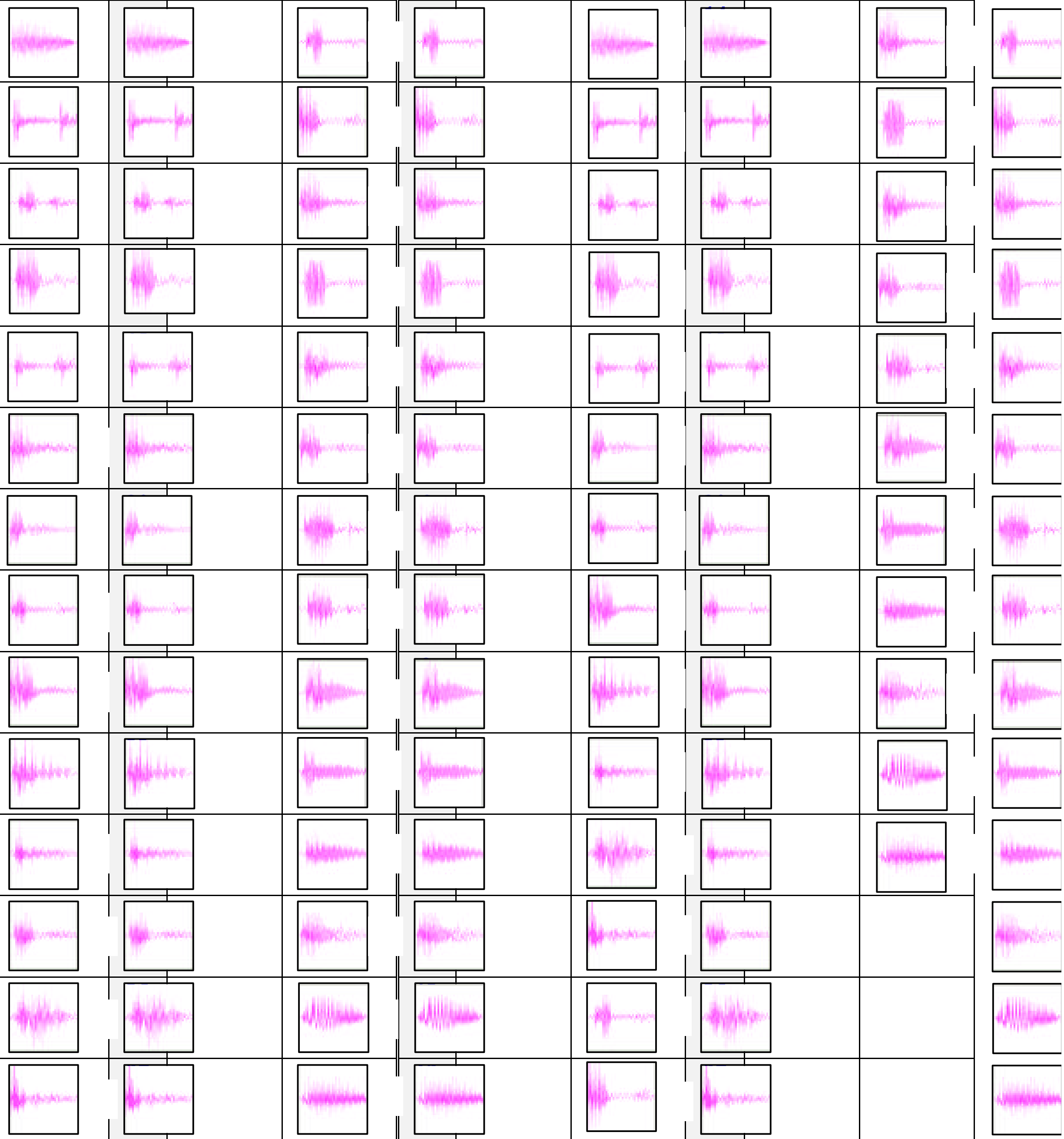
On experiments 1,2,15,16 and 18, ART2 learns and groups the 28 letter patterns into 28 stable categories within one presentation and it can recall 28 letters. MART2 also can learn and recall the same inputs patterns with the same parameters. In this case, ART2 searches for a resonant node and train it, while MART2 initially all nodes are in resonant state and the search mechanism is applied to the competitive reset wave which is generated from orienting subsystem toward F2.

When ART2 is exhausted and can not learn, MART2 can learn, cluster and recall all patterns. This is an advantage which characterises MART2, although it utilises the same parameters and without vigilance parameter. This is illustrated on experiments 3 and 5.

6 Conclusions

MART2 neural network is designed to maximise generalisation and minimise classification error in response to Arabic letter signals. Eminently, the network can achieve better accuracy than ART2 without using bottom-up LTM traces, vigilance parameter and match-reset loops. Although the diversity variants have introduced to basic

1	2	3	4	5	6	7	8	9	10	11	12
	ART2	**MART2**		**ART2**	**MART2**		**ART2**	**MART2**		**ART2**	**MART2**
#Node	**a=8,b=8, c=0.5, d=0.5, e=0.001, ñ=0.998, epoch=1**	**a=8,b=8, c=0.5, d=0.5, e=0.001, ---, epoch=1**	**#Node**	**a=8,b=8, c=0.5, d=0.5, e=0.001, ñ=0.998, epoch=1**	**a=8,b=8, c=0.5, d=0.5, e=0.001, ---, epoch=1**	**#Node**	**a=2,b=2, c=0.7, d=0.3, e=0.001, ñ=0.999, epoch=10**	**a=2,b=2, c=0.7, d=0.3, e=0.001, ---, epoch=1**	**#Node**	**a=2,b=2, c=0.7, d=0.3, e=0.001, ñ=0.999, epoch=10**	**a=2,b=2, c=0.7, d=0.3, e=0.001, ---, epoch=1**
0											
1											
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											



ART2	**28 nodes trained** **28 nodes recognised** **0 nodes failed**	**ART2**	**28 nodes trained** **25 nodes recognised** **3 nodes failed**

MART2	**28 nodes trained** **28 nodes recognised** **0 nodes failed**	**MART2**	**28 nodes trained** **28 nodes recognised** **0 nodes failed**

Fig.6, category learning of Arabic letter signals by ART2 and MART2.

Experiments on Arabic Letters						**ART2**					**MART2**		
						Learning			**Recall**		**Learning**		**Recall**
No	**A**	**b**	**c**	**d**	**e**	**ñ**	**epoch**	**No of Nodes**	**Success**	**Failure**	**Epoch**	**No of Nodes**	**Success**
1	8	8	0.5	0.5	0.001	0.998	1	28	28 Stable	0	1	28	28 Stable
2	10	10	0.5	0.5	0.001	0.998	1	28	28 Stable	0	1	28	28 Stable
3	2	2	0.3	0.7	0.1	0.99	1	Exhausted	-	-	1	28	28 Stable
4	2	2	0.7	0.3	0.01	0.99	1	28	26	2	1	28	28 Stable
5	2	2	0.1	0.9	0.001	0.998	1	Exhausted	-	-	1	28	28 Stable
6	2	2	0.7	0.3	0.001	0.999	10	25	-	-	1	28	28 Stable
7	9	9	0.6	0.4	0.001	0.999	5	28	27	1	1	28	28 Stable
8	9	9	0.9	0.1	0.0001	0.999	10	26	-	-	1	28	28 Stable
9	9	9	0.2	0.8	0.001	0.999	1	26	-	-	1	28	28 Stable
10	8	8	0.2	0.8	0.001	0.95	5	27	-	-	1	28	28 Stable
11	8	8	0.8	0.2	0.001	0.999	1	28	27	1	1	28	28 Stable
12	8	8	0.7	0.3	0.001	0.999	1	28	27	1	1	28	28 Stable
13	8	8	0.7	0.3	0.001	0.998	1	28	27	1	1	28	28 Stable
14	10	5	0.4	0.6	0.001	0.97	10	27	-	-	1	28	28 Stable
15	10	5	0.5	0.5	0.001	0.998	1	28	28 Stable	0	1	28	28 Stable
16	7	3	0.4	0.6	0.001	0.999	1	28	28 Stable	0	1	28	28 Stable
17	9	10	0.7	0.3	0.0001	0.999	1	28	27	1	1	28	28 Stable
18	9	10	0.7	0.3	0.0001	0.998	1	28	28 Stable	0	1	28	28 Stable

Table 1 ART2 and MART2 Experimental Results

architecture of ART2 and the simpler dynamics of MART2, it still conserves the ability of learning stable, in the sense of getting new knowledge, without forgetting prior information and efficiently, in the sense of utilising less training data set and few epochs to reach a stable performance on the same data set. Additionally, in ART2, F0 and F1 parameters play an important role in classifying input patterns, while in MART2 is not affected by them and still clusters stable and consistent categories in F2. This is an advantage of MART2 over ART2 because determining appropriate parameters values is considered a time-consuming. Experimental results show the effectiveness of MART2 network. It requires less number of data, simpler architecture, easier learning rules and gives better accuracy than ART2.

References:

[1] Carpenter G.A., *Neural Network Models for Pattern Recognition and Associative Memory*, Neural Networks, Vol.2, 1989, pp.243-257.

[2] Carpenter G.A. and Grossberg S., *ART2: Self-Organization of stable category y recognition codes for Analog Input Patterns*, Applied Optics, Vol. 26, 1987, pp.4919-4930.

[3] Carpenter G.A. and Grossberg S., *A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine*, Computer Vision, Graphics, and Image Processing, Vol.37, 1987, pp.54-115.

[4] Carpenter G.A. and Grossberg S., *ART2-A An Adaptive Resonance Algorithm for Rapid Category Learning and Recognition*, 1987, pp.493-504.

[5] Carpenter G.A. et al, *Fuzzy ART: Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance System*, Technical Report, Boston University, CAS/CNS-TR-91-015, 1987.

[6] Geoffrey E.Hinton, Terrence J. Sejnowski and Peperback, *Unsupervised learning: Foundation of Neural Computation* (Computational neuroscience series), MIT press; ISBN: 026258168X, 1999.

[7] Grossberg S., *Competitive learning: from Interactive Activation to Adaptive Resonance*, Cognitive Sci. 1987, pp. 11, 23.

[8] Grossberg S., *The link between brain learning, attention, and consciousness*, Consciousness and Cognition, Vol 8, 1999, pp. 1-44.

[9] Grossberg S., *The complementary Brain: A Unifying View of Brain Specialization and Modularity*, 2000, Trends in Cognitive Sciences, in press.

[10] Zonst Anders E., *Understanding the FFT*, Second Rev edition, citrus Press; ISBN:0964568152, 2000.