2012 International Conference on Image, Vision and Computing (ICIVC 2012)
IPCSIT vol. 50 (2012) © (2012) IACSIT Press, Singapore
DOI: 10.7763/IPCSIT.2012.V50.50

Fast Signal Clustering Based on a Compound Adaptive Resonance Model

Liao Huirong[1+], Li Guolin[1], Mao Weiping[1] and Zhou Wensong[2]

[1]Graduate Students' Brigade, Naval Aeronautic and Astronautic University, NASU, Yantai, China
[2]Training Department, Naval Aeronautic and Astronautic University, NASU, Yantai, China

Abstract. Design a new algorithm for fast signal clustering based on ANN via compound adaptive resonance model. The compound model use modular architecture， and its input layer is composed of several subordinated artificial neural networks, which extract each feature of data sets to form knowledge remember after supervised learning, and convert the original input data into inner normal input on learned knowledge base, to reduce the complexity and calculated amount of adaptive resonance when running. The pattern matching and unsupervised learning in adaptive resonance procedure are all based on comprehensible logic rulers. The system has obvious extensibility and evolutionary. Comparative simulation experiments are used to investigate the performance of new algorithm and ART2 on signal clustering problem. The result indicates that the new algorithm generally exhibits faster learning and high tolerance to noise interference. And it decreases the probability of false conviction when low vigilance values are used in case of low SNR.

Keywords: artificial neural network (ANN); compound adaptive resonance model (CARM); signal clustering; ELINT.

1. Introduction

The fast signal clustering is one of the most important and difficult problems in signal reconnoitre. Classical statistic methods are not adaptive to complex electromagnetic environment because of requiring too much priori information. Recently several clustering algorithms on artificial neural network spring up with the development of computational neural science. These self-organizing clustering neural networks which be most often used are Hopfield ANN, Learning Vector Quantization (LVQ), Self-organizing Feature Map (SOFM) and Adaptive Resonance Theory(ART)[1]. Hopfield ANN does not have good tolerance to noise interference because of the requirement that the pattern vectors to be recognized must be perpendicular. LVQ and SOFM are not applicable in situations where the number of clusters is generated by the system in an autonomous way. And the performance of clustering is connected with the primary condition of their neural network.

The clustering method on self-organizing neural network, ART2, is considered having obtained marked effect. ART2 was proposed by Carpenter and Grossberg to satisfy design principles derived from the analysis of neural net works [2]. The concepts of competitive learning and interactive activation are applied in a manner that leads to self-organization of stable category recognition codes for analog and binary input patterns. However, the ART2 network convergence speed is very slow because the procedure of neural kinetics is too complex. And the ART2 don't tolerate the deformation of representative members in each class by a

Corresponding author. Tel.: +13475356127.
E-mail address: Liaohuirong8205@163.com.

prescribed amount of signal-to-noise ratio on the determination of threshold values. Actually the vigilance parameter in ART2, which determines how closely an input pattern should match the stored patterns, does not explicitly correlate to the SNR or any other measure that relates to the input data [3]. No direct or effective method to determine the value of vigilance based on a priori information of the data has yet been available.

This paper systematizes the ART2 and improves the input layer to compose a transforming cluster of subordinated artificial neural networks, which adopt supervised learning independently to obtain knowledge base. The network achieves normalization of input space and establishes comprehensible logical structure of rules for pattern matching, and then improves the clustering performance.

2. A Compound Adaptive Resonance Model

The study of neurobiology indicate that human cerebral cortex have a regional structure for functions. There are many different neural functional regions in the cerebral cortex, and each complete the respective function [4]. The theory of working and learning in the regions is discord. They contact each other using normalization interface and inform compound neural functional system.

2.1. System Structure and Model

The structure of the compound neural network designed in this paper is consisting of four subsystems as show in the figure 1: input mapping subsystem, attention subsystem, orienting subsystem and learning subsystem. The compound adaptive resonance model architecture inherits the basic components of the adaptive resonance theory, and do some adjustment and improvement on the basis. For example, the original simple input layer is changed into input mapping subsystem which obtains the conversion function like interface [5]. The attention subsystem contains an input representation field F1 and a category representation field F2, and the orienting subsystem interacts with the attention subsystem to perform an internally controlled search process. The input representation field F1 and the category representation field F2 are linked by bottom-up F1 to F2 adaptive filter and top-down F2 to F1 adaptive filter.

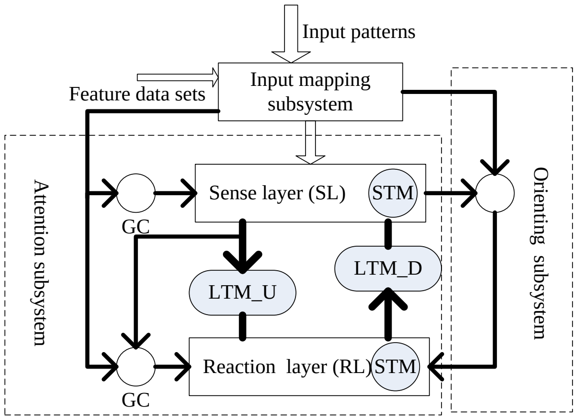
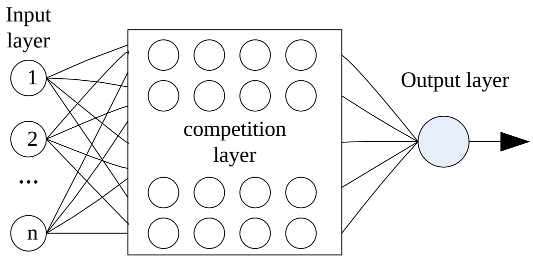


Fig. 1 The systematization structure of CARM

- Input mapping subsystem. The main work of the input layer is processing feature extraction from input signal and complete the form converting intend to transforming the higher dimensional signal to fixed low dimensional inner signal. The neural signals transmitted from sense system to cerebral cortex in the neurobiology system are provided with fixed inner rulers and dimensions. The sense system of neurobiology via which external input signals are mapped into an sequential feature chart is provided with self-organization topology feature mapping function and acquired nonlinear clustering property parameters of the data sets. Input mapping subsystem adopts supervised learning rule to extract and remember the feature by modifying the contacting weight of the self-organization mapping network. When running it wants to determine if the input pattern has the feature and output the decision value.
- Attention subsystem, which is consist of sense layer and response layer and long time memory and short time memory, processes mostly the learned patterns. The sense layer recept the activation and is provided with short time memory function. The response layer responds the activation and is provided with short time memory function too. The long time memory, as shown as the figure 1, is consisting of

up-long-time-memory (LTM_U) and down-long-time-memory (LTM_D). The gain controller (GC) control the down gate of learned expecting and make the attention subsystem distinguish the up consciousness activation and the down consciousness activation.

- Orienting Subsystem, which processes the non-learned patterns (the new and learning patterns), check and measure the disaccord of the up consciousness activation and the down consciousness, and restrain the learned pattern which doesn't match the up consciousness activation in LTM_D.
- Learning Subsystem, which achieve the ART learning, modify and adjust the long time memory, include the LTM_U and LTM_D.



\

Fig. 2 Input mapping patterns

The self-organizing capability of compound adaptive resonance network is based on the idea of competitive generation and elimination of attraction wells. Each attraction well is characterized by its center, width (threshold), and depth. The width of a well is determined from the lower bound of a tolerable signal-to-noise ratio (SNR). The depth of a well is a measure of confidence on a learned pattern (or stored cluster), and influences the inertia of the center of a well. The greater the depth of a well is, the lesser the well center moves towards a new position during the update process [6].

The similarity between patterns in the compound adaptive resonance is measured in terms of the inner-product of a normalized input pattern vector and learned exemplar vectors. If the magnitude of each element of the input pattern vectors is considered an important feature by itself, classes of patterns which are separable by more complicated boundary shapes can be discriminated through self-organization by using different metrics instead of the inner product.

2.2. The Formation of Knowledge base in the Input Mapping Subsystem

Given an universe of discourse U, for example the signal space that we want to cluster, and a cluster of equivalence relations (S) on U, so the two-tuples K=(U,S) is called knowledge base with U. The knowledge is the ability of clustering the objects in this domain. So the equivalence relations on U represent the clustering and knowledge. In that way the knowledge base is consist of the various knowledge which is derived from the equivalence relations(attribute property and the finite intersection), that is clustering or patterns and represent the ability of clustering the universe of discourse in the same time.

The work that the input mapping subsystem want to achieve is building up the equivalence relations on U by non-linear mapping and inform the knowledge base from that. The ability of non-linear mapping in ANN is obtained via learning.

The learning procedure, namely organization procedure, includes primarily tree important nodes:

1) Competition Node; When the sense domain V1 of input mapping subsystem is stimulated by the signal $x=(x_1,x_2,...,x_n)^T \in D_{SOM}^{(n)}$, the neurons of response layer export the signal $o_j(x)$. The neurons of input mapping subsystem compete with each other and represent mostly the strongest response when stimulated by the input signal X.

The rule of competition is that:

$$\text{If } o_{i(x)}(x) = \max_{1\le j\le N}\left\{o_j(x)\right\} \tag{1}$$

Then $v_{i(x)}$ became the winning unit for x. The output o_j of v_j is the inner product of the input vector x and synapse weight vector w_j： $o_j = w_j^T x$ （j=1, 2…, N）.

According the best matching criterion, the maximization of the inner product is equivalent to the minimization of the Euclidean distance when the two vectors are normalizing vectors.

That is when $\|x\| = \|w_j\| = 1(j = 1, 2..., N)$, we can get：

$$\|x - w_j\|^2 = (x - w_j)^T (x - w_j) = 2 - 2w_j^T x \tag{2}$$

Obviously, according to the (2), determining that the maximization of $o_j = w_j^T x$ is equivalent to determining that the minimization of $\|x - w_j^T\|$.

So, the winning unit is also estimated via the equation ， and adopting this equation can avoid the cumbersome normalizing processing for input vector x and feed-forward synapses vector $w_j (j = 1, 2, ...N)$.

$$i(x) = \arg\max_{1 \le j \le N} \|x - w_j^T\| \tag{3}$$

2）Cooperation Node: After the neuron of input mapping subsystem became the winning neuron unit that will establish a topology neighbourhood to center around itself. The situation was like a biological neural system that neurons stimulated in some similar to the Mexican hat function distribution in the vicinity of synaptic excitation efficiency nerve cells, together with its and their own excitement.

The focal functions is defined to calculate the $\omega_{i(x)j}$. Assuming $d_{i(x)j}$ is the lateral distance from v_j to $v_{i(x)}$, then $\omega_{i(x)j}$ is designed to unimodal function of the lateral distance $d_{i(x)j}$, which meet: symmetry, monotonicity, convergence condition: $\omega_{i(x)j} \to 0(d_{i(x)j} \to \infty)$.

3）Synaptic Adaptation： the adaptation procedure of the input mapping subsystem is the synaptic adaptation and organization processing that modifying or adjusting the connection efficiency $w_{ij} (i = 1, 2, ..., n; j = 1, 2, ..., N)$ of neuron synapses.

If $v_{i(x)}$ is the winning unit for the given stimulus signals $x = (x_1, x_2, ..., x_n)^T \in D_{SOM}^{(n)}$ ， the adjustment Δw_j of the feed-forward connection weight w_j for any neuron v_j of the input mapping subsystem is consist of two parts：Hebb learning paragraph, $\alpha o_j x$； Forgetfulness paragraph: $-q(o_j)x$.

In that, $\alpha > 0$ is the learning rate ,and just when $o_j = 0$， $q(o_j) \ge 0, q(o_j) = 0$.

If o_j is simply assumed with $\omega_{i(x)j}$, and set that $q(o_j) = \alpha o_j$ ， then the rule of the input mapping subsystem is that:

$$\begin{cases} \Delta w_j(t) = \alpha(t)\omega_{i(x)j}(t)\left(w(t) - w_j(t)\right) \\ w_j(t+1) = w_j(t) + \Delta w_j(t) \qquad (j = 1, 2, ..., N) \end{cases} \tag{4}$$

Among them, $t = 0,1,...$ is discrete-time, and learning rate $\alpha(t)$ is the function of the time t， which is exponential decay with time t like the effective radius σ of the synergistic neighborhood in Gaussian neighborhood function: $\alpha(t) = \alpha_0 \exp(-\kappa_\alpha t)$, where the learning rate for the initial time is $\alpha_0 = \alpha(0)$.

2.3. Design and Normalization of the Input Space

When the input signal access to the artificial neural network, it is firstly converted to a variety of feature subspace, the feature subspace data is then converted to pattern neuron signal on inner rules by the input mapping subsystem.

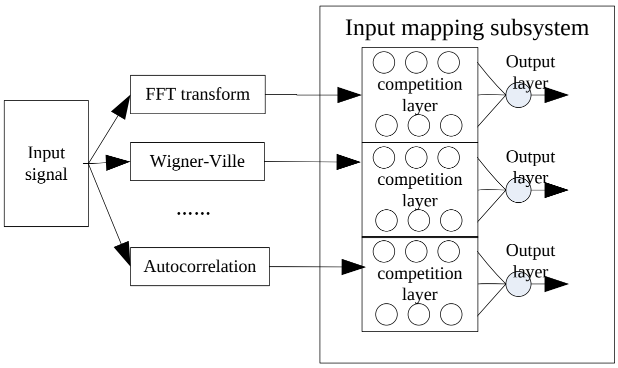


Fig. 3 The transforming of input space

Input signal is transformed to corresponding feature space domain by the FFT transform, autocorrelation operator, Wigner-Ville transforming. The input mapping subsystem extracting the feature and normalized into standard inner input signal. The input layer of the input mapping subsystem, as the retina in the biological visual system, receives the signal x from the input node, and its basic function is the normalization with the input signal.

The input normalization of the sense layer is that for any $x \in [0,+\infty)^{n\times 1}$, the pattern signal $o^{(s)} \in [0,1]^{n\times 1}$ of the network sense layer represent the normalizing ability to the signal with wide various range.

3. The Running Procedure of the Adaptive Resonance Network

The block diagram in Figure 4 shows the running procedure of patterns matching when the adaptive resonance network is running. The preprocessing stage performs normalization on input patterns.

3.1. The Running Procedure of the Network

Let x_n be an input pattern presented to compound adaptive resonance at the n-th iteration. To update the state of the network, execute the following steps.

1) *Determine the threshold value:*

$$threshold = \frac{1}{\sqrt{1+10^{-\frac{SNR}{10}}}} \tag{5}$$

Where SNR (in dB) is the lower bound of a tolerable signal-to-noise ratio, above which MARM is designed to operate properly.

2) *Calculate the attraction well width(in radians):*

$$\theta = \arccos(threshold) \tag{6}$$

3) *Pre-process and normalize the n-th input pattern* x_n .

4) *Compute the similarity:* Compute the similarity between the normalized input pattern x_n and each stored exemplar.

5) *Determine the update status* c_n *of each well from the result, which selects the exemplar to which the input pattern has maximum similarity. The value* c_n *of an attraction well is determined by the following rule:*

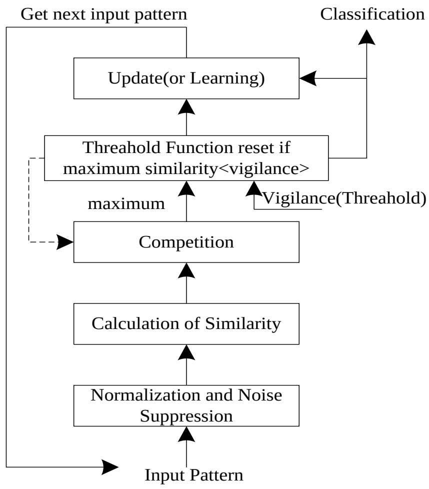


Fig. 4 A sample line graph using colors which contrast well both on screen and on a black-and-white hardcopy

6) *Update the depth of a well according to* $d_n = d_{n-1} + c_n$, *with an initial condition* $d_0 = 0$.

The stored exemplar is updated by using the following rules:

$$e_n = \frac{c_n}{d_n} x_n + \frac{d_{n-1}}{d_n} e_{n-1} \tag{7}$$

With an initial condition $e_0 = 0$, where d_{n-1} is the depth of an attraction well after the presentation of the previous input pattern x_{n-1}.

7) *Eliminate attraction wells which do not have any interaction in a desired time period.*
8) *Repeat Steps 3-7 for next pattern.*

3.2. Symbol Rules of the System Structure

The common defect of the artificial neural network model is that the knowledge hided in a large number of connection weights and it is difficult to be understood or given a clear easy reasoning judge process. However, we can create a more uniform guiding framework to guide the construction of artificial neural network model and the weight modification.

Neural network training requires a lot of data. If there is a large enough sample set and determined by trained network, the outcome will include both right and wrong results. But whether correct or not, these results fully reflect the performance of the network. If we can extract the rules from the original input data and the sample set which is consist of determining results of the neural network, the effect of using these rules will be similar to the original neural network model and these rules can describe the function of the original neural network model and give out the understandable symbol discrimination on the determine results of the neural network[7].

Let U be a non-empty finite set consisted of the objects we are interested in, called a universe. Any one subset X of U in the universe, $X \subseteq U$, can be considered as a concept or category of the universe U. The two-tuple K = (U, S) is a knowledge base on the domain U, $R \in IND(K)$ represent a group or individual system parameters that describes the system features. Set $\pi(U) = \{X_1, X_2, ..., X_n\}$ as a division of the domain U independently. The importance degree of the set X on the parameter R is defined as

$$sig_R(X) = \frac{|U - bn_R(X)|}{|U|} \tag{8}$$

The importance degree of the division $\pi(U)$ with system parameter R as

$$sig_R(\pi(U)) = \frac{\sum_{i=1}^{n} |U - bn_R(X_i)|}{n|U|} \tag{9}$$

The set $bn_R(X) = \bar{R}(X) - \underline{R}(X)$ is called R-boundary region of X, $\bar{R}(X)$ and $\underline{R}(X)$ are respectively the upper approximation and lower approximation.

The greater the importance degree of the system parameter R is, the smaller the R-boundary region of the set X is. This indicates that the classification ability of the parameter is stronger than other, so the input connections weigh of this parameter get stronger accordingly [8].

4. Simulation of Clustering Performance

To investigate the clustering performance of compound adaptive resonance and ART2, specific SNR and threshold (vigilance) values are chosen for each individual simulation. The SNR is varied from -10dB to +15dB. In compound adaptive resonance model simulations, the threshold ranges from 0.38 to 0.98. In ART2 simulations, the vigilance ranges from 0.837 to 0.99. The percentage of successful clustering during the simulations is illustrated.

4.1. The Structure of the Simulation and Producing Input Data

In the experiment, eight simulation signals $1.0 + \cos(2\pi f_i t)$ with different integer frequencies $f_i = 1,2,...,8Hz$, are used as noise-free exemplars. Each signal is sampled at the Nyquist rate of the highest frequency. The sampled noise-free exemplars used in the experiments become $1.0 + \cos(2\pi f_i n/16)$, where n=1,2,…,16. The data set of input patterns to the networks are generated by adding Gaussian noise to those noise-free exemplars. For each class, 125 noisy input patterns are generated. There are totally 1000 testing input patterns in the data set. In every learning epoch, each of the 1000 noisy input patterns is presented once to the networks. Input patterns are presented in random order.

4.2. Simulation Results Analysis

The compound adaptive resonance model simulation results are compared with the ART2 simulation results on the same data set. For the evaluation of clustering performance of the networks, the index of each created cluster is labelled with the class index of the majority of patterns classified in the cluster.

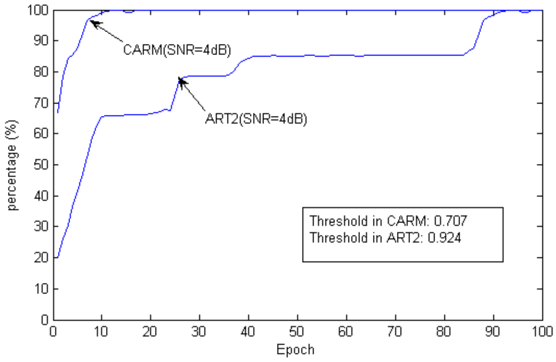


Fig. 5 The percentage of successful clustering versus epoch

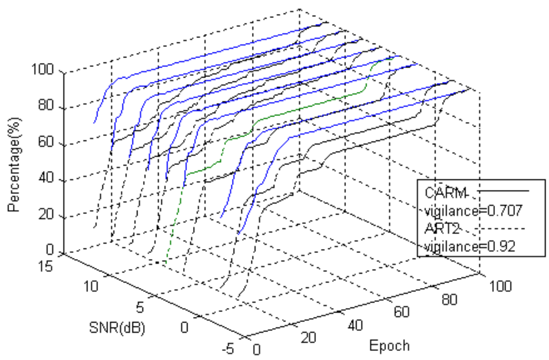


Fig. 6 Tracing of percentage of successful clustering in various SNRs

Figure 6 shows the percentage of correct classification for consecutive training epochs at different SNRs for compound adaptive resonance model and ART2. Percentage of successful clustering of compound adaptive resonance model is represented by solid lines, and that of ART2 is by dashed lines.

From the simulation results, it is seen that compound adaptive resonance model exhibits considerably faster learning of the set of input patterns, and it needs fewer epochs than the ART2 to reach a stable clustering result. An observation from Figure 6 is that the learning rate of ART2 reduces as the SNR increases.

5. Acknowledgment

This research was supported by military project and laboratory of institute. Thanks to the experts for discussion and help.

6. References

[1] G. A. Carpenter and S. Grossberg, “ART2: Self-organization of stable category recognition codes for analog input patterns,” Applied Optics, vol.26, December 1987, pp. 4919-4930 .

[2] G. A. Carpenter and S. Grossberg, “The ART of adaptive pattern recognition by a self-organizing neural network,” Computer Magazine, March 1988, pp. 77-88.

[3] C.-D. Wann and S. C. A. Thomopoulos, “Clustering with unsupervised learning neural networks: a comparative study,” in Proceedings of SPIE-Applications of Artificial Neural Networks IV , vol.1965, (Orlando, FL), April 1993 , pp.346-357.

[4] Handl J., Knowles, J., and Dorigo, M. (2006). Ant-based clustering and topographic mapping. *Artificial Life*, 12(1):35–61.

[5] Parsons, l., Haque, E., and Liu, H.(2004). Subspace clustering for high dimensional data: a review. SIGKDD Explorations, Newsletter of the ACM Special Interest Group on Knowledge Discovery and Data Mining, 6(1):90-105.

[6] Fabien Picarougne, Hanene Azzag, Gilles Venturini, Christiane Guinot. A New Approach of Data Clustering Using a Flock of Agents. Evolutionary Computation Fall 2007, Vol. 15, No. 3, pp. 345-367.

[7] Gleb Basalyga, Emilio Salinas. When Response Variability Increases Neural Network Robustness to Synaptic Noise. Neural Computation June 2006, Vol. 18, No. 6, pp. 1349-1379: 1349-1379.

[8] Arun Rao, Mark R. Walker, Lawrence T. Clark, L. A. Akers, R. O. Grondin. VLSI Implementation of Neural Classifiers. Neural Computation Spring 1990, Vol. 2, No. 1, pp. 35-43: 35-43.