Chaotic Resonance Theory, a New Approach for Pattern Storage and Retrieval in Neural Networks[1]

Radu Dogaru, A.T. Murgan

"Politehnica" University of Bucharest, Applied Electronics Department,
Spl. Independentei Nr. 313, Sector 6, Bucharest, ROMANIA,
Tel: +40-1-4105400/ ext. 140
E-Mail: radu_d@lmn.pub.ro ; atmurgan@vala.elia.pub.ro

ABSTRACT

A new architecture and methods for information storage in neural networks are presented. Behaving as Adaptive Resonance Theory neural networks, the proposed architecture is based on a different operation principle called Chaos Resonance Theory . According to this theory, standard neural units were replaced with small recurrent neural networks with chaotic dynamics, placed in a two layer architecture. The storage and retrieval of patterns are essentially based on chaos synchronization, and there are very few connections between units and layers, the architecture being attractive for VLSI implementation. Numerical simulations proved the possibility to store high amounts of patterns using relatively few units and synapses, the proposed architecture being also a plausible model for the neuro biological function of memory. The chaotic coding of information could also explain telepathic phenomena.

1. Introduction

Pattern storage and retrieval is one of the most important feature of the artificial neural networks. While it is not yet clear how this process is achieved in brains, there are a lot of studies and models for artificial neural networks memories. According with some opinions [7] chaos can play an important role in the process of storage and retrieval of information. The Hopfield model considers the prototype patterns (to be stored) as attractors in the state space of a neural system with symmetric weights. Generally speaking, they can be viewed as centroids which best characterize all the patterns belonging to a cluster centered on the prototype pattern associated with the centroid. Simple learning rule based on Hebb's law assure patterns storage while the retrieval is based on fast convergence to the closest fixed point solution, considering the input pattern presented as the initial condition. The clusters associated with a pattern are thus basins of attraction. There are however some disadvantages related with spurious attractors and low memory capacity. There is no possibility to control the radius of each cluster, and thus some overlapping phenomena can occur. The ART theory [1] which is much inspired from neuro-biology and neuro-psychology gives solutions to overcome the major drawbacks of the Hopfield associative memory model, being adaptive to new stimuli due to an unsupervised built-in learning mechanism. The "vigilance" parameter allows to globally control the clusters radius. The clustering process occur in a competitive layer of standard neurons and the learned information is stored in the synaptic weights between the two major layers of such a network. In this paper we introduce the Chaotic Resonance Theory (CRT) as a new principle for pattern storage and retrieval, essentially based on exploiting the phenomena of chaos synchronization [5]. A completely new neural architecture based on CRT is presented. This architecture can be used for clustering, having a behavior almost similar with the ART neural networks, while the different principle used to achieve this behavior gives some opportunities regarding the VLSI implementation and .

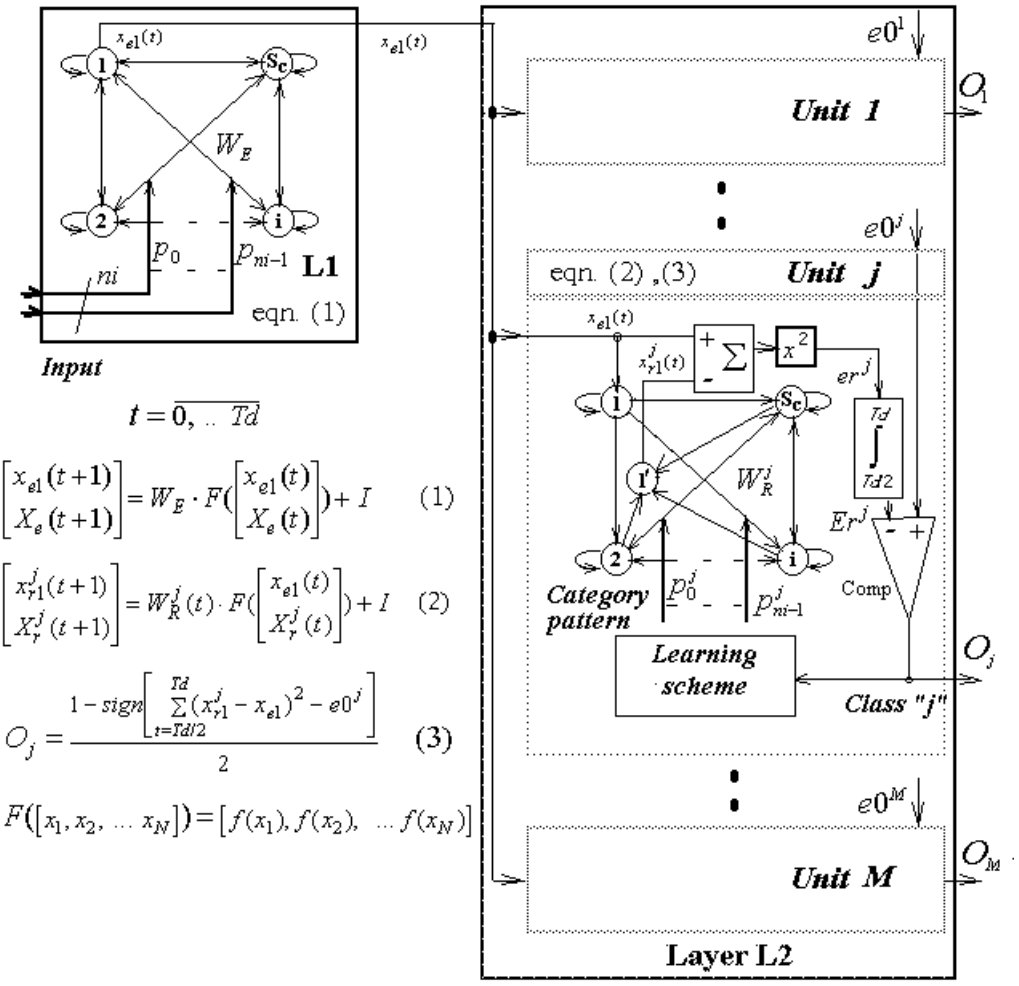
2. The Chaotic Resonance Theory neural network

2.1. Architecture and operation principle

According to Fig.1, the architecture of a Chaotic Resonance Theory (CRT) neural network is

[1] Published in ICNN'95 Proceedings, Vol. 6, pp. 3048-3052

based on an ***input layer*** (L1) and a ***pattern storage and retrieval layer*** (L2). The first layer is composed by only one full connected recurrent neural network having the size (number of neurons) S_c. There are ni continuos valued inputs corresponding with some previous selected interconnection synapses of the neural network. The first layer recurrent network must be designed to achieve synchronizable chaotic behavior for the whole set of possible input patterns. The discrete-time dynamics of this network is described by equation (1) in Fig.1 and it will be more detailed explained in section 2.2. The state variable $x_{e1}(t)$ associated with the neuron "1" is the output of this layer and it codifies the whole information about the input pattern, only one signal path being required for connecting the two layers. This path replace the standard web of synapses connecting layers in standard neural networks and it can be implemented using a single wire connection or a electromagnetic field. In the last case, the information can be exchanged between the layers of the network without spatial position constraints for the two layers, this situation being similar with the telepathy phenomena.



$$t = \overline{0, \ldots Td}$$

$$\begin{bmatrix} x_{e1}(t+1) \\ X_e(t+1) \end{bmatrix} = W_E \cdot F\left(\begin{bmatrix} x_{e1}(t) \\ X_e(t) \end{bmatrix}\right) + I \qquad (1)$$

$$\begin{bmatrix} x_{r1}^j(t+1) \\ X_r^j(t+1) \end{bmatrix} = W_R^j(t) \cdot F\left(\begin{bmatrix} x_{e1}(t) \\ X_r^j(t) \end{bmatrix}\right) + I \qquad (2)$$

$$O_j = \frac{1 - sign\left[\sum_{t=Td/2}^{Td}(x_{r1}^j - x_{e1})^2 - e0^j\right]}{2} \qquad (3)$$

$$F([x_1, x_2, \ldots x_N]) = [f(x_1), f(x_2), \ldots f(x_N)]$$

Fig. 1 The architecture and the operation principle of the CRT neural network

The ("telepathic") signal is fed simultaneously to M units of the second layer (L2) all of them acting as ***chaotic memory cells*** and working on the principle of chaos synchronization. Each unit is based on the same full recurrent neural network structure as the one used in the first layer, but here the first neuron acts only as a source of signal for the others, its state being identical with x_{e1}. A duplicate (1') neuron is introduced for each unit in order to obtain the reconstructed signal $x_{r1}^j(t)$ for each unit "j". When unit "j" synchronizes with the "telepathic" signal the states of the neurons 1 and 1' are the same. The weights used as inputs in the first layer are now replaced by the output ("j" category pattern) vector

$\mathbf{p^j} = \left[p_0^j, p_1^j, \ldots p_{ni-1}^j \right]$ given by a learning scheme which acts in the sense of tuning the unit to synchronize with the input layer oscillator. The learning scheme can act in a supervised or unsupervised manner.

When the learning is supervised, it occurs in a "pattern storage" phase, each unit being assigned to a centriod (prototype) represented by the vector $\mathbf{p^j}$, while in the retrieval phase no further tuning of the L2 units occur. The simulations presented in this paper are based on supervised learning.

The unsupervised manner of operation implies no distinct storage and retrieval phases while for each new pattern presented at the input, the unit in L2 which best synchronizes is selected to be tuned in order to achieve minimum synchronization error. The selected unit acts like the winner neuron in standard ART networks and the "learning" of a new category pattern associated with that unit, corresponds with the ART resonance. The ART category pattern is replaced by the tunable vector $\mathbf{p^j}$.

Each unit from the L2 layer has an output O_j having value 1 if the unit synchronizes with the input pattern or 0 if not and thus giving the information about the membership of the input pattern to the cluster defined by the category pattern stored in that unit. The value of the output is computed after a period T_d from the moment of the input pattern presentation, according to equation (3) in Fig. 1. For the first $T_d / 2$ time periods no synchronization error is counted in order to avoid the transitory regime. Experimental results proved that the duration of this transitory regime is small (usually few clock periods) and depends only on the size S_c of the ***generic structure***.

A generic structure corresponds with the complete definition of the neural network used in L1 layer which includes the size, the nonlinear activation function, biases and the interconnection weights values with the specification of the ones used as inputs. For each unit in the L2 layer a ***tolerance*** parameter $e0^j$ is defined, having an opposite significance with the ART ***vigilance*** parameter, while very small values (closed to 0) allows synchronization (i.e. retrieval) only for very good matching between the input and the category patterns. Instead of ART networks, in the CRT based ones, the radius of the centroids can be independently controlled for each unit.

The resonance process which occurs in ART networks due to the feedback between layers is now replaced with a chaos synchronization process which does not involve feedback between layers, only one physical connection being required to transmit the chaotic inter-layer signal $x_{e1}(t)$.

2.2. The principle of chaos synchronization extended to neural networks

A system based on chaos synchronization [5] [6] is composed by a transmitter subsystem which is a nonlinear chaotic oscillator and by at least one receiving subsystem which is also a nonlinear system. The transmitter's evolution equation is described by two groups of differential equations: the driving subsystem and the response subsystem, the last one being the same both for transmission and receiving systems.

According to [6] while the conditional Liapunov exponents of the response subsystem are negative, the receiving subsystem can track the state of the transmitter i.e. synchronize with it. The above mentioned statement is known as the Pecora-Carroll theorem.

This principle can be extended to neural networks while they are particular cases of nonlinear systems. Equation (1) in Fig.1. describes the dynamics of the transmitter system used in layer L1, while equation (2) is associated with the dynamics of the unit "j" . When the weight matrices of the transmitter and receiving unit "j" are identical ($W_E = W_R^j$), unit "*j*" perfectly synchronizes (i.e. the squared error is 0) with the chaotic signal from the layer L1 as long as the condition imposed by the Pecora-Carroll theorem is satisfied.

While the principle of the CRT implies input pattern presentation as variable weights, an appropriate design of the generic structure must be considered in order to assure the Pecora-Carroll theorem is satisfied for the whole set of input patterns.

We define the ***storage set*** $Ss \subset R^{ni}$ as the compact set of weight parameters used as inputs or category patterns for which chaos synchronization can occur. It is yet a difficult task to determine this set with analytical tools. However, experimental methods such as the weight space exploration [3] [8] were efficiently used. Fig. 2 presents four ***sensitivity maps*** obtained for a generic structure with *Sc=3* having the inputs (or category patterns) associated with the *p0* and *p1* parameters. Each pixel in a sensitivity map corresponds to a well defined neural network structure, the values of the two parameters being scaled to the vertical and horizontal position of the pixel. Upper left corner corresponds with the minimum values of the parameters while the lower right corner with the maximum ones. In this example the variation domain for both *p0* and *p1* is [-1,2]. While the brightness of each pixel is proportional with the number of positive Liapunov exponents of

the associated network, black zones correspond to structural stability (0 positive Liapunov exponents) and the others corresponds with chaos. Fig. 2a. shows the map obtained from the transmitter system (the recurrent network in L1 layer) and Fig. 2b, c, d shows the maps obtained for the response subsystem when three different values (including the extremes) of the input signal were considered in order to emphasize the conditional Liapunov exponents. Using visual inspection according to the Pecora-Carroll theorem, the ***storage set*** was determined. It corresponds with the inside region of the rectangle "1" drawn on the sensitivity maps.

3. Simulation results

A CRT based neural network having the generic structure and storage set already defined was considered for simulations. It corresponds with a two continuos input network, the inputs being associated with the *p0* and *p1* parameters. The network is operated in retrieval mode and the units 1 and 2 from were previously tuned using the supervised learning strategy. The following five patterns from the storage set were consequently presented at the input of the network:

P0 = (-0.5 , 0.5) ;P1= (-0.45 , 0.5); P2=(-0.3, 0.5) P3=(-0.3, 0.7) and P4= (-0.5 , 0.7).

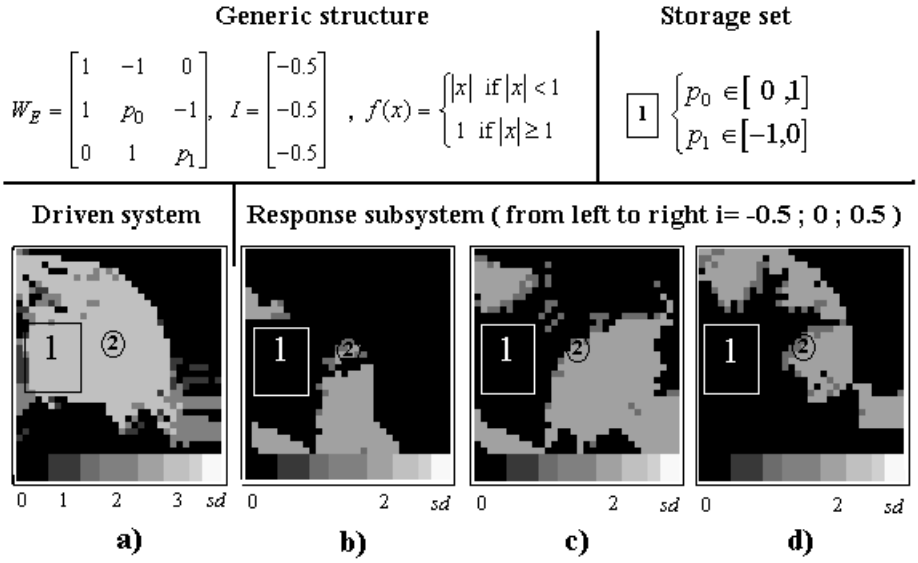


Fig.2. Generic structure used for a CRT based network.

Unit 1 was tuned to recognize the category pattern P1 while unit 2 was tuned to P2. For each input pattern a presentation period *Td=64* was considered. If the input patterns are ordered by the Euclidean distance between them and P1, the increasing order is: P1, P0, P2, P4, P3.

In Fig.3 the instantaneous quadratic error signals for units 1 and 2 are displayed. For each pattern presentation the quadratic error is higher at the beginning while after a small transitory regime (about 10 iterations) the quadratic error signal mean becomes constant and it indicates the degree of synchronization (or matching) between the input and the category patterns. Even in the presence of additive noise in the inter-layer connection (SNR=30dB) additive white noise was considered in the above simulations), the synchronization occurs. The table in Fig. 3 (obtained for unit 1) shows that by increasing the levels of the ***tolerance*** parameter, patterns with increasing distances from category pattern are assigned to the same class allowing one to control the "volume" of each centroid assigned to a category pattern. When the tolerance is smaller than 0.02 the inter-layer noise influences the decision and thus no output unit activates. Without noise in the inter-layer connection this threshold can be lowered to 0. Thus, the memory capacity is determined only by the ratio between the "volume" of the storage area and the "volume" of the most smallest cluster which can be defined in some given conditions (inter-layer or technological noise). For the above presented example and selecting low tolerance values (0.025) the "volume" of the smallest cluster was determined to be approx. 0.05x0.05 while the storage "volume" is 1x1 and thus the maximum capacity which can be achieved is about 400. To effectively implement this capacity, 400 ***independent*** L2 units and 1 transmitter unit are required, each of them with only 7 weights, thus a total amount of 2807 weights is needed, while for the same memory capacity using a Hopfield model a considerably large ***fully interconnected*** structure is necessary.

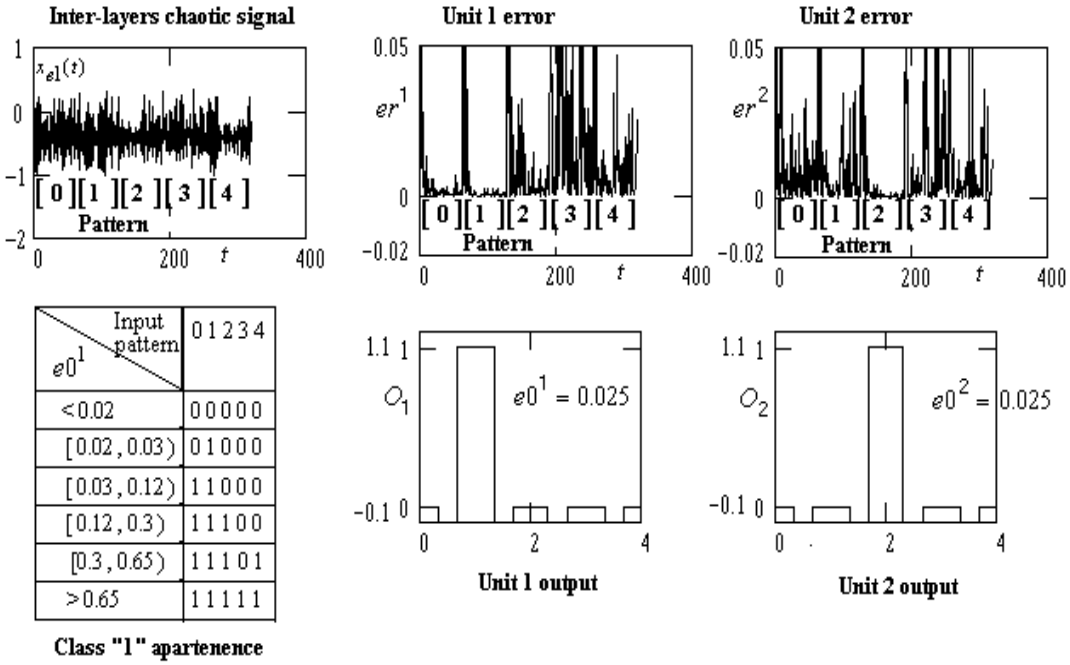


Fig. 3 Computer simulations of a CRT based neural network

It must be also noticed that Hopfield-based neural networks can store only binary patterns while the CRT based network can store any continuos-valued pattern in the storage set.

4. Conclusions

The principle of chaotic resonance theory and an artificial neural network for clustering based on this principle were presented in the paper. Simulation results show that chaos can be efficiently exploited to store and retrieve patterns in networks of small recurrent chaotic oscillators.

The behavior of this network is almost similarly with the behavior of an ART network while the principle of chaos synchronization make the structure very attractive for VLSI implementations, especially due to the single path needed for information exchange between the units of the network.

Future research will focus on finding better design strategies for the chaos synchronizable units and for the unsupervised learning strategies.

References

[1] G.A. Carpenter & S. Grossberg , "The ART of adaptive pattern recognition by a self organizing neural network ", in Computer , Vol. 21 , pp. 77-88, 1988.

[2] L.O. Chua, L Kocarev, K. Eckert and M. Itoh, " Experimental chaos synchronization in Chua's circuit " , in Int. Journal of Bifurcations and Chaos, vol.2 pp. 705-708, 1992.

[3] R. Dogaru, A.T. Murgan, "Software tools for analyzing oscillatory dynamics in neural networks", in "Fuzzy Systems and Artificial Intelligence", Vol.4 No. 1, Romanian Academy Publishing House, 1995.

[4] Fogelman-Soulié , F. "Neural network architectures and algorithms: A perspective." in Artificial Neural Networks, Vol.2, Amsterdam, North-Holland, 1991.

[5] M.J. Ogorzalek, "Taming Chaos - Part I; Synchronization", IEEE Tr. on Circuits and Systems - I , Vol. 40, No. 10, October 1993, pp. 693- 699.

[6] L.M. Pecora and T.L. Caroll, "Synchronized Chaotic Signals and Systems" in Proc. IEEE ICASSP, San Francisco CA, Vol. 4, pp. 137-140, 1992

[7] C.A. Skarda and W. J. Freeman, " How brains make chaos in order to make sense of the world" , in Brain and Behavioral Science 10, pp. 161, 1987.

[8] R. Dogaru, A.T. Murgan, I. Daniel, “Robust oscillations and bifurcations in cellular neural networks” in Proceedings of CNNA-94 (Third IEEE International Workshop on Cellular Neural Networks and their Applications), Rome, ISBN 0-7803-2070-0, pp. 303-308, 1994.