A MODIFIED CHAOTIC ASSOCIATIVE MEMORY SYSTEM FOR GRAY-SCALE IMAGES

Ke Li, Wenjiang Pei, Luxi Yang and Zhenya He, Fellow, IEEE

Dept. of Radio Engineering,

Southeast University, Nanjing 210018, China

ABSTRACT

Globally coupled map (GCM) model can evolve through chaotic searching into several stable periodic orbits under properly controlled parameters. This can be exploited in information processing such as associative memory and optimization. In this paper, we propose a novel covariance learning rule for multivalue patterns and apply it in memorization of gray-scale images based on modified GCM model (S-GCM). Analysis of the retrieval results are given finally.

1. INTRODUCTION

Recently, further understanding of chaotic dynamics and the mechanism of cognition in mammalian brain provide extensive theoretical basis for the application of chaos in information processing [1-3]. In [2] it is reported that in rabbit olfactory bulb, limit cycle activities occur for perceptible specific odors but chaotic activities occur for novel odors. Kaneko [4] proposed the "globally coupled map (GCM)" model whose main property is that by adjusting the parameters, all the units of the model will split into several periodic attractors called *cluster frozen attractors*. The units belonging to the same cluster come to follow an identical orbit. This can be viewed as a procedure evolving from chaotic searching to memory locking and to some extends it is in accordance with the biological experiment mentioned above.

Ishii et al. proposed a modified GCM model (S-GCM) [5] which can be applied to associative memory of characters. But it is invalid for multivalue patterns. Although we can convert multivalue patterns into binary ones by hierarchical coding and memorize as binary patterns, it will enlarge the network scale greatly. In addition, any retrieval error of significant bits will leads to isolate singular points. Based on the characteristics of S-GCM, we construct an associative memory system for gray-scale images. Finally, we discuss the retrieval results.

2. NEURAL NETWORK MODEL

The only difference between S-GCM and GCM model is that the former employs a cubic function instead of the logistic map. The model is described as:

$$\begin{cases} x_i(t) = (1-\varepsilon) f(x_i(t-1)) + \dfrac{\varepsilon}{N}\displaystyle\sum_{j=1}^{N} f(x_j(t-1)) & (1) \\ f(x) = \alpha x^3 - \alpha x + x \qquad x \in [-1,1] & (2) \end{cases}$$

where $x_i(t)$ denotes the ith unit's value at time t, and N the number of units. The summation part in (1) is the average feedback from all the units. ε is coupling strength. Each unit's dynamics is almost given by the cubic function described in (2). α is bifurcation parameter of $f(x)$ and for $\alpha \in [3.3, 4]$ the function is in chaotic state. The cubic function has two extrema in its range when $\alpha > 2$.

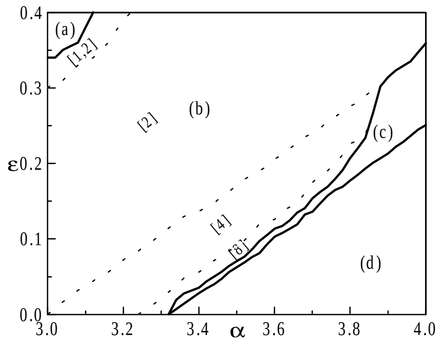


Fig. 1. Phase diagram of S-GCM

The spatiotemporal features of the S-GCM attractors are determined by both α and ε. As α increases, the S-GCM becomes chaotic and as ε increases, it becomes coherent, or stable. Fig. 1 shows a rough phase diagram of the S-GCM, where the attractors are classified according to their spatiotemporal features [4-5]:

1. Coherent phase (see area (a) in Fig. 1): When α is small and ε is large, all the units fall into the same orbit called "coherent attractor".

2. Ordered phase (area (b)): Cluster frozen attractors occur in this phase. It is divided into several subareas by the dominant number of clusters marked in square brackets in Fig. 1. The number of clusters increases in the manner of 2,4,8,... as α increases or ε decreases. The borders between these subareas become vague as ε increases. The S-GCM has a much larger ordered area than the GCM [5].

3. Partially ordered phase (area (c)): Attractors fall into a large number of clusters in some cases and a small number of clusters in other cases, i.e. the attractors vary

depending on their initial states.

4. Turbulent phase (area (d)): When α is large and ε is small, the dominant component of feedback is from each unit itself, and each unit follows its own chaotic orbit.

In the follows we analyze the dynamics of the model by exploiting the Lyapunov exponent. For the coherent attractor with x_i=x_j for all i, j, the motion is governed just by the single cubic map. The Jacobi matrix J_t for Eq. (1) is

$$J_t = f'(x(t))[(1-\varepsilon)I + (\varepsilon/N)D] \quad (3)$$

where I is identity matrix and D the constant matrix whose elements are all 1. Then the stability of this attroctor is calculated by:

$$\Gamma_j = \lim_{t\to\infty}(j\text{th eigenvalue of } \prod_{k=1}^{t} J_k)^{1/t},\ j=1,...,N \quad (4)$$

After simple algebra, we obtain the eigenvalues $\Gamma_1 = \gamma$ and $\Gamma_j = \gamma(1-\varepsilon)\ (j = 2,...,N)$. Here γ=exp($\lambda_0$) and λ_0 is the Lyapunov exponent of cubic map. The eigenvector corresponding to the eigenvalue γ is given by $(1/\sqrt{N})(1,1,...,1)^T$. Thus the amplification of a disturbance along this eigenvector does not destroy the coherence. Eigenvectors for the other N-1 eigenvalues are not uniform and the amplification (which occurs if $|\Gamma_j|>1$) along these vectors will destroy the coherence. Thus the stability condition of the coherent attractor is given by $|\gamma(1-\varepsilon)|<1$.

If our system falls into a k-cluster attractor with (N_1,N_2,...,N_k) where N_i is the number of elements for the ith cluster and $N_1 \geq N_2 \geq \cdots \geq N_k$, our system will never goes out of the state with this clustering if the motion for x_i and x_j at time $t'>t$ are governed by exactly the same dynamics, i.e. $x_i(t')=x_j(t')$ exists if $x_i(t)=x_j(t)$. Therefore, to destroy the coherence within a cluster, amplification of a small disturbance x_i-x_j must occur in it. This condition is again calculated by the products of Jacobi matrices for (1), as in Eq. (4). If all the absolute values of eigenvalues of (4) are less than 1, the state is obviously stable (the attractor is periodic). If all the elements of an eigenvector corresponding to the eigenvalue whose absolute value is larger than unity take an identical value in each cluster, the amplification of a disturbance along this eigenvector does not destroy the coherence within each cluster as well.

In the partial ordered area and turbulent area, numerical calculation gives that the largest Lyapunov exponent is large and its value is positive [5], which means that the S-GCM in those areas is chaotic.

3. MULTIVALUED PATTERN ASSOCIATIVE MEMORY

When the parameters are in the ordered area, the output of the network is almost equal to the input which means the network is in preserving mode. With some parameter values in the turbulent area, each unit shows a chaotic searching motion, i.e. the network is in searching mode. In binary pattern associative memory system [5], parameter α or ε is adjusted according to a match criterion to control each unit switching between this two modes. The time when all the units evolve into preserving mode and fall in cluster frozen attractors, the output of the network is the memorized pattern required.

Let each cluster represent one kind of values in a pattern, and modify the learning rule accordingly, the S-GCM can be extended to process multivalue patterns. Here we show this process by exploiting four-cluster frozen attractor to memorize four-gray-scale images.

Because the state value of each unit is continuous and the element value of pattern is discrete, conversions are needed at the input and output of the network. Assume that the final state values of the four clusters are x_a, x_b, x_c and x_d, respectively in decreasing order, the correspondent element values in pattern are 3, 2, 1 and 0. The convert function of the input is defined as:

$$s_i(x) = \begin{cases} x_a + rnd & x=3 \\ x_b + rnd & x=2 \\ x_c + rnd & x=1 \\ x_d + rnd & x=0 \end{cases} \quad (5)$$

in which rnd is a small random number. The output of $s_i(x)$ is also the initial state of the network.

Given a set of N-dimensional sample patterns $\{p^1,p^2,...,p^m\}$, $p_i^k \in \{0,1,2,3\}$, i=1,2,...,N. We construct a multivalue covariance learning rule similar to Hebb rule:

$$w_{ij} = \frac{1}{m}\sum_{k=1}^{m}\text{sgn}(p_i^k - p_j^k)\cdot\sum_{k=1}^{m}\left[1-\frac{1}{3}\left|p_i^k - p_j^k\right|\right] \quad (6)$$

where sgn(x)=1 or -1 for $x \geq 0$ or $x<0$, respectively.

Since the value of α in each unit is adjusted continuously and unit-wisely in the procedure of retrieval, it is time-vary and different in each unit. Therefore, the model should be rewritten as:

$$\begin{cases} x_i(t) = (1-\varepsilon)f(x_i(t-1)) + \dfrac{\varepsilon}{N}\displaystyle\sum_{j=1}^{N} f(x_j(t-1)) & (7) \\ f(x) = \alpha_{i;t}x^3 - \alpha_{i;t}x + x \qquad x\in[-1,1] & (8) \end{cases}$$

The value of $\alpha_{i;t}$ is adjusted according to the energy function of each unit which is defined as:

$$E_i = \sum_{j=1}^{N}\left|w_{ij} - s_{ij}\right|, \quad i = 1,2,...,N \quad (9)$$

$$s_{ij} = \text{sgn}[s_o(x_i) - s_o(x_j)]\cdot(1-1/3\,|\,s_o(x_i) - s_o(x_j)\,|) \quad (10)$$

in which $s_o(x)$ is the inverse of $s_i(x)$. It converts the state of network to the output pattern:

$$s_o(x)=\begin{cases}3 & x\ge (x_a+x_b)/2\\ 2 & (x_b+x_c)/2\le x<(x_a+x_b)/2\\ 1 & (x_c+x_d)/2\le x<(x_b+x_c)/2\\ 0 & x<(x_c+x_d)/2\end{cases} \quad (11)$$

$\alpha_{i;t}$ is adjusted according to the matching degree of present state value x_i to sample pattern which is evaluated by E_i:

$$\alpha_{i;t}=\{\alpha_{i;t}+(\alpha_{i;t}-\alpha_{min})\cdot\tanh[\gamma(1-E_i)(E_i-Th)]\}^{\#} \quad (12)$$

Each α is controlled to be between α_{min} and α_{max}. $f^{\#}=(f\vee\alpha_{min})\wedge\alpha_{max}$ is cutoff function, γ is a factor controlling the adjustment step. Th is the threshold of energy function, here Th=0.5. In order to arrive at a tradeoff between sufficiently searching in state space with present α of each unit and a much fast convergence rate, the modifying of α is done once every 2 time steps.

If E_i is greater than Th which means that the unit value x_i does not suit the pattern to be retrieved, α_i increases and this unit will evolve into chaotic motion and search for the proper state within the range. When the unit suits the pattern, E_i becomes small and α_i decreases. In this case, the unit-wisely processing mode will change from searching to preserving. When every unit falls into preserving mode, all the α_i will reduce to α_{min} eventually and the whole system reaches cluster frozen attractor. Thereby the output of the network equals to the pattern to be retrieved. If the input contains too high a noise as to run out of the basin volume of the corresponding pattern or the input isn't one of the memorized patterns at all, the system will always be in chaotic searching and the energy keeps a high value.

Therefore, we have the criterion for termination of iterating: If the system's total energy is great than a threshold in certain time steps t_s (t_s is determined by trial and error on considering the noise level and network scale), the association procedure fails and will be stopped; If not, the procedure will be terminated till all the α_i decrease to α_{min} and the output is just the retrieved pattern.

4. EXPERIMENTAL RESULTS

Given four 32×32 four-gray-scale images as learning samples (see Fig. 2), the corresponding network scale is N=32×32=1024. Let ε=0.1, α_{min}=3.4, α_{max}=4.0. The input is pattern A disturbed with 10% random noise. Fig. 3 shows the output of network at different time. Fig. 4 gives the evolution of several parameters and indices of the network in which the normalized total energy of the system is $E(t)=1/N\sum_{i=1}^{N}E_i(t)$ and the Hamming distance of multivalue patterns is

$$D_h(x,y)=\sum_{i=1}^{N}|x_i-y_i|,\quad x_i,y_i\in\{0,1,2,3\} \quad (13)$$

In Fig. 4, highly chaotic motions are observed at the early association state. As time elapses, these motions become quiet, and the association is completed when the system falls into a four-cluster frozen attractor after about 400 times of iteration. The Hamming distance between pattern A and the output of the system decreases from 102 to 5 finally. According to Amit [6], the network succeeds in memorizing m patterns, if the error rate of elements in retrieved pattern after the transient period is less than 1.5% when the input is one of the memorized patterns. So the retrieval is successful in this context.

Tab. 1 gives the numerical results of the four patterns under different noise levels. The denominator and numerator in each item are the Hamming distances of the input and final output pattern with the corresponding noise free pattern. We can see that the association is successful even under 20% noise.

Tab. 1 Association of patterns under different noise levels

	Pattern A	Pattern B	Pattern C	Pattern D
10% noise	5/102	7/131	6/119	6/108
20% noise	9/218	8/208	11/230	7/210

5. CONCLUSION

In this paper we analyze the properties of S-GCM model and extend the binary pattern associative memory system to multivalue pattern case. The experiments show that by properly adjusting parameters, the associative memory network based on multivalue covariance learning can successfully memorize and retrieve gray-scale images. Although there are still a few pixels cannot be restored completely, the difference is fairly small. In case of big images with more gray scale levels, the influence of this difference to the whole image is nearly neglectable according to human visual psychophysics.

6. REFERENCES

[1] C.A. Sakarda and W.J. Freeman, How brains make chaos in order to make sense of the world, *Behavioral and Brain Sciences*, vol.10, no.1, pp.161-195, 1987

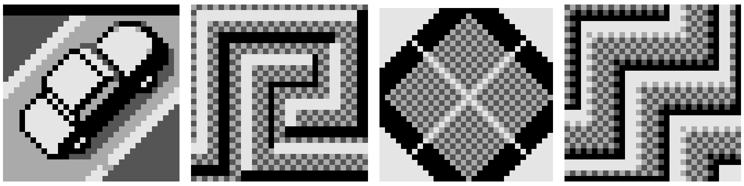
[2] Y. Yao and W.J. Freeman, Model of biological pattern recognition with spatially chaotic dynamics, *Neural Networks*, vol.3, no.1, pp.153-170, 1990

[3] Y.V. Andreyev, Y.L. Belsky, A.S. Dmitriev and D.A. Kuminov, Information processing using dynamical chaos: neural networks implementation, *IEEE Trans.on Neural Networks*, vol.7, no.2, pp.290-299, 1996

[4] K. Kaneko, Clustering, coding, switching, hierarchical ordering, and control in a network of chaotic elements, *Physica D*, vol.41, no.1, pp.137-172, 1990

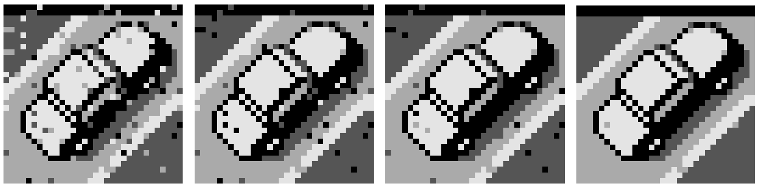
[5] S. Ishii, K. Fukumizu and S. Watanabe, A network of chaotic elements for information processing, *Neural Networks*, vol.9, no.1, pp25-40, 1996

[6] D.J. Amit, H. Gutfreund and H. Sompolinsky, Statistical mechanics of neural networks near saturation. *Annals of Physics*, vol.173, no.1, pp.30-67, 1987



Pattern A Pattern B Pattern C Pattern D

Fig. 2 Sample patterns



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

Fig. 3 Association procedure of the network. (a) is the input pattern A with 10% random noise; (b), (c) and (d) are the output of the network after 100, 200 and 400 times of iteration, respectively.

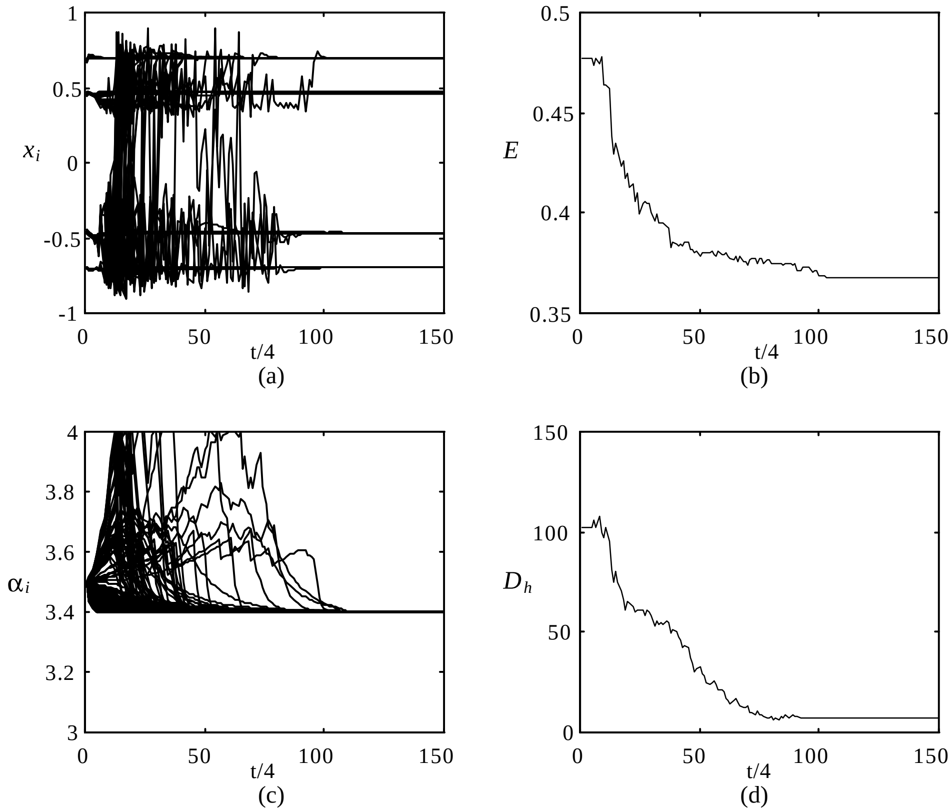


Fig. 4 Evolution of parameters and indices in the network. (a) Time series of each unit's output; (b) Time series of normalized total energy E; (c) Time series of α_i ; (d) Time series of Hamming distance