Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Bit Importance Strategy For Bidirectional Associative Memory

Mo Wei and Jose B. Cruz, Jr., *Life Fellow, IEEE*

***Abstract*—In this paper, a strategy named Bit Importance Strategy (BIS) is introduced to improve the performance of Bidirectional Associative Memory (BAM) in cases in which differences in importance among bits must be considered. The main advantage of this new strategy lies in the improvement of the recall success ratio. Simulation results are given, with an application to a simplified version of a military battlefield operation, based on a theoretical analysis.**

I. Introduction

ACCORDING to Kosko [1-3], Bidirectional Associative Memory (BAM) is a class of 2-level nonlinear artificial neural networks. In general, BAM uses a matrix to store input/output pairs, Kosko [1-3], Cruz and Stubberud [4], and then identifies the closest possible output according to the input, which might be noisy. The advantage of BAM is that it can filter noise in the recall procedure and much effort have been devoted to improving its encoding strategies, storage capacity, retrieval success ratio, and other indicators of performance.

Leung [5] used projections on convex sets to refine the strategy of BAM. Araujo and Haga [6] proposed unlearning and delta rule to improve the performance of BAM. Haines and Hecht-Nielsen [7] proposed an approach to increase the storage capacity of BAM. Wang and Don[8] showed that exponential non-linearity in the evolution equations can improve the storage capacity of BAM. Kawabata, Daido and Hattori [9] investigated the capacity of second-order BAM in the presence of noise bits. Leung [10] investigated optimum learning for BAM capacity. Tai [11] proposed a higher-order BAM so that the capacity and error-correcting capability can be improved. Wang, Cruz, and Mulligan [12] proposed and investigated multiple training in BAM. Leung, Wang, and Lee [13-14] studied the influences of dummy elements in BAM. Cao [15] investigated Exponential stability of delayed BAM. Liao and Wong [16] analyzed robust stability of interval BAM with delays. Lopez-Aligue and Acevedo-Sotoca [17] implemented a handwritten pattern recognition system via BAM and attained very high retrieval success ratio.

For the basic procedure of BAM retrieval, refer to the appendix. More details are found in Wang, Cruz, and Mulligan [18]. The retrieval procedure with conventional binary encoding is indicated in Fig. 1.1.

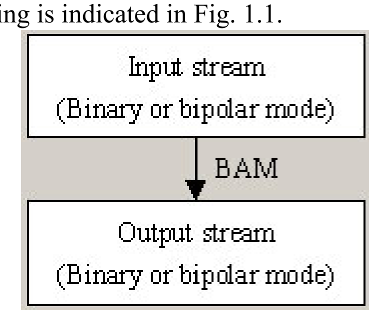

Figure 1.1 Retrieval process with conventional binary encoding

BAM strategy treats all bits in the information streams identically and it is most suitable for recalling uniform bit importance training pairs such as digital pictures. Here uniform bit importance means that in the information stream that constructs the training pairs, all the information bits have the same information importance.

However, there are many problems in which differences in importance of the bits need to be considered. For example, for a decimal number 128, the left-most bit 1 is more important than the right-most bit 8.

In these situations, BAM might show similar performance for different amplitudes of noises. For an original decimal number 128, a decimal noise 100 might cause the same recall error rate as a much smaller decimal noise 10, for both of them will cause the original number 128 to change just one bit. Since pure conventional BAM calculation is actually for uniform-importance, the two kinds of bit changes are considered at the same information level, although from the perspective of a human the smaller noise should be considered "weaker" and easier to filter.

In other words, conventional BAM strategy does not consider the information importance differences among the bits in the input/output information streams. To get better recall performance, we propose a new encoding/decoding algorithm which can reflect the differences in information importance.

Manuscript received March 6, 2005. This work was partially sponsored by the Defense Advanced Project Agency (DARPA) under ContractF33615-01-C3151 issued by the AFRL/VA.

M. Wei is with The Ohio State University, Columbus, OH 43210 USA (email: wei.73@osu.edu).

J. B. Cruz, Jr. is with The Ohio State University, Columbus, OH 43210 USA (email: jbcruz@ieee.org).

II. Bit Importance Strategy and Nonlinear Recovery Strategy for BAM

Bit Importance Strategy (BIS) can be stated as follows.

0-7803-9568-9/05/$20.00 ©2005 IEEE

Definition 2.1 *If there is a set of bits arranged in an order such as* $(a_n, a_{n-1}, \cdots, a_{i+1}, a_i, \cdots, a_2, a_1)$ *in the BAM input/output training pair, and among the elements there are differences in importance such as:*

$$weight(a_{i+1}) > weight(a_i) \qquad (2.1)$$

Then assign $2 \times i - 1$ *identical bits to represent the original bit* a_i*. This strategy is called Bit Importance Strategy (BIS).*

For example, if we have a binary number 0010, after applying this bit importance strategy, we get 0000000000001110. The first 7 zeroes stand for the first 0 in the original number 0010, and the next 5 zeroes stand for the next 0 in the original number, and the next three 1's represent the 1 in the original number 0010, etc.

We use *Nonlinear Recovery Strategy (NRS),* corresponding to *Bit Importance Strategy*, to decode a number after BAM filtering.

Definition 2.2 *When the BAM calculation ends (we get an equilibrium) and we need to decode the corresponding output, we check what value (1 or 0) occupies the majority in this bit section, and choose this corresponding bit to be the recovered output bit. This strategy is called Nonlinear Recovery Strategy (NRS).*

Here is an example for NRS. Suppose we get a recalled stream

001010000 1110011 01000 101 0

and it should be recovered into a binary number $a_5 a_4 a_3 a_2 a_1$.

By applying NRS we will have

001010000 $--->$ 0 is the majority $--->$ recover as 0
1110011 $--->$ 1 is the majority $--->$ recover as 1
01000 $--->$ 0 is the majority $--->$ recover as 0
101 $--->$ 1 is the majority $--->$ recover as 1
0 $--->$ 0 is the majority $--->$ recover as 0

thus get 0 1 0 1 0, where $a_5 = 0$, $a_4 = 1$, $a_3 = 0$, $a_2 = 1$, $a_1 = 0$.

After applying BIS and NRS to a conventional BAM, the new retrieval/recall procedure is indicated in Figure 2.1.

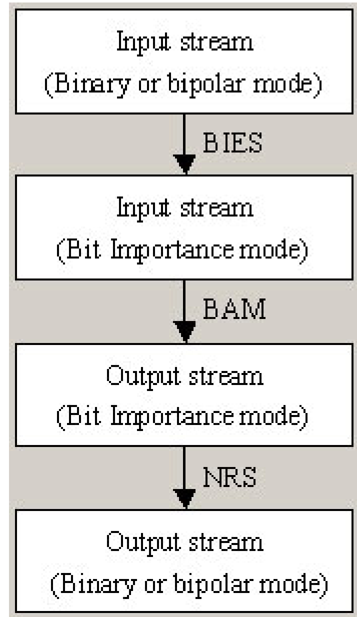

Figure 2.1 Retrieval process with BIS/NRS

For example, suppose we have a training pair (5,4). 5 is the input, and 4 is the corresponding output. We want to retrieve the output from the input. When there is no noise, a successful retrieval might be

input 0101 $\xrightarrow{\text{BitImportance}}$ 0000000111110001 $\longrightarrow$
$\xrightarrow{BAM}$ 0000000111110000 $\xrightarrow{\text{NonlinearRecover}}$ 0100 output

If the input is noisy, say the amplitude of noise is 2, a successful retrieval might be

input 0111 $\xrightarrow{\text{BitImportance}}$ 0000000111111111 $\longrightarrow$
$\xrightarrow{BAM}$ 0000000111010100 $\xrightarrow{\text{NonlinearRecover}}$ 0100 output

III. Theoretical analysis of Bit Importance Strategy

The idea for this strategy is from two basic assumptions. In the next section, experiments will show that conclusions based on these two assumptions are correct.

Assumption 1: *The higher the percentage of error bits in the input stream, the higher the probability of recalling wrong outputs.* In other words, if there are more errors in the inputs, it is more likely to get wrong outputs.

Assumption 2: *For a BAM system, each bit of the recalled result (after BAM recalling rotation ends, also before applying the nonlinear recovery strategy), has the same bit-error probability to be a wrong bit. This bit-error probability increases when input error percentage increases.* This assumption relates the input error percentage to the output bit-error probability.

To analyze BIS and NRS theoretically, we start from the case in which there is only one wrong bit in the original binary stream. If a noise affects more than one bit, the wrong bits can be considered one by one by exactly the same procedure.

Theorem 3.1: *If assumptions 1 and 2 hold and there are weight differences among the bits in the input stream, there exists* $r^* > 0$ *such that if the bit-error probability* r *satisfies* $0 < r < r^*$*, the Bit-Importance strategy improves the retrieval success ratio for the same bit-error probability* r.

Proof: Suppose we have a recalled bit importance mode number such as

$$b_{m,2m-1}b_{m,2m-2}\cdots b_{m,1}\cdots\cdots\cdots b_{n,2n-1}\cdots b_{n,1}\cdots\cdots\cdots b_{1,1}$$

and it corresponds to a common binary mode number after nonlinear recovering strategy

$$b_m^{'}\cdots b_n^{'}b_{n-1}^{'}\cdots b_1^{'}$$

That is, $b_{k,2k-1}b_{k,2k-2}\cdots b_{k,1}$ corresponds to $b_k^{'}$.

If we directly apply conventional BAM strategy to get this recalled binary mode number without using the Bit Importance Strategy and Nonlinear Recovery Strategy, we see that for the conventional binary encoding approach, to get a correct recall for a number (all bits in this number should be correct), the probability will be

$$P_{\text{conventional\_correct}} = \prod_{n=1}^{m}(1 - P_{\text{conventional\_1biterror}}) = (1-r)^m \tag{3.1}$$

If we use the BIS, we need to use NRS to get this binary mode number. After applying the nonlinear recovery strategy, the total probability that we will get a wrong bit value for the bit $b_n^{'}$ via the bit-error probability r will be

$$P_e(n) = C_{2n-1}^{0}r^{2n-1} + C_{2n-1}^{1}r^{2n-2} + \cdots + C_{2n-1}^{n-1}r^{n} \tag{3.2}$$

Here the first term $C_{2n-1}^{0}r^{2n-1}$ is equal to $C_{2n-1}^{2n-1}r^{2n-1}$, which corresponds to all the bits being wrong. Similarly, $C_{2n-1}^{n-1}r^{n}$ stands for the case that n bits in all $(2n-1)$ bits are wrong.

We can rewrite it as

$$P_e(n) = r^n(C_{2n-1}^{0}r^{n-1} + C_{2n-1}^{1}r^{n-2} + \cdots + C_{2n-1}^{n-1}\times r^0) \tag{3.3}$$

We see that

$$\begin{aligned} &P_e(n) + C_{2n-1}^{0}r^{n-1} + C_{2n-1}^{1}r^{n-2} + \cdots + C_{2n-1}^{n-1}\times r^0 \\ &= C_{2n-1}^{0}r^{2n-1} + C_{2n-1}^{1}r^{2n-2} + \cdots + C_{2n-1}^{n-1}r^{n} \\ &+ C_{2n-1}^{0}r^{n-1} + C_{2n-1}^{1}r^{n-2} + \cdots + C_{2n-1}^{n-1}\times r^0 \\ &= (1+r)^{2n-1} = P_e(n) + r^{(-n)}P_e(n) \end{aligned} \tag{3.4}$$

Thus the desired final error probability resulting from calculation noise for the n-th bit in the recovered binary mode stream is

$$P_e(n) = r^n(1+r)^{2n-1}/(1+r^n) \tag{3.5}$$

For the Bit-Importance Strategy, to get a correct recall, all the bits in the recovered binary mode stream should be correct. The probability will be

$$P_{\text{BitImportance\_correct}} = \prod_{n=1}^{m}(1 - P_e(n)) \tag{3.6}$$

From (3.5), we can get $P_e(1) = r$, which is same as $P_{\text{conventional\_1biterror}}$ in (3.1). For $n \geq 1$, we can see

$$\begin{aligned} P_e(n+1) &= r^{n+1}(1+r)^{2n+1}/(1+r^{n+1}) \\ &= P_e(n)\times(r(1+r)^2(1+r^n)/(1+r^{n+1})) \end{aligned} \tag{3.7}$$

Since r is the bit error probability, thus $0 \leq r \leq 1$. This means

$$P_e(n+1) \leq P_e(n)\times r(1+r)^2(1+r^0)/(1+r) \tag{3.8}$$

If we can find a $0 < r^* < 1$ such that when $0 < r < r^*$ and $n \geq 1$

$$2r(1+r) \ < 1 \tag{3.9}$$

We will get $P_e(n+1) < P_e(n)$. This implies that if $r < r^*$, $P_e(n)$ is a decreasing function, thus $P_e(n) < r$ for all $n > 1$. Thus we can see that for any $m > 1$, that is, there is an importance difference in the BAM stream, we will have

$$P_{\text{BitImportance\_correct}} = \prod_{n=1}^{m}(1 - P_e(n)) > (1-r)^m = P_{\text{conventional\_correct}} \tag{3.10}$$

which means that the Bit-Importance strategy improves the retrieval success ratio.

r^* is very easy to find and can be selected as the positive root of equation $2r(1+r) \ = 1$. Since derivative of $2r(1+r)$ in the interval $[0, r^*]$ is always positive, (3.9) will always hold for $0 < r < r^*$. *End.*

Theorem 3.2: *If the index of the highest importance bit in one number is* m*, the complexity of calculation for the Bit Importance Strategy is* m^2 *times more than that for the conventional binary strategy.*

Proof: Since for BAM the main calculations are matrix calculations, we know that the complexity of BAM calculation changes according to the square of the length of the stream. Now, suppose the index for the highest importance bit in one number in the conventional binary stream is m, and the length of the conventional binary stream is L, then the length of the new bit importance strategy will become mL, the complexity of the calculation will be of m^2L^2 order, that is, m^2 times bigger than the conventional binary stream. *End.*

IV. SIMULATIONS

The goal for this experiment is to test the success rate of BAM with BIS/NRS and compare it with the success rate of BAM with conventional binary strategy.

There are two forces, Blue and Red, confronting each other. The Blue Unmanned Aerial Vehicles (UAVs) would like to fly to the Red Area to attack Red troops. Between the blue airbases and the Red Area there are some Red objects threatening Blue UAVs. Since the coordinates of Red objects found by Blue UAVs are often noisy, Blue UAVs will use BAM to retrieve an offline-calculated optimal path so that the Blue UAVs can avoid directly crossing over or attack the Red objects. In training pairs, coordinates of Red objects construct A_i 's and coordinates of waypoints in the stored optimal path construct B_i 's (see the Appendix).

A sample scenario of the battlefield is shown in Figure 3.1.

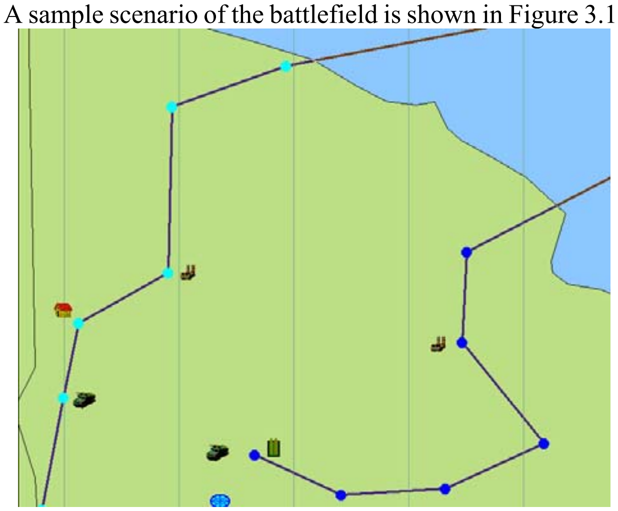

Figure 3.1 Sample scenario of the battlefield

We randomly assign the positions of the red objects and the waypoints and apply the same proportion of dummy elements. For comparison we apply the same mode of normal noise for BAM with BIS/NRS and BAM with conventional binary strategy. If the recalled output (after NRS if BIS is used) is exactly the same as the output in the training pair, it is regarded as a successful recall; if not it is counted as a failure.

The simulation results are shown in Figure 3.2. In the figure, the X axis is the percentage of the wrong bits in the input stream when it is still in binary mode. For Bit Importance encoding simulations, this is the percentage before applying the Bit Importance strategy. It is seen that the success ratio using BIS is higher than the success ratio with conventional coding, as a function of the percentage of wrong bits in the input stream.

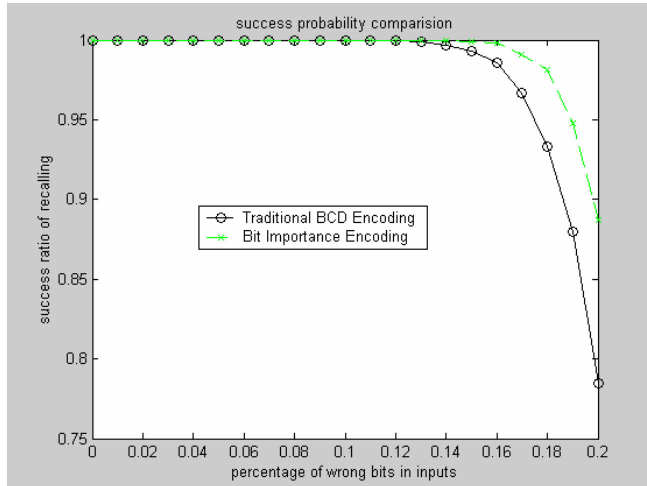

Figure 3.2 Success ratio comparison

V. CONCLUSIONS

In this paper we analyzed Bidirectional Associative Memory (BAM) strategies and discussed the influence of importance differences among the bits in the input stream. Bit Importance Strategy (BIS) and Nonlinear Recovery Strategy (NRS) are introduced to overcome the influence and improve the success ratio for BAM. Experiments support the effectiveness of the new approach.

APPENDIX

This appendix is a simplified description of the BAM retrieval procedure and is adapted from Wang, Cruz, and Mulligan [18].

Suppose there are N training pairs

$$\{(A_1,B_1),(A_2,B_2),\ldots,(A_i,B_i),\ldots,(A_N,B_N)\}$$

where

$$A_i=(a_{i1},a_{i2},\ldots,a_{in})$$
$$B_i=(b_{i1},b_{i2},\ldots,b_{ip})$$

here a_{ij},b_{ij} are in binary mode or bipolar mode:

$$\text{Binary mode: } a_{ij},b_{ij}=\begin{cases}1(on)\\0(off)\end{cases}$$

$$\text{Bipolar mode: } a_{ij},b_{ij}=\begin{cases}1(on)\\-1(off)\end{cases}$$

The correlation matrix (or weight matrix) M is

$$M=\sum_{i=1}^{N}X_i^T Y_i \qquad (6.1)$$

where $X_i(Y_i)$ is the bipolar mode of $A_i(B_i)$, respectively.

Given a test pair (α, β), to retrieve the nearest $\left(A_i, B_i\right)$ training pair from all the training pairs, the strategy is: starting with the (α, β), determine a finite sequence $(\alpha^{'}, \beta^{'})$, $(\alpha^{"}, \beta^{"})$, $(\alpha^{(i)}, \beta^{(i)})$ ….., until a final equilibrium point $(\alpha_F \ , \beta_F \ )$ is reached, where

$$\beta^{'} = \phi(\alpha M) \tag{6.2}$$

$$\alpha^{'} = \phi(\beta^{'} M^T) \tag{6.3}$$

$$\phi(F) = G = (g_1, g_2, ..., g_n) \tag{6.4}$$

where

$$F = (f_1, f_2, ..., f_n) \tag{6.5}$$

$$g_i = 1 \text{, if } f_i > 0 \text{;} \tag{6.6}$$

$$g_i = \text{ previous } \quad g_i \text{, if } f_i = 0 \tag{6.7}$$

$$g_i = \begin{cases} 0 \quad (binary), \quad if \quad f_i < 0 \\ -1 \quad (bipolar), \quad if \quad f_i < 0 \end{cases} \tag{6.8}$$

Also, we define the energy function

$$E = -\alpha^{(i)} M \beta^{(i)T} \tag{6.9}$$

Kosko [3] once showed that each cycle of decoding will decrease the energy E. Furthermore, the number of pairs $(\alpha^{'}, \beta^{'})$ in the $(\alpha^{'}, \beta^{'})$ sequence remains finite. Thus, there should be a finite number of cycles of decoding after which the pair $(\alpha_F \ , \beta_F)$ will become the final result, and the following will be a local minimum:

$$E = -\alpha_F M \beta_F^{\ T} \tag{6.10}$$

That is to say, the points, which are one Hamming distance from $(\alpha_F \ , \beta_F \ )$, will not yield a lower energy E.

One must notice that there is an important consequence of this behavior with regard to recalling a trained pair $\left(A_i, B_i\right)$. If the energy E calculated using the coordinates of the original pair $\left(A_i, B_i\right)$, that is,

$$E = -A_i M B_i^T \tag{6.11}$$

does not produce a local minimum, then the point can not be recovered even if one directly starts with $\alpha = A_i$ (no noise).

ACKNOWLEDGMENT

This research was partially sponsored by the Defense Advanced Research Project Agency (DARPA) under Contract F33615-01-C3151 issued by the AFRL/VAK. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the DARPA or AFRL.

The authors acknowledge helpful discussions with Dr. Genshe Chen.

REFERENCES

[1] B. Kosko, "Constructing an associative memory", *Byte: The Small Systems Journal,* 12(10), 137-144, Sept. 1987.

[2] B. Kosko, "Adaptive bi-directional associative memories", *Applied Optics,* vol. 26, no. 23, pp. 4947-4960, Dec. 1972.

[3] B. Kosko, "Bidirectional associative memories", *IEEE Trans. Syst., Man, Cybern.,* vol. 18, no. 1, pp.49-60, Jan./Feb. 1988.

[4] J. B. Cruz, Jr., and A. R. Stubberud, "Intelligent control of variably configured systems using neural networks", *Proc. IEEE TENCON 87*, pp. 893-989. , Aug. 1987.

[5] C.S. Leung, "Encoding method for bidirectional associative memory using projection on convex sets", *IEEE Transactions on Neural Networks*, Volume 4 , No. 5, pp.879- 881, Sept. 1993.

[6] A.F.R. Araujo, G.M. Haga, "Two simple strategies to improve bidirectional associative memory's performance: unlearning and delta rule", *International Conference on Neural Networks*, 1997, Volume 2, pp. 1178 – 1182, 9-12 June 1997.

[7] K. Haines, R. Hecht-Nielsen, "A BAM with increased information storage capacity", *IEEE International Conference on Neural Networks*, 1988, pp. 81-191, 24-27 July 1988.

[8] C.-C. Wang, H.-S. Don, "High-capacity discrete exponential BAM", *International Joint Conference on Neural Networks*, Volume 2, pp. 178-183, 7-11 June 1992

[9] Y. Kawabata, Y. Daido, S. Hattori, "Capacity of second-order bidirectional associative memory in the presence of noise bits", *International Joint Conference on Neural Networks*, 1999, Volume 2, pp.781-786, 10-16 July 1999.

[10] Chi Sing Leung, "Optimum learning for bidirectional associative memory in the sense of capacity", *IEEE Transactions on Systems, Man and Cybernetics*, Volume 24 , No. 5, pp. 791-796, May 1994.

[11] H. M. Tai, "High-order bidirectional associative memory", *Electronics Letters*, Volume 25, No. 21, pp.1424 -1425, 12 Oct. 1989.

[12] Y.-F. Wang, Cruz, J.B., Jr., J.H. Mulligan, Jr., "On multiple training for bidirectional associative memory", *IEEE Transactions on Neural Networks*, Volume 1, No. 3, pp. 275-276, September 1990

[13] C.S. Leung, "The performance of dummy augmentation encoding", *Proceedings of 1993 International Joint Conference on Neural Networks*, 93-Nagoya, Volume 3, pp. 2674 - 2677, 25-29 Oct. 1993.

[14] W.-J. Wang, D.-L. Lee, "A modified bidirectional decoding strategy based on the BAM structure", *IEEE Transactions on Neural Networks*, Volume 4, No. 4, pp. 710 - 717 , July 1993.

[15] Jinde Cao, "Exponential stability of delayed bi-directional associative memory networks", *Proceedings 5th International Conference Signal Processing, 2000.* WCCC-ICSP 2000, Volume 3, pp.1730 -1734, 21-25 Aug. 2000.

[16] Xiaofeng Liao, Kwok-wo Wong, "Robust stability of interval bidirectional associative memory neural network with time delays", *IEEE Transactions on Systems, Man and Cybernetics*, Part B, Volume 34, No. 2 , pp. 1142 - 1154, April 2004.

[17] F.J. Lopez-Aligue, I. Acevedo-Sotoca, C.G. Orellana, H.G. Velasco, "Associative memories for handwritten pattern recognition", *Proceedings of the International Joint Conference on Neural Networks*, 2003, Volume 2, pp. 1441-1446, 20-24 July 2003.

[18] Yeou-Fang Wang, Jose B. Cruz, Jr., and James H. Mulligan, Jr., "Two Coding Strategies for Bidirectional Associative Memory", *IEEE Trans. On Neural Networks*, Vol. 1, No. 1, pp. 81-92, March 1990.