Neural Processing Letters (2006) 24:67–80
DOI 10.1007/s11063-006-9012-y

Distance-Based Sparse Associative Memory Neural Network Algorithm for Pattern Recognition

LEI CHEN,[1,2] and SONGCAN CHEN[1,*]
[1]*Department of Computer Science & Engineering, Nanjing University of Aeronautics and Astronautics, 29 YuDao Street, Nanjing 210016, China. e-mail: chenlei@njupt.edu.cn*
[2]*Department of Computer Science & Technology, Nanjing University of Posts and Telecommunications, 66 XinMoFan Road, Nanjing 210003, China. e-mail: s.chen@nuaa.edu.cn*

Abstract. A sparse two-Dimension distance weighted approach for improving the performance of exponential correlation associative memory (ECAM) and modified exponential correlation associative memory (MECAM) is presented in this paper. The approach is inspired by biological visual perception mechanism and extensively existing sparse small-world network phenomenon. By means of the approach, the two new associative memory neural networks, i.e., distance-based sparse ECAM (DBS-ECAM) and distance-based sparse MECAM (DBS-MECAM), are induced by introducing both the decaying two-Dimension distance factor and small-world architecture into ECAM and MECAM's evolution rule for image processing application. Such a new configuration can reduce the connection complexity of conventional fully connected associative memories so that makes AM' VLSI implementation easier. More importantly, the experiments performed on the binary visual images show DBS-ECAM and DBS-MECAM can learn and recognize patterns more effectively than ECAM and MECAM, respectively.

Key words. associative memory (AM), neural network, sparse connection architecture, exponential correlation associative memory (ECAM), distance based training algorithm, pattern recognition

1. Introduction

Associative memory (AM) [1] is an important neural network model that can be employed to mimic human thinking and machine intelligence. So far, many researchers have been putting their efforts to explore AMs' theories and applications. The reasons why AMs draw the attention of so many researchers lie in the following aspects: (1) they have massively parallel-distributed configuration which leads to fast computation; (2) they can deal with noisy or incomplete information; (3) they have content access ability, and (4) they are strongly robust to noise. Pattern recognition is the most important practical application of AMs. The task is to produce a clear, noise-free pattern at the output when the input vectors are noisy versions of the trained patterns.

*Corresponding author.

In the past decades, researchers have proposed many AMs, among which Hopfield associative memory (HAM) [2] is most popular. HAM has yielded a great impact on the development of neural networks. The storage capacity of the HAM has been found, both empirically [2] and theoretically [3], to scale less than linearly (approximately $N/\log N$) with the number of components in memory patterns. Psaltis and Park [4], Dembo and Zeitouni [5, 6], and Sayeh and Han [7] all proposed new architectures, i.e. High-Order Correlation Associative Memory (HOCAM) and Potential-Function Correlation Associative Memory (PFCAM), respectively, that utilize nonlinear circuits and correlations between memory patterns and the input pattern to improve HAM's storage capacity. However, the storage capacity of the HOCAM is only asymptotically proportional to N^q, where q is greater than 1 and denotes an order of the high-order polynomial used. Although the PFCAM's capacity grows exponentially with the number of components in memory patterns, the primary disadvantage of this model is that hardware implementation can be cumbersome [8]. Later, Chieuh and Goodman [8] proposed Exponential Correlation Associative Memory (ECAM) with comparable storage capacity to PFCAM which has an advantage of easier hardware implementation. On the other hand, many efforts have also been made and focused on the two aspects: (1) one is to enlarge the processed data (value) range from just binary to multi-value; (2) the other is to change recalling mode from auto-association to hetero-association to realize multi-directional association. Among these improved models, the typical include Multi-valued Exponential Correlation Associative Memory (MVECAM) [9], Modified Multi-valued Recurrent Correlation Associative Memory with Exponential function (MEMRCAM) [10], Bidirectional Associative Memory (BAM) [11], High-Order Bidirectional Associative Memory (HOBAM) [12], Exponential Bidirectional Associative Memory (eBAM) [13], and Improved Exponential Bidirectional Associative Memory (IeBAM) [14], and so on.

Here, we mainly focus on the binary auto-associative memory neural networks for pattern recognition. Although ECAM holds the advantages of enough high-storage capacity and easy hardware implementation, it is neither robust enough nor realistic for real-world pattern recognition because its recalling performance reduces sharply when noise is added to a certain extent and its VLSI implementation becomes very difficult when the number of neurons in the fully connected ECAM is large. Therefore, in this paper, we tend to seek for a robust and workable neural network algorithm performed on binary visual images, that is to say, incorporating two-Dimension distance factor to improve noise-tolerance performance and introducing sparse configuration to reduce connection complexity. Firstly, a new Modified Exponential Correlation Associative Memory (MECAM) comparable performance to ECAM is proposed by changing similarity measurement in its evolution equation. However, ECAM and MECAM are still not sufficiently robust to salt and pepper noise. Inspired by the idea of "lateral feedback" mechanism in Self-Organizing Map [15], we construct Distance-based Exponential Correlation Associative Memory (DB-ECAM) and Distance-based Modified Exponential Correlation

Associative Memory (DB-MECAM) by introducing linear or nonlinear two-Dimension decaying distance factors which embody certain thought of neurophysiology into ECAM and MECAM for image processing applications. Although DB-ECAM and DB-MECAM achieve more robust recognition performance than ECAM and MECAM, respectively, for network model with thousands or even more neurons, the fully connected complex configuration will result in the difficulty of hardware implementation. So, how to reduce the complexity of neural networks but still keep their good recall or recognition performance becomes a very urgent and realistic problem. The idea of "small-world networks" described recently by Watts and Strogatz [16] provided such a possibility of relaxing this issue. Therefore, we construct two new sparse neural networks, Distance-based Sparse ECAM (DBS-ECAM) and Distance-based Sparse MECAM (DBS-MECAM), by incorporating the thinking of small-world architecture (SWA) into DB-ECAM and DB-MECAM. Experiments performed on bi-value visual images including 52 characters and 10 numbers show that the new approach can yield almost the same retrieval performance as the fully connected configuration while only using 40% of the total connectivity or synapses. Furthermore, the simulation results on visual images with bigger resolution reveal that DBS-MECAM with sparse configuration has distinctively better noise-tolerance performance than both MECAM and DB-MECAM, which means that the effect of sparse configuration may be more outstanding when neurons of the neural network is more.

The rest of the paper is organized as follows. In Section 2, DB-ECAM and DB-MECAM are introduced in detail. And SWA-inspired DBS-ECAM and DBS-MECAM are proposed in Section 3. In Section 4, we present simulation results on bi-value visual images including 52 characters, 10 numbers, and 50 other bi-value visual images. Conclusions as well as some discussion of future work are given in Section 5.

2. Distance-Based Exponential Correlation Associative Memory and Distance-Based Modified Correlation Associative Memory

2.1. EXPONENTIAL CORRELATION ASSOCIATIVE MEMORY AND MODIFIED EXPONENTIAL CORRELATION ASSOCIATIVE MEMORY

Suppose MN-dimensional bipolar (+1 or −1) patterns to be stored in ECAM $\{X^i, i = 1, 2, \ldots, M\}$ are given. Let X be the N-bit bipolar current-state pattern and X' be next-state pattern in the dynamic mode. Then the evolution equation (motion equation) of the ECAM is defined as:

$$X' = \mathrm{sgn}\left(\sum_{i=1}^{M} X^i * b^{\langle X^i, X\rangle}\right) \quad (b > 1), \tag{1}$$

where sgn(u) is a signum function and takes 1 if $u \geqslant 0$, and −1 otherwise. Chieuh and Goodman [8] have demonstrated that ECAM is asymptotically stable in both synchronous and asynchronous update modes by defining a Liapunov energy function. They have also shown that the asymptotic storage capacity of the ECAM scales exponentially with the length of memory patterns theoretically. Moreover, a VLSI chip based on the ECAM model with few neurons was fabricated. The ECAM chip was shown to perform almost as well as the computer simulation of ECAM.

If the inner product $\langle X^i, X\rangle$ in (1) is regarded as similarity measurement between the two patterns X^i and X, then we can substitute negative Manhattan distance (L1 distance) measurement $d_{\mathrm{L1}}(X^i, X) = -\sum_{j=1}^{N} \left|X_j^i - X_j\right|$ for $\langle X^i, X\rangle$. Therefore, we construct Modified ECAM (MECAM) as formulated below:

$$X' = \operatorname{sgn}\left(\sum_{i=1}^{M} X^i * \gamma^{-\sum\limits_{k=1}^{N}\left|X_k^i - X_k\right|}\right) \quad (\gamma > 1). \tag{2}$$

Certainly, we also may utilize other similarity measurements, e.g. Hamming distance, Euclidean distance, Direction cosine measurements, and so on, to modify the evolution equation of ECAM.

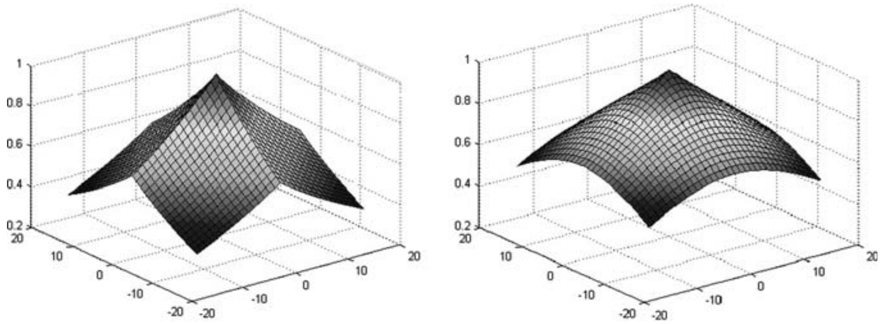
2.2. DISTANCE FACTOR

Neural networks are information processing systems made up of processing nodes and interconnections between nodes. The behaviors of a neural network depend on the structure of how the nodes are connected. Many previous works [17, 18] showed that good recognition generalization of neural networks can be obtained by designing a network architecture that contain amount of prior knowledge about the problem itself. Inspired by the thinking of "lateral feedback" mechanism in Self-Organizing Map, which shows that the synaptic weight between two neurons is proportional to the distance between them, we introduce two-Dimension decaying linear or nonlinear distance factors into ECAM and MECAM. Here, the value of the distance factor varies based on the spatial location of the neuron with respect to the neuron under consideration, and the effect of synaptic weights from farther neurons to the neuron under consideration is smaller.

Suppose the stored patterns are images with size $m \times n$ pixels, we arrange the neuron networks for two-Dimension array with m rows and n columns and define linear and nonlinear distance factors from the neuron in the ith row and jth column to the neuron in the kth row and lth column as follows:

(1) Nonlinear distance factor (NLDF):

$$d(i, j, k, l) = \eta^{-\omega\left(\left|\frac{i-k}{m}\right| + \left|\frac{j-l}{n}\right|\right)} \quad (\omega \geq 1, \eta > 1). \tag{3}$$



a) Nonlinear distance factor **b)** Linear distance factor

Figure 1. Distance functions. Note that the effect of the neighboring neurons on a particular neuron is reduced as the distance is increased.

(2) Linear distance factor (LDF):

$$d(i, j, k, l) = 1 - \frac{\beta}{\sqrt{2}}\sqrt{\left(\frac{i-k}{m}\right)^2 + \left(\frac{j-l}{n}\right)^2} \quad (0 \le \beta \le 1)\,. \tag{4}$$

Two-Dimension representations of Equations (3) and (4) are shown in Figure 1(a) and (b), respectively, where $m = n = 32$, $\omega = 1$, $\eta = e$, $\beta = 1$.

2.3. DISTANCE-BASED EXPONENTIAL CORRELATION ASSOCIATIVE MEMORY AND DISTANCE-BASED MODIFIED EXPONENTIAL CORRELATION ASSOCIATIVE MEMORY

Here, for convenience of description, we convert the above-mentioned distance factors in two-Dimension neurons array into corresponding distance factors in one-Dimension neuron vector array concatenated row by row. In other words, linear or nonlinear distance factor $D(i, j)$ between the ith neuron and the jth in one-Dimension neuron networks should be computed as (3) or (4), respectively. When, we incorporate a decaying distance matrix D into ECAM and MECAM, the evolution equations of DB-ECAM and DB-MECAM can, respectively, be formed as follows:

(1) DB-ECAM

$$X'_j = \mathrm{sgn}\left(\sum_{i=1}^{M} X_j^i * b^{\langle X^i .* D(:,j), X .* D(:,j)\rangle}\right) \quad (b > 1). \tag{5}$$

(2) DB-MECAM

$$X'_j = \mathrm{sgn}\left(\sum_{i=1}^{M} X_j^i * \gamma^{-\sum\limits_{k=1}^{N}\left|X_k^i * D(k,j) - X_k * D(k,j)\right|}\right) \quad (\gamma > 1), \tag{6}$$

where X'_j, X^i_j are the jth component of vector X', X^i, respectively. And $D(:, j)$ is the jth column of decaying distance matrix D.

3. Distance-Based Sparse Exponential Correlation Associative Memory and Distance-Based Sparse Modified Exponential Correlation Associative Memory

The conventional AMs usually adopted fully connected configuration. Though such neural networks exhibit, in general, a good performance, they are biologically unrealistic, as it is unlikely that natural evolution leads to such a large connectivity. In particularly, from a viewpoint of implementation perspective, realistic AMs must have sparse connectivity. So far, networks with sparse connectivity have been studied in detail [19–23], but they did only involve simple HAM and said nothing of performance comparison among the networks with sparse connectivity and full connection in the literatures. In this section, our goal is to borrow the thinking of SWA to build sparse neural networks, i.e. DBS-ECAM and DBS-MECAM, so that reduce their connection complexity greatly for ease of hardware implementation.

3.1. SPARSE NEURAL NETWORK ARCHITECTURE: SMALL WORLD ARCHITECTURE

The notion of small-world phenomenon was first introduced in a social experiment by Milgram [24]. Milgram showed that, despite the high amount of clustering in social networks (meaning that two acquaintances are likely to have other common acquaintances), any two individuals could be "linked" through a surprising small number of others. In fact, a lot of networks are shown to have a small-world topology. Examples include social networks such as acquaintance networks and collaboration networks, technological networks such as the Internet, the World Wide Web, and power grids, and biological networks such as neural networks, and metabolic networks, etc.

More recently, Watts and Strogatz [16] described a formulation for an analytical model of the small world. Starting with a periodic, locally connected lattice with n vertices and k edges per vertex (where $n \gg k \gg \ln(n) \gg 1$), edges are rewired with probability p to any random vertex in the network. For low p, this creates a small world network with primarily local connectivity and a few randomly placed long-range connections termed "short cuts" (see Figure 2).

3.2. DISTANCE-BASED SPARSE EXPONENTIAL CORRELATION ASSOCIATIVE MEMORY AND DISTANCE-BASED SPARSE MODIFIED EXPONENTIAL CORRELATION ASSOCIATIVE MEMORY

In DBS-ECAM and DBS-MECAM, neural interconnectivity is specified by a topology graph, G on the vertices $[N] \times [N]$ where a connection, or synapse, from neuron i to j exists if $\{i, j\} \in G$. Initially, the network is a one-dimensional lattice

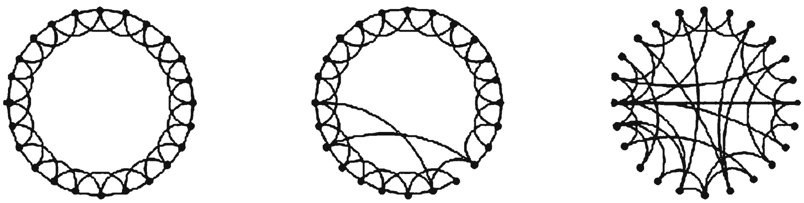


Figure 2. Network connection topologies.

with periodic boundary conditions (a ring), with each neuron feeding its output to its $k\,(=\alpha^* N)$ nearest–neighbors, where α is the overall connectivity of the network. Each vertex (neuron) is then visited, and with probability p, each edge is rewired to a randomly chosen vertex in the network, as described in [16].

Suppose matrix C denotes connectivity of the topology graph G. Let $C(i,j)$ be the element of the ith row and jth column in matrix C and defined as follows:

$$\mathrm{C}(i,j)=\begin{cases}0, & \text{if } \{i,j\}\notin G,\\ 1, & \text{if } \{i,j\}\in G.\end{cases} \tag{7}$$

Then the evolution equations of DBS-ECAM and DBS-MECAM with sparse configuration are shown as follows:

(1) DBS-ECAM

$$X'_j=\operatorname{sgn}\left(\sum_{i=1}^{M} X^i_j * b^{\langle X^i .* D(:,j) .* C(:,j),\, X .* D(:,j) .* C(:,j)\rangle}\right) \quad (b>1). \tag{8}$$

(2) DBS-MECAM

$$X'_j=\operatorname{sgn}\left(\sum_{i=1}^{M} X^i_j * \gamma^{-\sum_{k=1}^{N}\left|X^i_k * D(k,j) * C(k,j) - X_k * D(k,j) * C(k,j)\right|}\right) \quad (\gamma>1). \tag{9}$$

It is worth noting that, strictly speaking, the network configuration remains with SWA only when both α and p are small (usually $\alpha \le 0.25$, $p \le 0.5$). But in this paper, we only utilize the thinking of SWA to embody sparseness of neural networks and take α as 0.2 or 0.4, p as 0.67 or 0.75.

4. Computer Simulations

The performance of an AM is usually measured in terms of storage capacity and noise tolerance. Chiueh and Goodman had analyzed that the capacity of ECAM scales exponentially with N (the length of the stored vectors) in [8], and as its variants, we conjecture that DB-ECAM, DBS-ECAM, MECAM, DB-MECAM, and

DBS-MECAM also possess comparable storage capacity to ECAM. But in fact, to prove this claim theoretically is difficult, if not impossible. Thus, we turn to use an experimental method to test the capacity. However, for the above networks with exponential capacity, even using experiment seems also difficult because to directly test their real capacities needs an exponential-scale number of training samples or/and their corresponding slightly noisy ones, especially for large dimension sample. Instead, we choose the noise tolerance under the salt and pepper noise as an evaluation criterion of measuring their performance. Throughout this paper, the salt and pepper noise with noise density q means that, for a binary image of pixel size $m \times n$, $q \times (m \times n)$ pixels are first randomly selected, and then a half of these selected pixels are randomly set to black, and the other half are set to white.

In this section, we firstly used 62 binary images of pixel size 32×32, including 26 uppercases, 26 lowercases, and 10 Arabic numbers as shown in Figure 3(a). A set of testing patterns was generated by adding the salt and pepper noise to the images in the training set. Figure 3(b), (c) and (d) show the images corrupted with 20, 40, and 60% salt and pepper noise, respectively. Notice that at 60% noise level, the recognition problem becomes extremely difficult even for the human. In the network processing, each one of the 62 prototype images is first turned into a corresponding vector by the row by row concatenation way and black pixels are encoded by -1 while white pixels by $+1$. Thus, each pattern is represented by a 1024 dimensional bipolar vector.

When a noisy or incomplete pattern is input to the neural networks, if the output pattern matches entirely the original training exemplar, then we call a successful recognition. The high-recognition rate means that the neural network has high-storage capacity and robust error-correcting capability.

4.1. COMPARISON OF RECOGNITION PERFORMANCE ON IMAGES WITH SALT AND PEPPER NOISE

To illustrate that the new networks incorporating distance factor and SWA work better than ECAM and MECAM, we will verify this point by examining their recognition capability. Figure 4 shows the average recognition rate of 5 times independent repeated runs when characters and numbers are recognized, respectively, using ECAM, MECAM, DB-ECAM, DB-MECAM, DBS-ECAM, and DBS-MECAM (where $\omega = 1, \beta = 1, \gamma = e, \eta = e, b = 4, \alpha = 0.4, p = 0.75$, and NLDF was selected).

It can be observed that both DB-ECAM and DB-MECAM with distance factor achieve far higher recognition accuracy than both ECAM and MECAM, respectively, when various amount of Salt and Pepper noise was introduced into the original images. Moreover, it is very interesting that both DBS-ECAM and DBS-MECAM with only 40% connectivity hold comparable noise-tolerance capability to both DB-ECAM and DB-MECAM with full connectivity when salt and pepper noise added in the original images is lower than 75%. Even for 85 and 90% salt and pepper noise, the recognition effect of both DBS-ECAM and DBS-ME-

(a) Original training image

(b) Testing images with 20% salt & pepper noise

(c) Testing images with 40% salt & pepper noise

(d) Testing images with 60% salt & pepper noise

Figure 3. 32 × 32 dimensional character and number images set and the corresponding patterns corrupted with various amounts of salt and pepper noise.

CAM is also high enough and acceptable. More importantly, the sparse configuration will result in less computational and storage costs and contributes to VLSI implementation.

4.2. EFFECT OF LINEAR DISTANCE FACTOR AND NONLINEAR DISTANCE FACTOR ON THE NEW NEURAL NETWORKS

Here, the effect of LDF and NLDF on DBS-ECAM and DBS-MECAM, in which the parameters are taken as the same the ones as in Section 4.1, is investigated when 60–90% salt and pepper noises are, respectively, added to the original characters and the numbers images. Figure 5 describes the average recognition accuracy of 5 times experiments.

Simulation results show that LDF and NLDF almost yield the same effect on the new networks for all the noise levels lower than 75%. But for salt and pepper noise levels of 80% or above, the new networks incorporating LDF achieve slightly

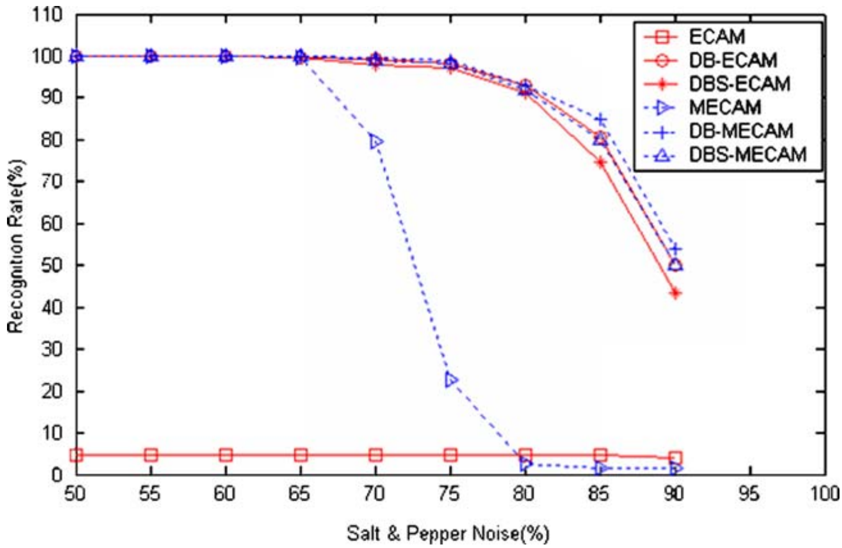


Figure 4. Recognition performance comparison of neural networks with different configurations.

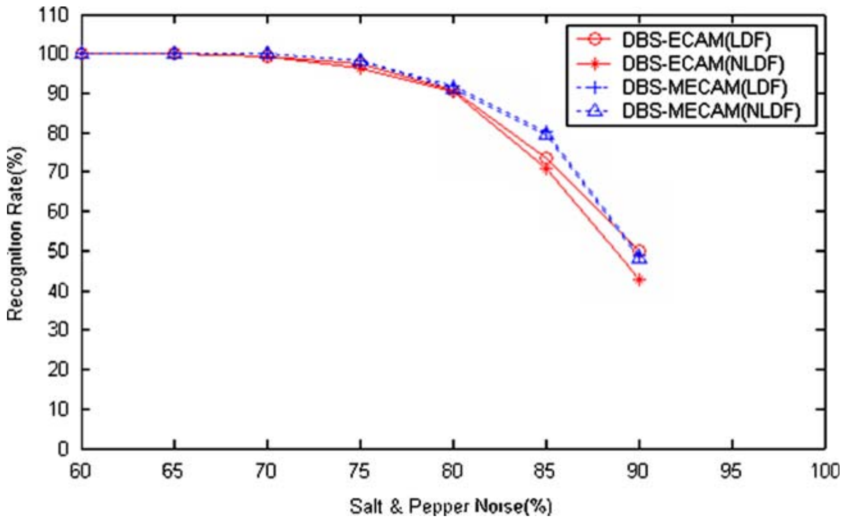


Figure 5. Recognition accuracy comparison of neural networks incorporating LDF or NLDF.

better noise-tolerance performance than the ones incorporating NLDF. Therefore, good recognition performance may be ensured by selecting an appropriate distance factor and adjusting the networks' arguments theoretically.

Hereafter, we only investigate the new MECAMs with various configurations due to ECAMs with the same configurations possess similar performance to the former.

4.3. INFLUENCE OF SPARSE CONNECTIVITY ON DISTANCE BASED SPARSE MODIFIED EXPONENTIAL CORRELATION ASSOCIATIVE MEMORY (NONLINEAR DISTANCE FACTOR)

Influence of sparse connectivity on the recognition performance of DBS-MECAM is explored in this sub-section. Simulation results of DBS-MECAM with 20 and 40% sparse connectivity are shown in Figure 6. From Figure 6, we can observe that DBS-MECAM with 20% sparse connectivity has also very high-recognition accuracy and its complexity is further reduced. In a real-world application, if some recognition accuracy is acceptable, then we usually select the network with simpler configuration because its hardware implementation is realistic and economical.

4.4. IMMUNE PERFORMANCE FROM CORRUPTED BLOCKS ON DISTANCE BASED SPARSE MODIFIED EXPONENTIAL CORRELATION ASSOCIATIVE MEMORY (NONLINEAR DISTANCE FACTOR)

In this subsection, we concentrates on examining influence of block noise on the restoring performance of the DBS-MECAM (NLDF, 40% sparse connectivity), aiming to investigate its immunity or insensitiveness to block noise as shown in Figure 7. It can be seen from the figure that the proposed network model not only is immune from corrupt blocks, what is more, but also converges and get stability in recall after only less than three iterations.

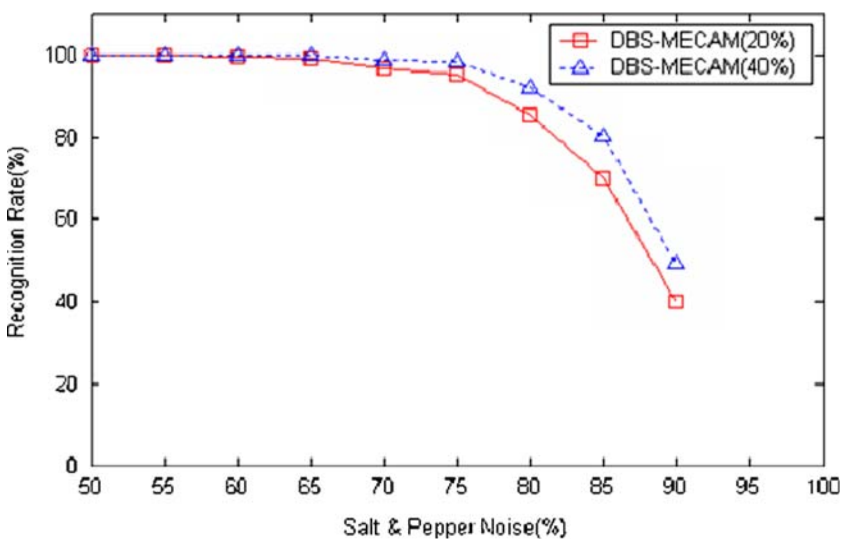


Figure 6. Recognition performance comparison of DBS-MECAM with 20 and 40% sparse connectivities.

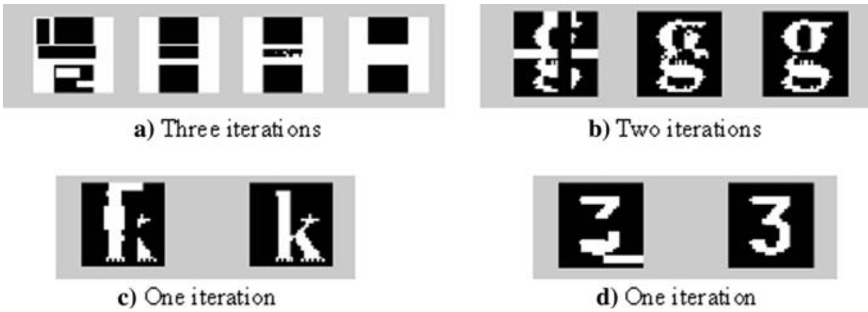


Figure 7. Restoration of characters and numbers from data with corrupt blocks.

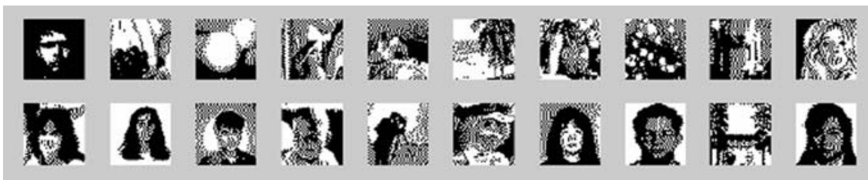


Figure 8. 50×50 dimensional binary visual images training set.

4.5. INFLUENCE OF NEURAL NETWORKS WITH DIFFERENT CONFIGURATIONS ON THE BINARY VISUAL IMAGES WITH HIGH RESOLUTION

Finally, Influence of neural networks with different configuration on visual images with high resolution is investigated. Here, we select randomly 50 binary images with 50×50 resolution as training set. Figure 8 shows some sample images from the training set. We carry out the experiments, respectively, for MECAM, DB-MECAM (NLDF), and DBS-MECAM (NLDF and 40% sparse connectivity) with the images added 10, 30, 50, 70, and 90% salt and pepper noises as their inputs respectively, the results further reveals that the higher the images' resolution is, the more obvious the advantage of distance factor and SWA are. Figure 9 shows their recall or recognition results on one of the images used. It can be observed that when the salt and pepper noise added is low, the three networks can all perfectly recall the original model; but with noise increasing to a certain extent, only the models incorporating NLDF can work well; especially for the testing images with higher noise level, DBS-ECAM exhibits the best noise-tolerance performance.

5. Conclusion and Future Work

In this paper, we propose sparse associative memory neural networks using a distance-based learning algorithm. The simulation results show that recognition or recall performances of the new neural networks are clearly superior to their corresponding fully connected ECAM and MECAM. The nature of the distance factor and sparse configuration contributes to remove a part of synaptic weights, which

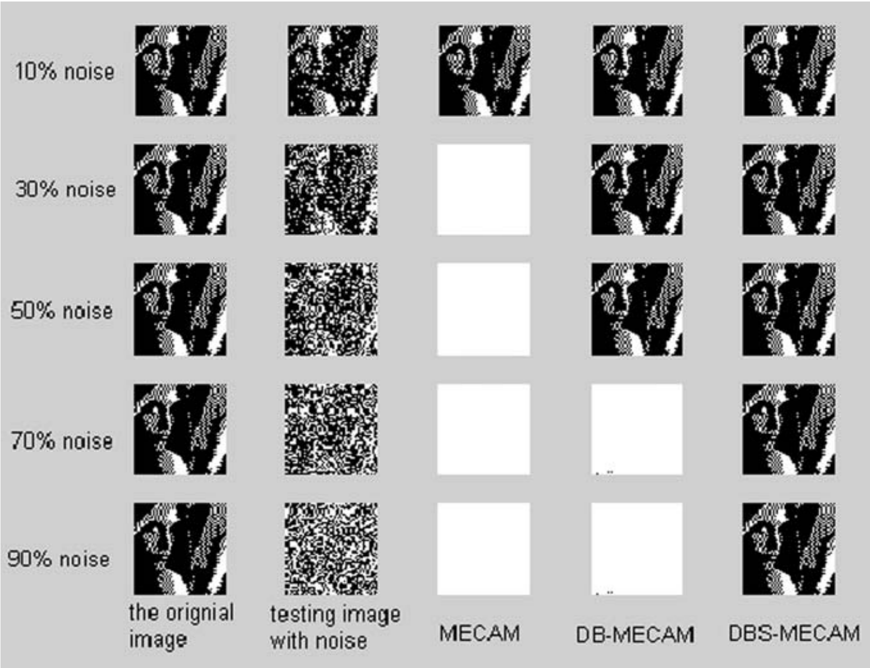


Figure 9. Simulation results of neural networks with different configurations when the pattern corrupted by various amounts of salt and pepper noise with high resolution was input.

leads to the reduction of complexity of the networks in terms of both software and hardware implementation.

In our future work, we will focus on extending the sparse configuration and distance factor to those multi-value associative memories and hetero-associative memories for recognizing more realistic gray level images and human face images.

Acknowledgments

We thank the referees for their insightful comments, and thank National Science Foundations of China under Grant No. 60271017, Jiangsu Natural Science Key Project (BK2004001), Jiangsu and NJUPT "QingLan" Project Foundation and the Returnee's Foundation of China Scholarship Council for partial supports, respectively.

References

1. Haykin, S.: *Neural Networks: A Comprehensive Foundation, 2nd ed*, Englewood Cliffs, NJ, Prentice-Hall, 1999.
2. Hopfield, J. J.: Neural network and physical systems with emergent collective computational abilities, Proceeding of the National Academic Science USA **79** (1982), 2554–2558.

3. McEliece, R. J., Posner, E. C., Rodemich, E. R. and Venkatesh, S. S.: The capacity of the Hopfield associative memory, *IEEE Transaction Information Theory* **IT-33** (1987), 461–482.
4. Psaltis, D. and Park, C. H.: Nonlinear discriminant functions and associative memories, In: J. S. Denker (ed), pp. 370–375, *Neural Networks for Computing*, American Institute of Physics, New York, NY, 1986.
5. Dembo, A. and Zeitouni, O.: High density associative memories, In: D. Z. Anderson (ed),*Neural Information Processing Systems*, pp. 211–218, American Institute of Physics, New York, NY, 1988.
6. Dembo, A. and Zeitouni, O.: General potential surfaces and neural networks, *Physics Review A* **37**(6) (1988), 2134–2143.
7. Sayeh, M. R. and Han, J. Y.: Pattern recognition using a neural network, In: *Proceeding SPIE Cambridge Symposium Optics and Optoelectronics Engineering* Cambridge, MA, 1987.
8. Chiueh, T. D. and Goodman, R. M.: Recurrent correlation associative memories, *IEEE Transactions on Neural Networks* **2**(2) (1991), 275–284.
9. Chiueh, T. D. and Tsai, H. K.: Multivalued associative memories based on recurrent networks, *IEEE Transactions on Neural Networks* **4**(2) (1993), 364–366.
10. Chen, Z. Y. and Kwong, C. P.: Recurrent correlation associative memories with multiple-value, In: *Proceedings of IEEE International Conference Neural Networks*, pp.1068–1073, 1994.
11. Kosko B.: Bidirectional associative memories, *IEEE Transactions on System, Man and Cybernetics* **18**(1) (1988), 49–60.
12. Tai, H. M.: High-order bidirectional associative memory, *Electronics Letters* **25** (1989), 1424–1425.
13. Jeng, Y. J., Yeh, C. C. and Chiueh, T. D.: Exponential bidirectional associative memories, *IEEE Transactions Neural Networks* **2** (1991), 275–284.
14. Chen, S. C., Gao, H. and Yan, W.: Improved exponential associative memory, *Electronics Letters* **33** (1997), 223–224.
15. Kohonen, T.: *Self-Organizing Maps*, 2nd edn, Springer-Verlag, Berlin, 1997.
16. Watts, D.J. and Strogatz, S.H.: Collective dynamics of 'small-world' networks, *Nature* **393** (1998), 440–442.
17. Moon, S. W. and Kong, S. G.: Block-based neural networks, *IEEE Transaction on Neural Networks* **12**(2) (2001), 307–317.
18. Seow, M. J. and Asari, V. K.: Learning using distance based training algorithm for pattern recognition, *Pattern Recognition Letters* **25**(2) (2004), 189–196.
19. Venkatesh, S.: Robustness in neural computation: random graphs and sparsity, *IEEE Transactions Information Theory* **38** (1992), 1114–1119.
20. McGraw, P. J. and Mezinger, M.: Topology and computational performance of attractor neural networks, *Physical Review E* **68** (2003), 047102.
21. Kim, B. J.: Performance of networks of artificial neurons: the role of clustering, *Physical Review E* **69** (2004), 045101.
22. Torres, J. J., Munoz, M. A., Marro, J. and Garrido, P. L.: Influence of topology on the performance of a neural network, *Neurocomputing* **58–60** (2004), 229–234.
23. Bohland, J. W. and Minai, A. A.: Efficient associative memory using small-world structure, *Neurocomputing* **38–40** (2001), 489–496.
24. Milgram, S.: The small-world problem, *Psychological Today* **2** (1967), 60–67.