Lattice Associative Memories That Are Robust In The Presence Of Noise

Gerhard X. Ritter
CISE Dept., Univ. of Florida
Gainesville, FL 32611-6120
Email: ritter@cise.ufl.edu

Gonzalo Urcid-Serrano
Optics Dept., INAOE
Puebla, Mexico
Email: gurcid@inaoep.mx

Mark S. Schmalz
CISE Dept., Univ. of Florida
Gainesville, FL 32611-6120
Email: mssz@cise.ufl.edu

***Abstract*— This paper presents a novel two-layer feedforward neural network that acts as an associative memory for pattern recall. The neurons of this network have dendritic structures and the computations performed by the network are based on lattice algebra. Use of lattice computation avoids multiplicative processes and, thus, provides for fast computation. The synaptic weights of the axonal fibers are preset, making lengthy training unnecessary. The proposed model exhibits perfect recall for perfect input vectors and is extremely robust in the presence of noisy or corrupted input.**

***Keywords*—artificial neural networks, dendritic computation, associative memories.**

I. Introduction

The concept of an associative memory is a fairly intuitive one: associative memory seems to be one of the primary functions of the brain. We easily *associate* the face of a friend with that friend's name, or a name with a telephone number. Artificial neural networks (ANNs) that are capable of storing several types of patterns and corresponding associations are referred to as *associative memories*.

These notions can be made more explicit by using mathematical formulation. Suppose that $X = \left\{\mathbf{x}^1, \ldots, \mathbf{x}^k\right\} \subset \mathbb{R}^n$ and $Y = \left\{\mathbf{y}^1, \ldots, \mathbf{y}^k\right\} \subset \mathbb{R}^m$, where $\mathbb{R}$ denotes the set of real numbers and $\mathbb{R}^k$ denotes the k-dimensional Euclidean vector space. A function $M : \mathbb{R}^n \to \mathbb{R}^m$ with the property that $M\left(\mathbf{x}^\xi\right) = \mathbf{y}^\xi \ \forall \xi = 1, \ldots, k$ is called an *associative memory* that associates pattern $\mathbf{x}^\xi$ with the pattern $\mathbf{y}^\xi$. If, for $\xi = 1, \ldots, k$, $\tilde{\mathbf{x}}^\xi$ denotes a distorted (noisy) version of $\mathbf{x}^\xi$, and $M\left(\tilde{\mathbf{x}}^\xi\right) = \mathbf{y}^\xi$, then M is said to be *robust in the presence of noise*. Finally, if $X = Y$, then M is called an *autoassociative memory*, and if $X \neq Y$, then M is called a *heteroassociative memory*.

Some of the earliest heteroassociative memories are the well-known correlation matrix memories established by T. Kohonen [1], [2]. These memories were linear associative memories. Many other popular models of associative memories also allow for a formulation using matrices [3], [4], [5], [6], [7], [8], [9]. Among the various autoassociative memories, the Hopfield network is the most widely known today [5], [10], [11]. A large number of researchers have exhaustively studied this network, its variations, and generalizations [6], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22].

In comparison to the early correlation matrix memories, the memories based on Hopfield's early work are nonlinear due to the application of a nonlinear activation function. However, some major problems concerning these memories are the limited storage capacity [20], [23], [24] and convergence problems. Unlike the Hopfield network, which is a recurrent network with the underlying mathematics based on the ring of real numbers $(\mathbb{R}, +, \times)$, the morphological model formulated by Ritter et al. [8], [9], [25] is based on the semiring $(\mathbb{R}_{-\infty}, \vee, +)$ or, equivalently, $(\mathbb{R}_{\infty}, \wedge, +)$. This model has no restriction on the number of stored patterns and converges in one pass through the network. In spite of these advantages, the morphological model also exhibits some bothersome deficiencies. For morphological autoassociative memories, perfect recall for perfect (non-corrupted) input is always guaranteed. Although these memories can be extremely robust in the presence of noise, perfect recall is not guaranteed. Even minute distortions in just a few pattern vector components of an input pattern may prevent perfect recall [25]. Morphological heteroassociative memories share this problem when faced with noisy input. Additionally, perfect recall association for perfect input patterns is not guaranteed unless certain specific conditions are satisfied by the prototype association patterns [8].

In this paper we present a novel heteroassociative memory that inherits the storage and computational advantages of the morphological associative memories but not their deficiencies mentioned above. The memory is grounded on previous research that employs neurons with dendritic structures [26], [27]. Since autoassociative memories are special cases of associative memories, the proposed model can be used as either a heteroassociative or autoassociative memory.

II. The Dendritic Model

The artificial neural model that employs dendritic computation has been motivated by the fact that several researchers have proposed that dendrites, and not the neurons, are the elementary computing devices of the brain, capable of implementing logical functions such as AND, OR, and NOT [28], [29], [30], [31], [32]. In the mammalian

Mathematical Methods in Pattern and Image Analysis, edited by
Jaakko T. Astola, Ioan Tabus, Junior Barrera, Proceedings of SPIE
Vol. 5916 (SPIE, Bellingham, WA, 2005) 0277-786X/05/$15 · doi: 10.1117/12.622589

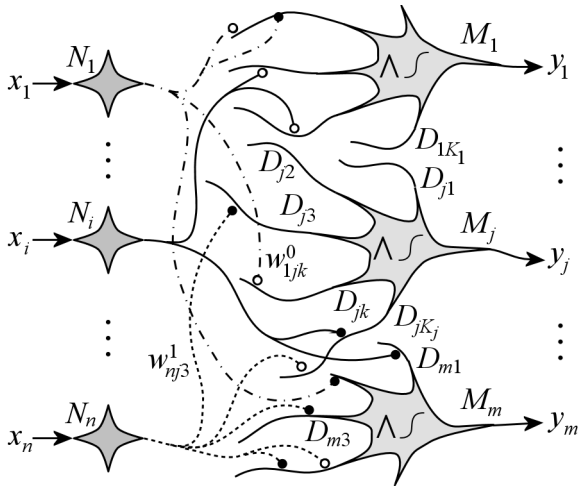

Fig. 1. Single-layer morphological perceptron with dendritic structures. The symbol D_{jk} denotes the kth dendrite of neuron M_j and K_j the number of dendrites of M_j. An axonal fiber of an input neuron can synapse on a dendrite with excitatory and/or inhibitory fibers. For instance, w^1_{nj3} is the weight of an excitatory fiber from neuron N_n to dendrite D_{j3} of neuron M_j, while w^0_{1jk} is the weight of an inhibitory fiber coming from neuron N_1 to dendrite D_{jk} of neuron M_j.

brain, dendrites span all cortical layers and account for the largest component in both surface area and volume. Thus, dendrites cannot be omitted when attempting to build artificial neural models.

Inspired by the neurons of the biological brain, we developed a model of *morphological neuron* that possesses *dendritic structures*. A number of such neurons can then be arranged on one layer, similarly to the classical single-layer perceptron (SLP), in order to build a single-layer morphological perceptron with dendritic structures (SLMP). This artificial model is described in detail in [26] and only briefly summarized below due to page limitation. The network is illustrated in Fig. 1.

Let $N_1, \ldots, N_n$ denote a set of input neurons, which provide synaptic input to the main layer of neurons with dendritic structures $M_1, \ldots, M_m$, which is also the output layer. The value of an input neuron N_i ($i = 1, \ldots, n$) propagates through its axonal tree to the terminal branches that make contact with the dendrites of the output neuron M_j ($j = 1, \ldots, m$). Each M_j has a finite number of dendrites D_{jk}, where $k = 1, \ldots, K_j$ and K_j denotes the total number of dendrites of M_j. The weight of an axonal branch of neuron N_i terminating on the kth dendrite of M_j, namely D_{jk}, is denoted by w^ℓ_{ijk}, where the superscript $\ell \in \{0, 1\}$ distinguishes between *excitatory* ($\ell = 1$) and *inhibitory* ($\ell = 0$) input to the dendrite. The kth dendrite, D_{jk}, of M_j will respond to the total input received from the neurons $N_1, \ldots, N_n$ and will either accept or inhibit the received input. The computation of the kth dendrite of M_j is given by

$$\tau^j_k(\mathbf{x}) = p_{jk} \bigwedge_{i \in I(k)} \bigwedge_{\ell \in L(i)} (-1)^{1-\ell} \left(x_i + w^\ell_{ijk}\right) ,$$

where $\mathbf{x} = (x_1, \ldots, x_n)'$ denotes the input value of the neurons $N_1, \ldots, N_n$ with x_i representing the value of N_i; $I(k) \subseteq \{1, \ldots, n\}$ corresponds to the set of all input neurons with terminal fibers that synapse on the kth dendrite of M_j; $L(i) \subseteq \{0, 1\}$ corresponds to the set of terminal fibers of N_i that synapse on the kth dendrite of M_j; and $p_{jk} \in \{-1, 1\}$ denotes the excitatory ($p_{jk} = 1$) or inhibitory ($p_{jk} = -1$) response of the kth dendrite of M_j to the received input.

It follows from the formulation $L(i) \subseteq \{0, 1\}$ that the ith neuron N_i can have at most two synapses on a given dendrite k [26]. Also, if the value $\ell = 1$, then the input $\left(x_i + w^1_{ijk}\right)$ is excitatory, and inhibitory for $\ell = 0$ since in this case we have $-\left(x_i + w^0_{ijk}\right)$.

The value $\tau^j_k(\mathbf{x})$ is passed to the cell body and the state of M_j is a function of the input received from all its dendrites. The total value received by M_j is given by

$$\tau^j(\mathbf{x}) = p_j \bigwedge_{k=1}^{K_j} \tau^j_k(\mathbf{x}) ,$$

where K_j denotes the total number of dendrites of M_j and $p_j = \pm 1$ denotes the response of the cell body to the received dendritic input. Here again, $p_j = 1$ means that the input is accepted, whereas $p_j = -1$ means that the cell rejects the received input. The *next* state of M_j is then determined by an activation function f, namely $y_j = f\left[\tau^j(\mathbf{x})\right]$. Typical activation functions used with the dendritic model include the hard-limiter and the pure linear identity function. The single-layer morphological perceptron usually employs the former.

For a more thorough understanding of this model as well as its computational performance, we refer the reader to examples and theorems given in [26], [27].

III. An Associative Memory based on the Dendritic Model

In this section we describe a new associative memory that is based on the dendritic model discussed in the previous section. This new memory can store the sets $X = \left\{\mathbf{x}^1, \ldots, \mathbf{x}^k\right\} \subset \mathbb{R}^n$ and $Y = \left\{\mathbf{y}^1, \ldots, \mathbf{y}^k\right\} \subset \mathbb{R}^m$ of associate patterns and has the ability to cope with random noise. The memory we are about to describe consists of the following: n input neurons $N_1, \ldots, N_n$; k neurons in the hidden layer, which we denote by $H_1, \ldots, H_k$; and m output neurons $M_1, \ldots, M_m$. Let

$$\mathrm{d}\left(\mathbf{x}^\xi, \mathbf{x}^\gamma\right) = \max\left\{\left|x^\xi_i - x^\gamma_i\right| : i = 1, \ldots, n\right\}$$

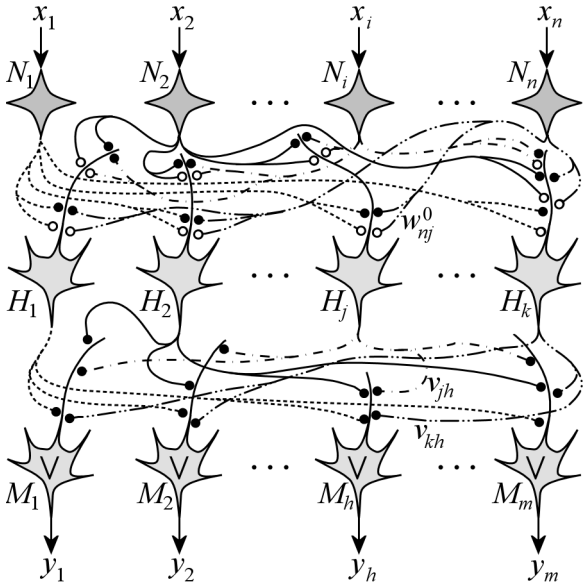

Fig. 2. The topology of the morphological associative memory based on the dendritic model. The network is fully connected; all axonal branches from input neurons synapse via two fibers on all hidden neurons, which in turn connect to all output nodes via excitatory fibers.

represent the Chebyshev (checkerboard) distance between the patterns $\mathbf{x}^\xi$ and $\mathbf{x}^\gamma$, and $d_{\min}$ denote the minimum Chebyshev distance between all pairs of patterns,

$$d_{\min} = \min\left\{\mathrm{d}\left(\mathbf{x}^\xi, \mathbf{x}^\gamma\right) : \xi < \gamma,\, \xi, \gamma \in \{1, \ldots, k\}\right\}\ ,$$

based on which we choose an allowable noise parameter α satisfying the inequality

$$\alpha < \frac{1}{2} d_{\min}\,. \tag{1}$$

For $\mathbf{x} \in \mathbb{R}^n$, the input for the neuron N_i will be the ith coordinate of $\mathbf{x}$. Each neuron H_j in the hidden layer has exactly one dendrite. Every input neuron N_i has exactly one pair of terminal axonal fibers terminating on the dendrite of H_j. The weights of the terminal fibers of N_i terminating on the dendrite of H_j are given by

$$w^\ell_{ij} = \begin{cases} -\left(x^j_i - \alpha\right) & \text{if } \ \ell = 1\,, \\ -\left(x^j_i + \alpha\right) & \text{if } \ \ell = 0\,, \end{cases} \tag{2}$$

where $i = 1, \ldots, n$ and $j = 1, \ldots, k$. For a given input $\mathbf{x} \in \mathbb{R}^n$, the dendrite of the hidden unit H_j computes

$$\tau^j(\mathbf{x}) = \bigwedge_{i=1}^{n} \bigwedge_{\ell=0}^{1} (-1)^{1-\ell} \left(x_i + w^\ell_{ij}\right). \tag{3}$$

The state of the neuron H_j is determined by the hard-limiter activation function

$$f(z) = \begin{cases} 0 & \text{if } \ z \geq 0\,, \\ -\infty & \text{if } \ z < 0\,. \end{cases} \tag{4}$$

Thus, the output of H_j is given by $f\left[\tau^j(\mathbf{x})\right]$ and is passed along its axonal fibers to the output neurons $M_1, \ldots, M_m$. The activation function defined by Eq. (4) is a hard-limiter in the algebra $\mathcal{A} = (\mathbb{R}_{-\infty}, \vee, +)$ since the *zero* of $\mathcal{A}$ is $-\infty$ (for the operation $\vee$) and the *unit* of $\mathcal{A}$ corresponds to 0. This mirrors the hard-limiter in the algebra $(\mathbb{R}, +, \times)$ $f(z) = 0$ if $z < 0$ and $f(z) = 1$ if $z \geq 0$, since in this algebra the zero is 0 and the unit is 1.

Similar to the hidden layer neurons, each output neuron M_h, where $h = 1, \ldots, m$, has one dendrite. However, each hidden neuron H_j has exactly one excitatory axonal fiber and no inhibitory fibers terminating on the dendrite of M_h. Figure 2 illustrates this dendritic network model. The excitatory fiber of H_j terminating on M_h has synaptic weight $v_{jh} = y^j_h$. The computation performed by M_h is given by

$$\tau^h(\mathbf{s}) = \bigvee_{j=1}^{k} \left(s_j + v_{jh}\right), \tag{5}$$

where s_j denotes the output of H_j, namely $s_j = f\left[\tau^j(\mathbf{x})\right]$, with f defined in Eq. (4). The activation function g for each output neuron M_h is the linear identity function $g(z) = z$.

Each neuron H_j will have the output value $s_j = 0$ if and only if $\mathbf{x}$ is an element of the hypercube

$$B^j = \left\{\mathbf{x} \in \mathbb{R}^n : x^j_i - \alpha \leq x_i \leq x^j_i + \alpha\right\} \tag{6}$$

and $s_j = -\infty$ whenever $\mathbf{x} \in \mathbb{R}^n \setminus B^j$. Thus, the output of this network will be $\mathbf{y} = (y_1, \ldots, y_m)' = \left(y^j_1, \ldots, y^j_m\right)' = \mathbf{y}^j$ if and only if $\mathbf{x} \in B^j$. That is, whenever $\mathbf{x}$ is a corrupted version of $\mathbf{x}^j$ with each coordinate of $\mathbf{x}$ not exceeding the allowable noise level α, then $\mathbf{x}$ will be associated with $\mathbf{y}^j$. If the amount of noise exceeds the level α, then the network rejects the input by yielding the output vector $(-\infty, \ldots, -\infty)'$. Obviously, each uncorrupted pattern $\mathbf{x}^\xi$ will be associated with $\mathbf{y}^\xi$.

Experiment

To illustrate the performance of this associative memory, we use a visual example consisting of the associated image pairs $P = \{\mathbf{p}^1, \mathbf{p}^2, \mathbf{p}^3\}$ and $Q = \{\mathbf{q}^1, \mathbf{q}^2, \mathbf{q}^3\}$ shown in Fig. 3. Each $\mathbf{p}^\xi$ is a 50×50-pixel 256-gray scale image, whereas each $\mathbf{q}^\xi$ is a 30×50-pixel 256-gray scale image, where $\xi = 1, 2, 3$. Using the standard row-scan method, each pattern image $\mathbf{p}^\xi$ and $\mathbf{q}^\xi$ was converted into a pair of pattern vectors $\mathbf{x}^\xi = \left(x^\xi_1, \ldots, x^\xi_{2500}\right)'$ and $\mathbf{y}^\xi = \left(y^\xi_1, \ldots, y^\xi_{1500}\right)'$ by defining $x^\xi_{50(r-1)+c} = p^\xi(r, c)$ for $r, c = 1, \ldots, 50$ and, respectively, $y^\xi_{50(r-1)+c} = q^\xi(r, c)$ for $r = 1, \ldots, 30$ and $c = 1, \ldots, 50$. Thus, $X = \left\{\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3\right\} \subset \mathbb{R}^{2500}$ and $Y = \left\{\mathbf{y}^1, \mathbf{y}^2, \mathbf{y}^3\right\} \subset \mathbb{R}^{1500}$.

The images were then distorted by randomly corrupting 95% of the coordinates of each $\mathbf{x}^\xi$ within a noise level α

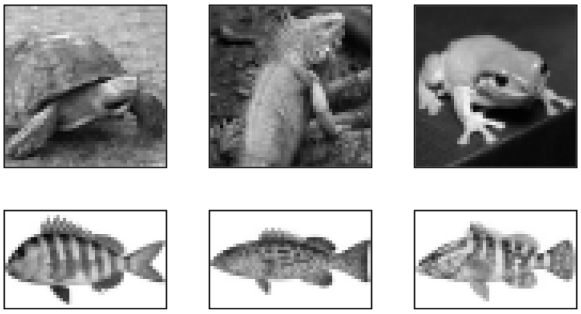

Fig. 3. The *top row* depicts the images $\mathbf{p}^1, \mathbf{p}^2, \mathbf{p}^3$, which were converted into the prototype patterns of the set $X = \{\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3\}$, while the *bottom row* shows the images used to generate the corresponding association patterns from the set $Y = \{\mathbf{y}^1, \mathbf{y}^2, \mathbf{y}^3\}$.

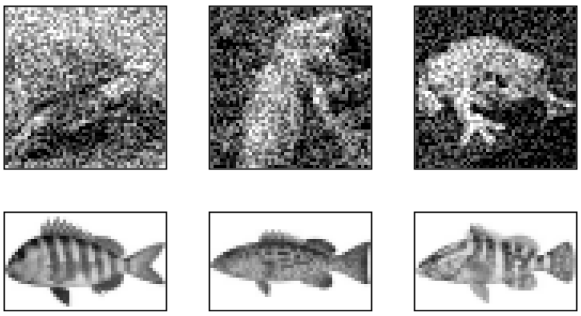

Fig. 4. The *top row* shows the prototype patterns $\tilde{\mathbf{x}}^1, \tilde{\mathbf{x}}^2, \tilde{\mathbf{x}}^3$ corrupted with random noise within the range $[-\alpha, \alpha]$ and the *bottom row* shows the perfect recall association achieved by the proposed model.

chosen to satisfy the inequality in (1). In this particular example, we set $\alpha = \frac{2}{5} d_{\min}$. Numerically, we have $\mathrm{d}(\mathbf{x}^1, \mathbf{x}^2) = 198$, $\mathrm{d}(\mathbf{x}^1, \mathbf{x}^3) = 211$, and $\mathrm{d}(\mathbf{x}^2, \mathbf{x}^3) = 188$, where the range of pixel values (vector features) is $[0, 255]$. Hence, $d_{\min} = \mathrm{d}(\mathbf{x}^2, \mathbf{x}^3) = 188$ and $\alpha = 75.2$. The top row of Fig. 4 shows the images thus corrupted while the bottom row shows the perfect recall association achieved by the network.

In this experiment, recall succeeds because the amount of distortion is controlled not to exceed the allowable level α. When the noise is above α, the memory described thus far fails to recognize the patterns and rejects them by outputting a vector $(-\infty, \ldots, -\infty)'$. Failure of recall is illustrated in Fig. 5, where the output vector components of $-\infty$ are conventionally depicted as white pixels.

IV. Increase of Noise Tolerance

The model can be refined to be more tolerant to random noise as follows. Instead of choosing a single parameter α for all prototype patterns in the set X, we set one allowable

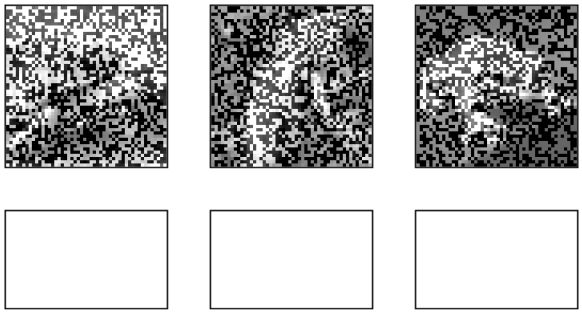

Fig. 5. The *top row* shows the prototype patterns $\tilde{\mathbf{x}}^1, \tilde{\mathbf{x}}^2, \tilde{\mathbf{x}}^3$ corrupted with random noise that exceeds the level $\alpha = 75.2$. The memory rejects these patterns and outputs vectors $(-\infty, \ldots, -\infty)'$ depicted in white pixels in the *bottom row*.

noise parameter $\alpha_\xi > 0$ for each $\xi \in \{1, \ldots, k\}$, satisfying

$$\alpha_\xi < \frac{1}{2} \min \left\{ \mathrm{d}\big(\mathbf{x}^\xi, \mathbf{x}^\gamma\big) : \gamma \in K(\xi) \right\} , \tag{7}$$

where $K(\xi) = \{1, \ldots, k\} \backslash \{\xi\}$. The model is similar to the one described above, with the exception that the weights of the axonal fibers of N_i terminating on the dendrite of H_j are now given by

$$w_{ij}^\ell = \begin{cases} -\big(x_i^j - \alpha_j\big) & \text{if } \ell = 1 , \\ -\big(x_i^j + \alpha_j\big) & \text{if } \ell = 0 , \end{cases} \tag{8}$$

where $i = 1, \ldots, n$ and $j = 1, \ldots, k$, instead of being set as previously in Eq. (2). For a given input $\mathbf{x} \in \mathbb{R}^n$, computation at the hidden and output layers proceeds as before according to Eqs. (3)–(5).

This time, each output neuron H_j will have a value $s_j = 0$ if and only if $\mathbf{x}$ is an element of the hypercube

$$B^j = \left\{ \mathbf{x} \in \mathbb{R}^n : x_i^j - \alpha_j \le x_i \le x_i^j + \alpha_j \right\} \tag{9}$$

and $s_j = -\infty$ whenever $\mathbf{x} \in \mathbb{R}^n \setminus B^j$. In contrast to Eq. (6), the hypercubes constructed at the hidden neurons are now of different sizes. All these hypercubes are at least as large as the one in Eq. (6) because the parameters α_ξ can be chosen greater than α and still satisfy the inequality in (7).

The output of the network having the weights set according to Eq. (8) will be $\mathbf{y} = (y_1, \ldots, y_m)' = \big(y_1^j, \ldots, y_m^j\big)' = \mathbf{y}^j$ if and only if $\mathbf{x} \in B^j$ with B^j as in Eq. (9). That is, whenever $\mathbf{x}$ is a corrupted version of $\mathbf{x}^j$ with each coordinate of $\mathbf{x}$ not exceeding the noise level α_j, then $\mathbf{x}$ will be associated with $\mathbf{y}^j$. As before, each uncorrupted pattern $\mathbf{x}^\xi$ will be associated with $\mathbf{y}^\xi$.

Experiment

The patterns illustrated in Fig. 5 were obtained by distorting 100% of the vector components of each $\mathbf{x}^\xi$ within a noise level α_ξ, chosen to

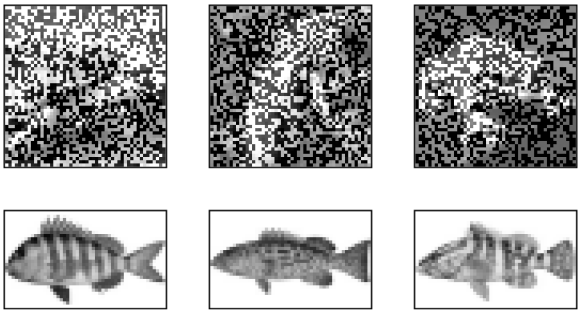

Fig. 6. The *top row* shows the same prototype patterns $\tilde{\mathbf{x}}^1, \tilde{\mathbf{x}}^2, \tilde{\mathbf{x}}^3$ from Fig. 5, which are correctly recognized by the refined model (*bottom row*).

satisfy the inequality in (7). Specifically, we set $\alpha_1 = 0.49 \min\left(\mathrm{d}\left(\mathbf{x}^1, \mathbf{x}^2\right), \mathrm{d}\left(\mathbf{x}^1, \mathbf{x}^3\right)\right) \approx 97$, $\alpha_2 = 0.49 \min\left(\mathrm{d}\left(\mathbf{x}^1, \mathbf{x}^2\right), \mathrm{d}\left(\mathbf{x}^2, \mathbf{x}^3\right)\right) \approx 92$, $\alpha_3 = 0.49 \min\left(\mathrm{d}\left(\mathbf{x}^1, \mathbf{x}^3\right), \mathrm{d}\left(\mathbf{x}^2, \mathbf{x}^3\right)\right) \approx 92$. The values were obtained by truncation to the nearest integer. Figure 6 shows that the refined model achieves perfect recall association, whereas the previous model does not, as seen in Fig. 5.

V. Discussion and Conclusion

In comparison to the Hopfield associative memory and the Kohonen matrix-based correlation memory and their various generalizations, the storage capacity of the lattice-based dendritic memory described here is unlimited; for any set of k associations, only k hidden neurons are required. Furthermore, dynamical convergence problems are nonexistent and lengthy training algorithms are absent. The memory is *hard-wired* via predetermined weights and the network can *learn* new pattern associations by simply adding one neuron for each new association in the hidden layer and grow additional axonal fibers using Eqs. (2) and (5) (or (8) and (5)). Thus far, we have been unable to find a memory that is superior in terms of storage capacity, speed, and recall capability in the presence of noise. The example provided in this paper hints at the extreme robustness of the lattice-based memory in the presence of noise. Testing the example presented here using other recently-proposed models [33], [34], [35], [36] we found that these other models either exhibited complete failure or much poorer performance. Of course, one example does not yield a statistical performance measure. However, our model is extremely simple and easy to analyze. Its recall capability is always perfect when the corrupted patterns remain within the boundaries given by Eq. (6) (or (9)). Additionally, several of the current memories we investigated operate only within the Boolean or bipolar domain. Our model has no such limitations.

References

[1] T. Kohonen, "Correlation matrix memory," IEEE Trans. Comp. **C-21**, 353–359 (1972).

[2] T. Kohonen and M. Ruohonen, "Representation of associative data by computers," IEEE Trans. Comp. **C-22**, 701–702 (1973).

[3] J. S. Albus, "A new approach to manipulator control: The cerebellar model articulation controller," J. Dynamic Systems and Measure Control, Trans. of the ASME **97**, 220–227 (1975).

[4] J. A. Austin Adam, "A distributed associative memory for scene analysis," Proc. IEEE 1st Internat. Conf. Neural Networks **4**, 285–286 (San Diego, CA, 1987).

[5] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," Proc. Nat. Acad. Sci. **79**, 2554–2558 (1982).

[6] B. Kosko, "Adaptive bidirectional associative memories," IEEE Trans. Systems, Man and Cybernetics, 124–136 (1987).

[7] K. Okajima, S. Tanaka, and S. Fujiwara, "A heteroassociative memory with feedback connection," Proc. IEEE 1st Internat. Conf. Neural Networks **II**, 711–718 (1987).

[8] G. X. Ritter, P. Sussner, and J. L. Diaz de Leon, "Morphological associative memories," IEEE Trans. Neural Networks **9**(2), 281–293 (1998).

[9] G. X. Ritter, J. L. Diaz de Leon, and P. Sussner, "Morphological bidirectional associative memories," Neural Networks **12**(6), 851–867 (1999).

[10] J. J. Hopfield, "Neurons with graded response have collective computational properties like those of two state neurons," Proc. Nat. Acad. Sci. **81**, 3088–3092 (1984).

[11] J. J. Hopfield and D. W. Tank, "Computing with neural circuits," Science **233**, 625–633 (1986).

[12] Y. Abu-Mostafa and J. St. Jacques, "Information capacity of the Hopfield model," IEEE Trans. Information Theory **7**, 1–11 (1985).

[13] D. Amit, H. Gutfreund, and H. Sompolinsky, "Storing infinite number of patterns in a spin-glass model neural networks," Physics Review Letters **55**(14), 1530–1533 (1985).

[14] H. Chen et al., "Higher order correlation model for associative memories," in *Neural Networks for Computing*, Ed. by J. S. Denker (AIP Proc. **151**, 1986).

[15] J. S. Denker, "Neural network models of learning and adaption," Physica **22**D, 216–222 (1986).

[16] V. Gimenez-Martinez, "A modified Hopfield auto-associative memory with improved capacity," IEEE Trans. Neural Networks **11**(4), 867–878 (2000).

[17] J. D. Keeler, "Basins of attraction of neural network models," in *Neural Networks for Computing*, Ed. by J. S. Denker (AIP Proc. **151**, 1986).

[18] B. Kosko, "Adaptive bidirectional associative memories," IEEE 16th Workshop Applied Images Pattern Recognition, 1–49 (1987).

[19] W.-J. Li, T. Lee, "Hopfield neural networks for affine invariant matching," IEEE Trans. Neural Networks **12**(6), 1400–1410 (2001).

[20] R. McEliece et al., "The capacity of Hopfield associative memory," Trans. Information Theory **1**, 33–45 (1987).

[21] T. Munehisa, M. Kobayashi, and H. Yamazaki, "Cooperative updating in the Hopfield model," IEEE Trans. Neural Networks **12**(5), 1243–1251 (2001).

[22] *Neural network study* (AFCEA Internat. Press, Fairfax, VA, 1988).

[23] G. Palm, "On associative memory," Biological Cybernetics **36**, 19–31 (1980).

[24] P. Sussner, "Observations on morphological associative memories and the kernel method," Elsevier Neurocomputing **31**, 167–183 (2000).

[25] G. X. Ritter, G. Urcid, and L. Iancu, "Reconstruction of noisy patterns using morphological associative memories," J. Math. Imaging and Vision **19**(2), 95–111 (2003).

[26] G. X. Ritter and G. Urcid, "Lattice algebra approach to single neuron computation," IEEE Trans. Neural Networks **14**(2), 282–295 (2003).

[27] G. X. Ritter and L. Iancu, "Lattice algebra approach to neural networks and pattern classification," Pattern Recognition Image Analysis **14**(2), 190–197 (2004).
[28] J. C. Eccles, *The Understanding of the Brain* (McGraw-Hill, New York, NY, 1977).
[29] C. Koch and I. Segev, Eds., *Methods in Neuronal Modeling: From Synapses to Networks* (MIT Press, Boston, MA, 1989).
[30] T. McKenna, J. Davis, and S.E. Zornetzer, Eds., *Single Neuron Computation* (Academic Press, San Diego, CA, 1992).
[31] B. W. Mel, "Synaptic integration in excitable dendritic trees," J. Neurophysiology **70**, 1086–1101 (1993).
[32] W. Rall and I. Segev, "Functional possibilities for synapses on dendrites and dendritic spines," in *Synaptic Function*, Ed. by G. M. Edelman, E. E. Gall, and W. M. Cowan, 605–636 (Wiley, New York, NY, 1987).
[33] D. Nowicki and O. Dekhtyarenko, "Kernel-based associative memory," Proc. IEEE Internat. Joint Conf. Neural Networks **1**, 741–744 (2004).
[34] D. Shen and J.B. Cruz, "Encoding strategy for maximum noise tolerance bidirectional associative memory," IEEE Trans. Neural Networks **16**(2), 293–300 (2005).
[35] C. S. Leung, "Optimum learning for bidirectional associative memory in the sense of capacity," IEEE Trans. Systems, Man and Cybernetics **24**(5), 791–796 (1994).
[36] B. Caputo, "Storage capacity of kernel associative memories," Proc. Internat. Conf. Artificial Neural Networks, 211–217 (Madrid, Spain, 2002).