Redundant Encoding of Patterns in Lattice Associative Memories

Gonzalo Urcid[1,⋆], Gerhard X. Ritter[2], and José-Angel Nieves-V.[1]

[1] Optics Department, INAOE, Tonantzintla, Pue. 72000, Mexico
{gurcid,janieves}@inaoep.mx
[2] CISE Department, University of Florida,
Gainesville, FL 32611–6120, USA
ritter@cise.ufl.edu

Abstract. Lattice matrix auto-associative memories also known as auto-associative morphological memories are artificial neural networks used to store and recall a finite set of binary or real valued patterns. They differ from other auto-associative memory models in the way exemplar patterns are encoded in the network as well as in the computation performed to recall a pattern. Both storage and recall mechanisms are based on minimax algebra operations that result in unique memory properties, such as, single step recall, perfect retrieval of all exemplar patterns, and infinite storage capacity. Two dual lattice matrix auto-associative memories have been developed so far. The min-memory is robust to erosive noise and the max-memory is robust to dilative noise; however, neither of these memories is able to recall patterns degraded by mixed or random noise. This paper introduces a redundant encoding of patterns based on the geometrical characterization of the set of fixed points common to both memories. Redundancy changes the size and shape of attraction basins of exemplar patterns and expands the set of fixed points, hence recall capability of patterns corrupted with random noise is possible using a simple scheme based on this type of memory networks.

Key words: Fixed point sets, Lattice associative memories, Minimax algebra, Morphological associative memories, Noisy patterns, Pattern recall, Redundant encoding

1 Introduction

Lattice matrix auto-associative memories approach the problem of pattern association from a minimax point of view. The Hebbian law of correlation encoding is still used to store a set of patterns but modified accordingly to be consistent with the mathematical framework of minimax algebra. Pattern retrieval is performed with minimax matrix products in a similar way as usual matrix multiplication of linear algebra is used in linear correlation associative memory models. Brief but complete summaries of the different developments in the field

⋆ Corresponding author. Fax: +52 (222) 247-2940; Tel.: +52 (222) 266-3100 Ext.8205

© Vassilis Kaburlasos, Uta Priss, Manuel Graña (Eds.): LBM 2008 (CLA 2008), pp. 45–57, ISBN 978–80–244–2112–4, Palacký University, Olomouc, 2008.

of morphological associative memories can be consulted in [6, 10, 17, 18, 23]. More recently, the canonical *lattice auto-associative memories* (LAMs), also known as *auto-associative morphological memories*(AMMs), have been treated within the more general framework of lattice transforms. In this framework, the min-$\mathbf{W}_{XX}$ and max-$\mathbf{M}_{XX}$ auto-associative memories are examples of lattice transforms between n-dimensional real vectors. The set of fixed points of a lattice transform is a key concept since it provides a complete description of the transform behavior. Fixed points of AMMs where first presented in [15] and discussed further in [10, 18]; one of the main results is that $\mathbf{W}_{XX}$ and $\mathbf{M}_{XX}$, have the same set of fixed points denoted by $F(X)$. Recently, an algebraic as well as a geometrical characterization of the set of fixed points $F(X)$ has been established in [10, 12]. This work proposes to exploit the infinite storage capacity and to take advantage of the geometrical nature of $F(X)$, to enable recall capability of lattice matrix auto-associative memories for input patterns degraded by random noise.

Our work in the present paper is organized as follows: Section 2 provides the necessary matrix operations borrowed from minimax algebra for dealing with lattice matrix associative memories. Also, an abstract overview of the main theoretical results obtained from previous research on lattice auto-associative memories is given very briefly. Section 3 describes the redundant encoding technique and simple examples with small sized matrix associative memories are given to make explicit its potential. Section 4 presents some computational experiments with a set of high-dimensional patterns, consisting of gray-scale images, that demonstrate the recall capability of canonical LAMs or AMMs when presented with noisy inputs. Finally, conclusions of this research are given in Section 5.

2 Mathematical Background

Mathematical definitions and key results about lattice matrix associative memories are given here to support the geometrical encoding technique described in the next section. For further material on the mathematical basis the reader is invited to read previous works on the subject [4, 10, 18]. Computation in morphological neural networks is performed within a mathematical framework that is based on lattice algebra systems [1, 13]. More specifically, the extended real minimax algebra [2, 5], denoted by $(\mathbb{R}_{\pm\infty}, \wedge, \vee, +, +')$, is used for lattice associative memories where $\mathbb{R}_{\pm\infty} = \mathbb{R} \cup \{-\infty, +\infty\}$, the binary operations $\wedge, \vee$ denote, respectively, the min and max arithmetical operations between two numbers, and $+, +'$ correspond respectively, to addition and dual addition of arbitrary elements in $\mathbb{R}_{\pm\infty}$. For finite values $x, y \in \mathbb{R}$, we have that $x + y = x +' y$. Matrix computations in minimax algebra are defined elementwise in a similar way as matrix computations are defined in linear algebra. Two fundamental minimax matrix operations are *matrix conjugation* and *matrix max-multiplication*; given a matrix $\mathbf{A}$ of size $m \times p$ and a matrix $\mathbf{B}$ of size $p \times n$ over $\mathbb{R}_{\pm\infty}$, for each

$i = 1, \ldots, m$ and $j = 1, \ldots, n$, they are defined as (t denotes transposition)

$$a^*_{ij} = -a_{ji} \quad \text{or} \quad \mathbf{A}^* = -\mathbf{A}^t \tag{1}$$

$$c_{ij} = \bigvee_{k=1}^{p} (a_{ik} + b_{kj}) \quad \text{or} \quad \mathbf{C} = \mathbf{A} \boxtimes \mathbf{B}. \tag{2}$$

Dually, the *min-product* of matrices $\mathbf{A}$ and $\mathbf{B}$, denoted by $\mathbf{A} \boxtimes \mathbf{B}$, is defined in terms of the generalized $\wedge$-min operation. There are two canonical lattice matrix associative memories known as the *min-memory*, denoted by $\mathbf{W}_{\mathrm{XY}}$, and the *max-memory*, denoted by $\mathbf{M}_{\mathrm{XY}}$. Following a minimax correlated weight rule, each memory stores a set of k associations $(\boldsymbol{x}^\xi, \boldsymbol{y}^\xi)$, represented by (X, Y), where $X = (\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k) \subset \mathbb{R}^n$ and $Y = (\boldsymbol{y}^1, \ldots, \boldsymbol{y}^k) \subset \mathbb{R}^m$. Left equations in (3) and (4) are given in entry format for $i = 1, \ldots, m$ and $j = 1, \ldots, n$; right equations are in matrix notation.

$$w_{ij} = (\mathbf{W}_{\mathrm{XY}})_{ij} = \bigwedge_{\xi=1}^{k} (y_i^\xi - x_j^\xi) \quad ; \quad \mathbf{W}_{\mathrm{XY}} = Y \boxtimes X^* \tag{3}$$

$$m_{ij} = (\mathbf{M}_{\mathrm{XY}})_{ij} = \bigvee_{\xi=1}^{k} (y_i^\xi - x_j^\xi) \quad ; \quad \mathbf{M}_{\mathrm{XY}} = Y \boxtimes X^* \tag{4}$$

If $Y \neq X$ then a memory is called *hetero-associative*, otherwise it is called *auto-associative*; this last case is where our attention is focused in this paper. We repeat two fundamental results on lattice auto-associative memories proved in [4]. First, $\mathbf{W}_{\mathrm{XX}}$ and $\mathbf{M}_{\mathrm{XX}}$ give *perfect recall* for *perfect input* in the sense that, $\mathbf{W}_{\mathrm{XX}} \boxtimes \boldsymbol{x}^\xi = \boldsymbol{x}^\xi$ (resp. $\mathbf{M}_{\mathrm{XX}} \boxtimes \boldsymbol{x}^\xi = \boldsymbol{x}^\xi$) for $\xi = 1, \ldots, k$. Second, since the value of k is not restricted in anyway, $\mathbf{W}_{\mathrm{XX}}$ and $\mathbf{M}_{\mathrm{XX}}$ have *infinite* storage capacity.

Missing parts, occlusions or corruption of exemplar patterns can be considered as "noise" and we speak of *random noise*, when alterations in pattern entries follow a probability density function. Recall capabilities of LAMs for *non-perfect* inputs, requires noise to be classified in three basic types. Let $I = \{1, \ldots, n\}$ then, a distorted version $\widetilde{\boldsymbol{x}}$ of pattern $\boldsymbol{x}$ has undergone an *erosive change* whenever $\widetilde{\boldsymbol{x}} \leq \mathbf{x}$ or equivalently if $\forall i \in I$, $\widetilde{x}_i \leq x_i$. A *dilative change* occurs whenever $\widetilde{\boldsymbol{x}} \geq \mathbf{x}$ or equivalently if $\forall i \in I$, $\widetilde{x}_i \geq x_i$. Let $L, G \subset I$ be two non-empty disjoint sets of indexes. If $\forall i \in L$, $\widetilde{x}_i < x_i$ and $\forall i \in G$, $\widetilde{x}_i > x_i$, then the distorted pattern $\widetilde{\boldsymbol{x}}$ is said to contain *mixed* noise (random or structured). In practical situations, the "perfect recall" requirement is relaxed, and we say that $\mathbf{W}_{\mathrm{XX}}$ (resp. $\mathbf{M}_{\mathrm{XX}}$) is an *almost perfect recall* memory for a set X of patterns if and only if there exists a small rational number $\varepsilon > 0$ close to zero, such that $\mu(\mathbf{W}_{\mathrm{XX}} \boxtimes \widetilde{\boldsymbol{x}}, \boldsymbol{x}) \leq \varepsilon$ (resp. $\mu(\mathbf{M}_{\mathrm{XX}} \boxtimes \widetilde{\boldsymbol{x}}, \boldsymbol{x}) \leq \varepsilon$) for all $\boldsymbol{x} \in X$ with respect to finite sets of noisy versions $\widetilde{\boldsymbol{x}}$ of $\boldsymbol{x}$, where $\mu(\cdot)$ is an adequate measure tailored to treat with binary or real valued patterns.

The lattice matrix auto-associative memories $\mathbf{W}_{\mathrm{XX}}$ and $\mathbf{M}_{\mathrm{XX}}$ can be viewed as lattice transforms of the real vector space $\mathbb{R}^n$ into itself [5, 10, 24]. Thus, for example, the min-memory is considered as a map $W_{\mathrm{X}} : \mathbb{R}^n \to \mathbb{R}^n$ for which $W_{\mathrm{X}}(\boldsymbol{x}) = \mathbf{W}_{\mathrm{XX}} \boxed{\vee}\, \boldsymbol{x}$ for each $\boldsymbol{x} \in \mathbb{R}^n$. Sometimes, to simplify notation we write W for the min-memory lattice transform instead of W_{X}, if reference to the set X is understood in the context of discourse. The *natural factorization* of the map W, shown below in commutative diagram format, provides in a compact way the action that W performs on input vectors.

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "W"] \arrow[d, "\pi"'] & \mathbb{R}^n \\ \mathbb{R}^n/\mathcal{R}_{\mathrm{W}} \arrow[r, "\eta"] & \mathrm{Im}_{\mathrm{W}}(\mathbb{R}^n) \arrow[u, "\iota"'] \end{tikzcd} \tag{5}$$

In diagram (5), $\pi : \mathbb{R}^n \to \mathbb{R}^n/\mathcal{R}_{\mathrm{W}}$ is the projection map (surjective) from the lattice transform domain $\mathbb{R}^n$ to the quotient set $\mathbb{R}^n/\mathcal{R}_{\mathrm{W}}$ obtained from the induced equivalence relation $\mathcal{R}_{\mathrm{W}}$ between pairs of n-dimensional vectors, i.e., vector $\boldsymbol{x}$ is related to vector $\boldsymbol{y}$ if and only if $W(\boldsymbol{x}) = W(\boldsymbol{y})$. Let $W(\boldsymbol{x}) = \boldsymbol{x}$, then $W(\boldsymbol{y}) = \boldsymbol{x}$ means that $\boldsymbol{y}$ is attracted to the fixed point $\boldsymbol{x}$; otherwise, let $W(\boldsymbol{x}) = \boldsymbol{x}' \neq \boldsymbol{x}$, then $W(\boldsymbol{y}) = \boldsymbol{x}'$ and since $\boldsymbol{x}' = W(\boldsymbol{x}')$ (one-step output recall), $\boldsymbol{y}$ is attracted to another fixed point $\boldsymbol{x}'$. The new fixed point $\boldsymbol{x}'$ is simply a minimax combination of the exemplar patterns $\boldsymbol{x}^\xi \in X$, a fact established in [10, 18]. The natural map (bijective), $\eta : \mathbb{R}^n/\mathcal{R}_{\mathrm{W}} \to \mathrm{Im}_{\mathrm{W}}(\mathbb{R}^n)$ establishes a one to one and onto correspondence between each equivalence class $\Omega_{\mathrm{W}}(\boldsymbol{x})$ and the single image value $W(\boldsymbol{x}')$ computed by W on each $\boldsymbol{x}' \in \Omega_{\mathrm{W}}(\boldsymbol{x})$. The equivalence class $\Omega_{\mathrm{W}}(\boldsymbol{x})$ corresponds to the *orbit* or *attraction basin* of a given input pattern $\boldsymbol{x}$; note that the input pattern may be an exemplar pattern $\boldsymbol{x}^\xi \in X$ and that $\boldsymbol{x}'$ may be a corrupted version $\widetilde{\boldsymbol{x}}^\xi$ of $\boldsymbol{x}^\xi$. The fact that $\mathrm{Im}_{\mathrm{W}}(\mathbb{R}^n)$ equals the set of fixed points of the W transform was proved in [10], it is denoted by $F(X)$ or alternatively by $F(W) = F(W_{\mathrm{X}})$. It turns out that $F(X)$ coincide for both auto-associative memories, i.e., $F(X) = F(W) = F(M)$, and consists of the same elements that belong to the linear minimax span of X. Under the natural map η, the inverse image of any fixed point $\boldsymbol{x}$ gives its orbit, i.e., $\eta^{-1}(\boldsymbol{x}) = \Omega_{\mathrm{W}}(\boldsymbol{x})$. Finally, $\iota : F(X) \to \mathbb{R}^n$ is just an immersion map (injective) that assigns each fixed point to itself within the original range of the W map.

To illustrate the previous ideas, we describe a simple example in detail for a two-dimensional memory that stores just one pattern, i.e., let $X = \{\boldsymbol{x}^1\}$ where $\boldsymbol{x}^1 = (x_1^1, x_2^1) \in \mathbb{R}^2$; then, the min-memory matrix $\mathbf{W}_{\mathrm{XX}}$ is given by

$$\mathbf{W}_{\mathrm{XX}} = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} = \begin{pmatrix} 0 & w_{12} \\ w_{21} & 0 \end{pmatrix}. \tag{6}$$

Without loss of generality we take $x_1^1 > x_2^1$, thus $w_{12} = x_1^1 - x_2^1 > 0$, $w_{21} = x_2^1 - x_1^1 < 0$, and $w_{21} = -w_{12}$. There is no need to verify that $\mathbf{W}_{\mathrm{XX}} \boxed{\vee}\, \boldsymbol{x}^1 = \boldsymbol{x}^1$, since this is a specific instance of the perfect recall property that lattice auto-associative memories have on X [4]. Another useful result is that each column vector of the matrix $\mathbf{W}_{\mathrm{XX}}$ is also a fixed point (a proof based on graph theoretical concepts appears in [18], for an alternative shorter argument see the Appendix). The recall stage for an input vector is explained next. Let $\boldsymbol{x} = (x_1, x_2)$ be an input vector different to the stored exemplar pattern $\boldsymbol{x}^1$, i.e., $\boldsymbol{x} \neq \boldsymbol{x}^1$ as shown in Fig. 1.

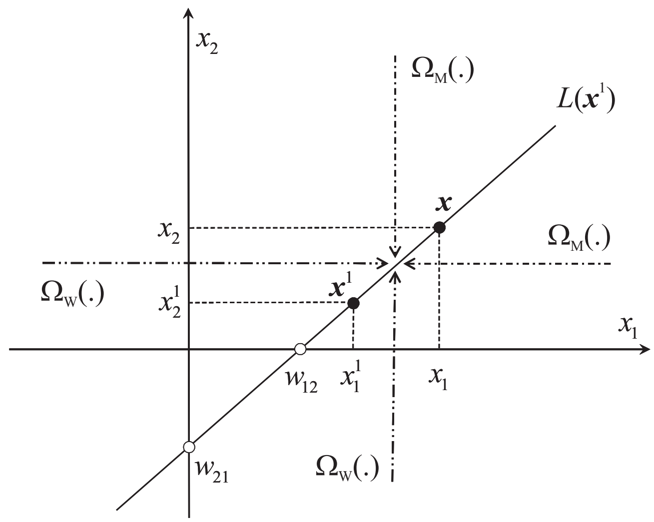

Fig. 1. The solid line $L(\boldsymbol{x}^1)$ of slope 1 equals $F(X)$, the set of fixed points of the min-W_{X} transform for $X = \{\boldsymbol{x}^1\}$. In the diagram, $\boldsymbol{x}^1, \boldsymbol{x} \in F(X)$ but $\boldsymbol{x} \neq \boldsymbol{x}^1$; the left and bottom half rays form a single orbit under W, the upper and right half rays form a single orbit under M, and the arrows show the sense of attraction towards $L(\boldsymbol{x}^1)$

The entries of the output vector $\boldsymbol{y}$ recalled with the min-memory are computed as

$$y_1 = (w_{11} + x_1) \vee (w_{12} + x_2) = x_1 \vee (x_1^1 - x_2^1 + x_2), \tag{7}$$
$$y_2 = (w_{21} + x_1) \vee (w_{22} + x_2) = x_2 \vee (x_2^1 - x_1^1 + x_1). \tag{8}$$

Consider the first coordinate y_1 of the recalled pattern (abscissa of the new point) and let $x_2 = x_2^1$ (same vertical position), then

$$y_1 = x_1^1 \Leftrightarrow x_1 < x_1^1 \tag{9}$$
$$\text{or} \quad y_1 = x_1 \Leftrightarrow x_1 > x_1^1, \tag{10}$$

similarly, for the second coordinate y_2 of the recalled pattern (ordinate of the new point), let $x_1 = x_1^1$ (same horizontal position), hence

$$y_2 = x_2^1 \Leftrightarrow x_2 < x_2^1 \tag{11}$$
$$\text{or} \quad y_2 = x_2 \Leftrightarrow x_2 > x_2^1. \tag{12}$$

From (9) and (11), the orbit of the fundamental memory $\boldsymbol{x}^1$, denoted by $\Omega_{\mathrm{W}}(\boldsymbol{x}^1)$, consists of all those points of $\mathbb{R}^2$ that lie in the *horizontal half-ray* determined by $x_1 \leq x_1^1$ (eroded 1st coordinate) and $x_2 = x_2^1$, together with the *vertical half-ray* determined by $x_1 = x_1^1$ and $x_2 \leq x_2^1$ (eroded 2nd coordinate). Note that, less than or equal signs are used since $\boldsymbol{x}^1$ is included in the orbit, therefore

$$\Omega_{\mathrm{W}}(\boldsymbol{x}^1) = \{\boldsymbol{x} \in \mathbb{R}^2 | x_1 \leq x_1^1, x_2 = x_2^1\} \cup \{\boldsymbol{x} \in \mathbb{R}^2 | x_1 = x_1^1, x_2 \leq x_2^1\}. \tag{13}$$

On the other hand, from (10) and (12), if the abscissa of the input pattern $\boldsymbol{x} \in \mathbb{R}^2$ is strictly to the right of x_1^1 with $x_2 = x_2^1$, or its ordinate is strictly above x_2^1 with $x_1 = x_1^1$, then the recalled pattern is a fixed point *different* to the stored exemplar. Hence, $\boldsymbol{x} \neq \boldsymbol{x}^1$ is such that $\mathbf{W}_{\mathrm{XX}} \boxtimes \boldsymbol{x} = \boldsymbol{x}$, or equivalently,

$$x_1 = x_1 \vee (w_{12} + x_2) \quad \text{and} \quad x_2 = x_2 \vee (w_{21} + x_1), \tag{14}$$

which are satisfied, respectively, if $x_2 = w_{21} + x_1$ and $x_1 = w_{12} + x_2$. However, each one of the last two expressions implies the other by the fact, stated earlier, that $w_{21} = -w_{12}$. Therefore, the coordinates of *any* new fixed point $\boldsymbol{x} \neq \boldsymbol{x}^1$ are related by a single linear equation given by

$$x_2 = x_1 + w_{21}, \tag{15}$$

which is a line of positive slope equal to 1 and intercept at the origin equal to w_{21}. The line passing through the points $(0, w_{21}), (x_1^1, x_2^1) = \boldsymbol{x}^1$, and $(x_1, x_2) = \boldsymbol{x}$, is denoted by $L(\boldsymbol{x}^1)$, to remind its relation to the exemplar pattern $\boldsymbol{x}^1 \in X$. Hence, the min-memory matrix transformation W_{X} has an infinite number of fixed points $\boldsymbol{x} \neq \boldsymbol{x}^1$, such that $\boldsymbol{x} \in L(\boldsymbol{x}^1)$, and the whole plane $\mathbb{R}^2$ is partitioned through their orbits, i.e.,

$$\mathbb{R}^2 = \bigcup \{\Omega_{\mathrm{W}}(\boldsymbol{x}) \mid \boldsymbol{x} \in L(\boldsymbol{x}^1)\}. \tag{16}$$

A similar analysis can be carried out for the max-memory transform M_{X} with $X = \{\boldsymbol{x}^1\}$. Although, the fixed points of M lie on the same line $L(\boldsymbol{x}^1)$, the orbit of the fundamental memory $\boldsymbol{x}^1$, $\Omega_{\mathrm{M}}(\boldsymbol{x}^1) \neq \Omega_{\mathrm{W}}(\boldsymbol{x}^1)$, since it is composed of points that lie in the horizontal half-ray determined by $x_1 \geq x_1^1$ (dilated 1st coordinate) and $x_2 = x_2^1$, together with the vertical half-ray determined by $x_1 = x_1^1$ and $x_2 \geq x_2^1$ (dilated 2nd coordinate), hence

$$\Omega_{\mathrm{M}}(\boldsymbol{x}^1) = \{\boldsymbol{x} \in \mathbb{R}^2 | x_1 \geq x_1^1, x_2 = x_2^1\} \cup \{\boldsymbol{x} \in \mathbb{R}^2 | x_1 = x_1^1, x_2 \geq x_2^1\}. \tag{17}$$

Also, $\Omega_{\mathrm{W}}(\boldsymbol{x}) \cap \Omega_{\mathrm{M}}(\boldsymbol{x}) = \{\boldsymbol{x}\}$ for any $\boldsymbol{x} \in L$, and the dual partition of the plane $\mathbb{R}^2$ is given by $\bigcup \{\Omega_{\mathrm{M}}(\boldsymbol{x}) \mid \boldsymbol{x} \in L(\boldsymbol{x}^1)\}$. We remark that, the line given in (15) can be obtained directly from the $\mathbf{W}_{\mathrm{XX}}$ matrix. Here, $\mathbf{W}_{\mathrm{XX}}^1 = (0, w_{21})^t$ is the intercept point on the x_2 axis and $\mathbf{W}_{\mathrm{XX}}^2 = (w_{12}, 0)^t$ is the intercept point on the x_1 axis. These two points, depicted as hollow dots on Fig. 1, define L. For this example, given a distorted version of $\boldsymbol{x}^1$, denoted by $\widetilde{\boldsymbol{x}}^1$, the min-memory $\mathbf{W}_{\mathrm{XX}}$ (resp. max-memory $\mathbf{M}_{\mathrm{XX}}$) is capable of perfect recall, i.e., $\mathbf{W}_{\mathrm{XX}} \boxtimes \widetilde{\boldsymbol{x}}^1 = \boldsymbol{x}^1$ (resp. $\mathbf{M}_{\mathrm{XX}} \boxtimes \widetilde{\boldsymbol{x}}^1 = \boldsymbol{x}^1$), if and only if $\widetilde{\boldsymbol{x}}^1 \in \Omega_{\mathrm{W}}(\boldsymbol{x}^1)$ (resp. $\widetilde{\boldsymbol{x}}^1 \in \Omega_{\mathrm{M}}(\boldsymbol{x}^1)$), meaning that $\widetilde{\boldsymbol{x}}^1$ must be eroded (resp. dilated) in one of its two coordinates but *not* both.

3 Redundant Encoding

Besides the algebraic characterization given for $F(X)$ in the previous section, a geometrical characterization of the set of fixed points has been recently developed in [10, 12]. The shape of $F(X)$ corresponds to an *m-dimensional convex prismatic beam* with $1 \leq m \leq n$, and the value of m depends on the nature of the underlying finite set X of exemplar pattern vectors. Schematic visualization of the set $F(X)$ is possible only for $n = 1, 2, 3$; for higher dimensions, used is made of concepts from n-dimensional geometry of which the necessary material together with a detailed treatment of the shape of $F(X)$ appears in [10, 12]. Simple numerical examples in lower dimensions (small sized memories) make evident the inherent limitations of any lattice matrix auto-associative memory, in the sense that, neither $\mathbf{W}_{\mathrm{XX}}$ nor $\mathbf{M}_{\mathrm{XX}}$ is able to recall correctly any exemplar pattern that has been corrupted with random noise. Also, the surprising robustness of the min-memory $\mathbf{W}_{\mathrm{XX}}$ with respect to erosive noise as well as the max-memory $\mathbf{M}_{\mathrm{XX}}$ against dilative noise is related to typical applications whose patterns are points in higher dimensional spaces, i.e., for $\xi = 1, \ldots, k$, $\boldsymbol{x}^{\xi} \in \mathbb{R}^n$ with $n \gg 1$. Complementary theoretical developments coupled with computational techniques such as the kernel methods [3, 4, 6, 14, 16, 21], fuzzy MAM models [19, 20, 25], enhanced memory schemes [22, 26], and dendrite based models [7–11], offer alternative solutions to robust pattern recall from noisy inputs. Recently, the technique of noise masking has been introduced to take advantage of the strong properties that single lattice matrix auto-associative memories have [23]. All the previous techniques share in common an increase in computational complexity or in the definition of the network topology; however, none of them exploits the infinite storage capacity property that $\mathbf{W}_{\mathrm{XX}}$ and $\mathbf{M}_{\mathrm{XX}}$ have. Therefore, we propose to use the aforementioned property together with the geometrical characterization of the fixed point set of LAMs, to introduce a simple procedure based on redundant encoding, to enable the recall of pattern approximations to exemplars from distorted inputs.

Given a set $X = \{\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^k\}$ of k exemplar patterns $\boldsymbol{x}^{\xi}$, with finite real entries, *redundant encoding* consists in the storage of additional, carefully design patterns to set X. The new patterns must be spatially related to the original ones and for each $\boldsymbol{x}^{\xi} \in X$, with $\xi = 1, \ldots, k$, the same number p of related patterns, denoted by $\{\boldsymbol{x}^{\xi_1}, \boldsymbol{x}^{\xi_2}, \ldots, \boldsymbol{x}^{\xi_p}\}$, can be stored in the new min- and max-memories defined on the augmented set X^{G} given by

$$X^{\mathrm{G}} = X \cup \bigcup_{\xi=1}^{k} \bigcup_{q=1}^{p} \{\boldsymbol{x}^{\xi q}\} \tag{18}$$

If $|X| = k$ then $|X^{\mathrm{G}}| = k(p+1)$ which poses no problem, since W and M have infinite storage capacity. The purpose to add redundant patterns to the original exemplar pattern set X is to increase the number of fixed points by changing the geometrical shape of $F(X)$, and to expand it, in the sense that $F(X^{\mathrm{G}}) \supset F(X)$. It is *essential* for this scheme to work that the new encoded patterns include erosive and dilative, or mixed approximations of each exemplar pattern. In other words, lattice matrix auto-associative memories should remember not only the exemplars but also some eroded and dilated, or mixed resemblances of them. This simple mechanism endows them with recall capability in the presence of arbitrary noisy inputs. In what follows, we consider the *minimum redundant* case, in which only a single approximate pattern is considered for each exemplar. Thus, $p = 1$ and only $2k$ patterns, $\{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k, \boldsymbol{x}^{11}, \ldots, \boldsymbol{x}^{k1}\}$, are stored in the augmented memories, $\mathbf{W}_{\mathrm{X^G X^G}}$ and $\mathbf{M}_{\mathrm{X^G X^G}}$.

To illustrate the basic idea behind the proposed technique, we elaborate further the example provided in Section 2. Thus, if set X has a single pattern $\boldsymbol{x}$ such that $x_1 = x_2$, then line $L(\boldsymbol{x})$ passes through the origin, and contains all the fixed points of the min- and max-transforms W_{X} and M_{X}, i.e., $F(X) = L(\boldsymbol{x})$ as shown by the dotted line at 45° with respect to the x_1 axis in Fig. 2. The points to the left, right, below, and above, respectively, denoted by, $\boldsymbol{x}^\ell, \boldsymbol{x}^r, \boldsymbol{x}^b$, and $\boldsymbol{x}^a$, represent close approximations to the exemplar pattern $\boldsymbol{x}$. For $\epsilon > 0$, let the coordinates of these points be defined by $\boldsymbol{x}^\ell = (x_1^\ell, x_2^\ell) = (x_1 - \epsilon, x_2)$, $\boldsymbol{x}^r = (x_1^r, x_2^r) = (x_1 + \epsilon, x_2)$, $\boldsymbol{x}^b = (x_1^b, x_2^b) = (x_1, x_2 - \epsilon)$, and $\boldsymbol{x}^a = (x_1^a, x_2^a) = (x_1, x_2 + \epsilon)$. Also, notice that, $\boldsymbol{x}^\ell$ and $\boldsymbol{x}^b$ are eroded, respectively, in its 1st and 2nd coordinates; similarly, $\boldsymbol{x}^r$ and $\boldsymbol{x}^a$ are dilated, respectively, in x_1 and x_2. Although we can adjoin all four points to X to build the augmented set X^{G}, the addition of *only* two points not on the same line, such as $\{\boldsymbol{x}^\ell, \boldsymbol{x}^r\}$, $\{\boldsymbol{x}^a, \boldsymbol{x}^b\}$, $\{\boldsymbol{x}^\ell, \boldsymbol{x}^b\}$, or $\{\boldsymbol{x}^a, \boldsymbol{x}^r\}$, give the same set of fixed points for X^{G}. Let $X^{\mathrm{G}} = \{\boldsymbol{x}, \boldsymbol{x}^a, \boldsymbol{x}^r\}$, then the lattice auto-associative memories of X^{G}, computed from (3) and (4), are given by

$$\mathbf{W}_{\mathrm{X^G X^G}} = \begin{pmatrix} 0 & -\epsilon \\ -\epsilon & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{M}_{\mathrm{X^G X^G}} = \begin{pmatrix} 0 & \epsilon \\ \epsilon & 0 \end{pmatrix}. \tag{19}$$

Alternatively, the max-memory matrix can be computed using the fact that $\mathbf{M} = \mathbf{W}^*$ together with (1). Next, to determine $F(X^{\mathrm{G}})$, we solve the equation $\mathbf{W}_{\mathrm{X^G X^G}} \boxdot \boldsymbol{x}' = \boldsymbol{x}'$ for $\boldsymbol{x}'$ not in X^{G}, equivalently, $x'_i = \bigvee_{j=1}^{2}(w_{ij} + x'_j)$ for $i = 1, 2$. The resulting expressions are given by $x'_1 = x'_1 \vee (x'_2 - \epsilon)$ and $x'_2 = x'_2 \vee (x'_1 - \epsilon)$, from which $x'_2 - \epsilon \leq x'_1$ and $x'_1 - \epsilon \leq x'_2$, consequently, $x'_1 - \epsilon \leq x'_2 \leq x'_1 + \epsilon$. The boundary values for x'_2 in this last compound inequality provide us with the line equations of slope 1, $x'_2 = x'_1 + \epsilon$ and $x'_2 = x'_1 - \epsilon$, labeled in Fig. 2 as $L_a = L(\boldsymbol{x}^a)$ and $L_r = L(\boldsymbol{x}^r)$, respectively. Since $x'_2 \in [x'_1 - \epsilon, x'_1 + \epsilon]$, the fixed point set of X^{G} consists of all points that belong to L_a and L_r, or lie in the infinite band between them. Mathematically, if $\bar{H}_a^-$ denotes the closed half plane on and *below* the line L_a and $\bar{H}_r^+$ denotes the closed half plane on and *above* the line L_r, then, $F(X^{\mathrm{G}}) = \bar{H}_a^- \cap \bar{H}_r^+$. Notice that, by definition, $\bar{H}_a^+ \cap \bar{H}_a^- = L_a$ and $\bar{H}_r^+ \cap \bar{H}_r^- = L_r$, hence the boundary of the fixed point set of X^{G} is given by $\partial F(X^{\mathrm{G}}) = L_a \cup L_r$. The geometric shape of $F(X)$ for n-dimensional patterns is fully developed in [12]. In this example, by encoding two redundant patterns that are spatially related to the single exemplar in X, an expanded fixed point set is obtained such that $F(X) \subset F(X^{\mathrm{G}})$. For pattern recall in the presence of noise, we may consider all fixed points $\boldsymbol{x}' \in F(X^{\mathrm{G}})$ that satisfy the inequality, $|x'_1 - x_1| + |x'_2 - x_2| \leq \epsilon$ (hatched rhombus shown in

Fig. 2), to be "good" approximations to a given exemplar pattern $\boldsymbol{x}$, in the sense described earlier for almost perfect recall using a single lattice auto-associative memory. The parameter ϵ gives us the possibility to increase or decrease the size of the rhombus shaped neighborhood and therefore to relax or constrain the admissible memory outputs.

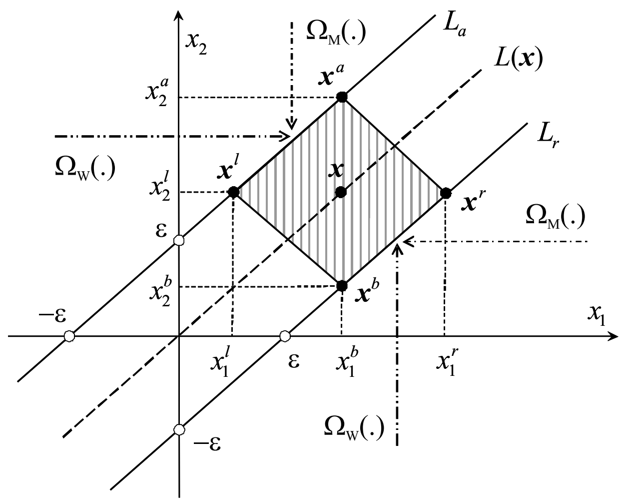

Fig. 2. The solid line $L(\boldsymbol{x})$ of slope 1 equals $F(X)$, the set of fixed points of the min-W_{X} transform for $X = \{\boldsymbol{x}\}$. The points, $\boldsymbol{x}^\ell, \boldsymbol{x}^r, \boldsymbol{x}^a, \boldsymbol{x}^b \in F(X^{\mathrm{G}})$ are eroded or dilated versions of $\boldsymbol{x}$ in a single coordinate. The left and bottom half rays are different orbits under $W_{\mathrm{X}^{\mathrm{G}}}$, the upper and right half rays are distinct orbits under $M_{\mathrm{X}^{\mathrm{G}}}$, and the arrows show the sense of attraction, respectively, towards L_a or L_r

It is reasonable to expect that for high dimensional patterns, redundant encoding can be achieved in various ways. The extension of the base examples, described previously, to higher dimensions ($n > 3$) is impossible to visualize. However, the basic procedure is still the same, since a subset of the coordinates of a given exemplar pattern in X can be eroded to a minimum prescribed value and another subset of coordinates, disjoint from the first, may be dilated to a maximum prescribed value. The rest of the coordinate values would remain the same as those in the original exemplar. In this manner, a *single* redundant pattern related spatially to each exemplar in X is encoded to build X^{G}. Adjoined patterns are then approximations to exemplars which also can be regarded as "noisy" versions of them, in the sense that, mixed *structured* noise is induced by eroding and dilating several coordinates. Clearly, the adequate selection of the coordinate subsets to be modified is determined by pattern dimensionality as well as their nature in a given application.

4 Application Example

To test our proposal, we conducted a computer experiment using 10 gray scale images of size 53×53 pixels as integer valued patterns of dimension 2809. Fig. 3 displays only four of these exemplar pattern images together with a single redundant encoded pattern obtained by assigning a minimum value (0 for extreme erosion) or a maximum value (255 for extreme dilation) to the gray values of a subset of the pixel coordinates, the same for each exemplar in X. Hence, using (3) and (4), the memory matrices $\mathbf{W}_{\mathrm{X^G X^G}}$ and $\mathbf{M}_{\mathrm{X^G X^G}}$, of size 2809×2809, store all associations between the 20 patterns in the augmented set X^{G}.

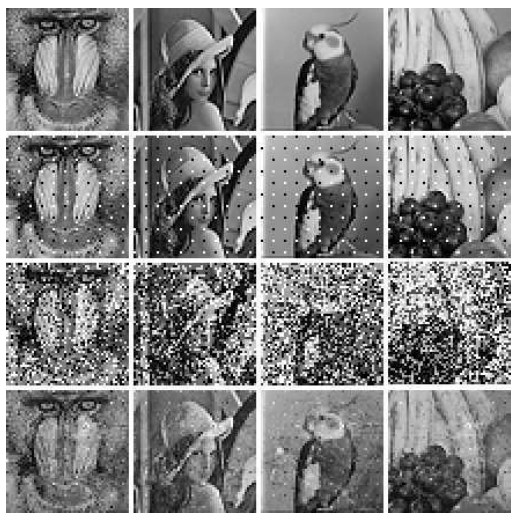

Fig. 3. 1st row: exemplar patterns $\boldsymbol{x}^1, \boldsymbol{x}^3, \boldsymbol{x}^6, \boldsymbol{x}^8 \in X$; 2nd row: redundant encoded patterns $\boldsymbol{x}^{11}, \boldsymbol{x}^{31}, \boldsymbol{x}^{61}, \boldsymbol{x}^{81}$ adjoined to X to form X^{G}; 3rd row: noisy versions of exemplars, respectively, with noise probabilities of $0.3, 0.4, 0.5$, and 0.6; 4th row: recalled pattern approximations.

Each pattern in X was contaminated 100 times by adding random noise with probabilities 0.1 to 0.9 and 128 as offset value. The measure $\mu(\cdot)$ used for almost perfect recall is the *normalized mean square error* (NMSE) defined by $\sigma(\boldsymbol{x}, \widetilde{\boldsymbol{x}}) = \sum_{i=1}^{n}(x_i - \widetilde{x}_i)^2 / \sum_{i=1}^{n} x_i^2$. Fig. 4 shows the performance curves obtained in the recall stage using as final output, the arithmetical mean of the min and max memory outputs, since $\mathbf{W}_{\mathrm{X^G X^G}}$ attracts points towards $F(X^{\mathrm{G}})$ by increasing coordinate values and $\mathbf{M}_{\mathrm{X^G X^G}}$ attracts points towards $F(X^{\mathrm{G}})$ by decreasing coordinate values. Also, if more and more pattern entries are affected by noise, these modified entries are less likely to be equal in the remaining patterns, thus it is no surprise that NMSE decreases as the probability of error increases. In spite of this abnormal behavior, adjoining redundant encoded patterns to the original pattern set is a useful alternative that works best for highly corrupted inputs.

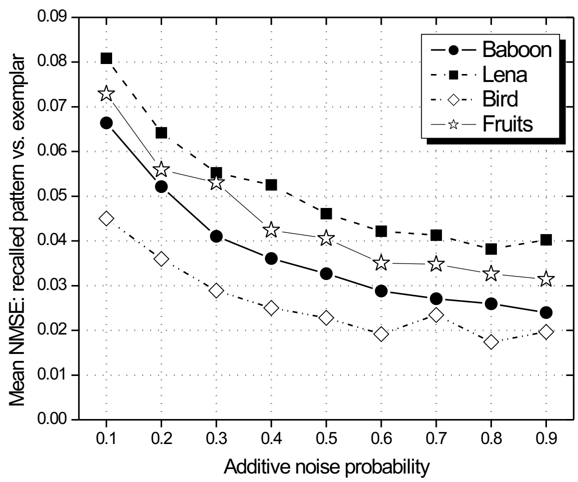

Fig. 4. Recall stage performance curves obtained over 100 trials of producing noise versions of four selected exemplars. NMSE is computed between recalled outputs and corresponding exemplars. In selected patterns, relative error range remains within the same order of magnitude (10^{-2}), for all noise levels

5 Conclusions

In this paper, we introduce a novel and interesting technique that encodes redundant patterns in order to endow lattice auto-associative memories with recall capability in the presence of noise. Basically, the set of fixed points generated by the min or max transforms on the original exemplar set is expanded to allow for the existence of local neighborhoods that might be considered as clusters of approximate versions to exemplars. A complete and simple description using two dimensional patterns has been provided to illustrate the encoding mechanism. Finally, a scenario for gray scale image storage and recall demonstrates the scope and usefulness of minimal redundant pattern encoding in LAMs. Future work will consider different geometrical configurations for pattern encoding, further study on fixed point neighborhood shapes, and more important, carefully designed tests – with or without additional encoding – to compare our proposed model against other associative memory models.

Appendix

Theorem 1. *Let $X = (\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k) \subset \mathbb{R}^n$ be a set of k exemplar patterns and let $(\boldsymbol{w}^1, \ldots, \boldsymbol{w}^\lambda, \ldots, \boldsymbol{w}^n)$ denote the column vectors of the corresponding min-memory $\mathbf{W}_{\mathrm{XX}}$ of size $n \times n$, then $\mathbf{W}_{\mathrm{XX}} \boxed{\vee} \boldsymbol{w}^\lambda = \boldsymbol{w}^\lambda$ for $\lambda = 1, \ldots, n$. Therefore, any column of the matrix W_{XX} is a fixed point for the min-memory transform.*

Proof. By definition, the entries of the column vector $\boldsymbol{w}^\lambda$ of $\mathbf{W}_{\mathrm{XX}}$ for $j = 1, \ldots, n$ are given by $w_j^\lambda = \bigwedge_{\xi=1}^{k} (x_j^\xi - x_\lambda^\xi)$, and the entries of the recalled output pattern $\boldsymbol{y}$ are computed as follows, $y_i = (\mathbf{W}_{\mathrm{XX}} \boxed{\vee} \boldsymbol{w}^\lambda)_i = \bigvee_{j=1}^{n} (w_{ij} + w_j^\lambda) = w_i^\lambda \vee \bigvee_{j \neq i} (w_{ij} + w_j^\lambda)$, for $i = 1, \ldots, n$. Since, for all i, $w_{ij} + w_{j\lambda} \leq w_{i\lambda} = w_i^\lambda$, by Lemma 5.1 in [10], then $\bigvee_{j \neq i} (w_{ij} + w_j^\lambda) \leq w_i^\lambda \Rightarrow y_i = w_i^\lambda$.

A similar result holds for the max-memory $\mathbf{M}_{\mathrm{XX}}$, i.e., any column of the $\mathbf{M}_{\mathrm{XX}}$ matrix is a fixed point for the max-memory transform. Since $\mathbf{W}_{\mathrm{XX}} \boxed{\vee} \mathbf{W}_{\mathrm{XX}} = \mathbf{W}_{\mathrm{XX}}$ and $\mathbf{M}_{\mathrm{XX}} \boxed{\wedge} \mathbf{M}_{\mathrm{XX}} = \mathbf{M}_{\mathrm{XX}}$, the previous result is equivalent to the fact that $\mathbf{W}_{\mathrm{XX}}$ and $\mathbf{M}_{\mathrm{XX}}$ are *idempotent*, respectively, under the max and min products.

Acknowledgment

Gonzalo Urcid thanks the National Research System (SNI) and the National Council of Science and Technology (CONACYT) in Mexico for the partial support provided for this work under grant 22036. José A. Nieves-Vázquez also thanks CONACYT for doctoral scholarship 181663.

References

1. Birkhoff, G.: Lattice Theory 3rd ed, Am. Math. Soc., Providence RI (1967).
2. Cuninghame-Green, R.: Minimax algebra and applications. In: Hawkes, P. (ed) Advances in Imaging and Electron Physics, **90**, Academic Press, New York (1995) 1–121.
3. Hattori, M., Fukui, A., Ito, H.: A fast method of constructing kernel patterns for morphological associative memories. Proc. 9th Int. Conf. on Neural Information Processing, Singapore China (2002) 1058–1063.
4. Ritter, G.X., Sussner, S., Diaz de Leon, J.L.: Morphological associative memories. IEEE Trans. on Neural Networks, **9**(2) (1998) 281–293.
5. Ritter, G.X.: Lattice algebra & Minimax algebra. In: Image Algebra, unpublished manuscript available via anonymous ftp, CISE Department, University of Florida, Gainesville FL (1999) 121–135.
6. Ritter, G.X., Urcid, G., Iancu, L.: Reconstruction of noisy patterns using morphological associative memories. J. of Math. Imaging and Vision, **19**(5) (2003) 95–111.
7. Ritter, G.X., Iancu, L.: A morphological auto-associative memory based on dendritic computing. Proc. IEEE (2004) 915–920.
8. Ritter, G.X., Iancu, L.: Lattice algebra approach to neural networks and pattern classification. Pattern Recognition and Image Analysis, **14**(2) (2004) 190–197.

9. Ritter, G.X., Urcid, G., Schmalz, M.S.: Lattice associative memories that are robust in the presence of noise. SPIE Proc., Mathematical Methods in Pattern and Image Analysis, **5916**, San Diego CA (2005) 59160Q1–Q6.
10. Ritter, G.X., Gader, P.: Fixed Points of Lattice Transforms and Lattice Associative Memories. In: Hawkes, P.(ed) Adv. in Imaging and Electron Physics, **144**. Elsevier, San Diego CA (2006) 165–242.
11. Ritter, G.X., Urcid, G.: Learning in Lattice Neural Networks that Employ Dendritic Computing. In: Kaburlasos, V.G. & Ritter, G.X.(eds) Computational Intelligence based on Lattice Theory, **67**. Springer, Heidelberg, Germany (2007) 25–44.
12. Ritter, G.X., Urcid, G.: Fast autonomous endmember determination using lattice algebra, *in preparation* for IEEE Trans. on PAMI.
13. Skornjakov, L.A.: Elements of Lattice Theory, Adam Hilger, Bristol UK (1977).
14. Sussner, P.: Observations on morphological associative memories and the kernel method. Neurocomputing, **31** (2000) 167–183.
15. Sussner, P.: Fixed points of autoassociative morphological memories. Proc. of the Int. Joint Conf. on Neural Networks, Como, Italy (2000) 611–616.
16. Sussner, P.: Associative morphological memories based on variations of the kernel and dual kernel methods. Neural Networks, **16**(5) (2003) 625–632.
17. Sussner, P.: Recall of patterns using binary and gray-scale autoassociative morphological memories. SPIE Proc., Mathematical Methods in Pattern and Image Analysis, **5916**, San Diego CA (2005) 59160M:1-10.
18. Sussner, P., Valle, M.E.: Gray-scale morphological associative memories. IEEE Trans. on Neural Networks, **17**(3) (2006) 559–570.
19. Sussner, P., Valle, M.E.: Implicative fuzzy associative memories. IEEE Trans. on Fuzzy Systems, **14**(6) (2006) 793–807.
20. Sussner, P., Valle, M.E.: Fuzzy Associative Memories and Their Relationship to Mathematical Morphology. In: Pedrycz, W., Skowron, A. & Kreinovich, V. (eds) Handbook of Granular Computing, Ch. 33, Wiley, (2008).
21. Urcid, G., Ritter, G.X.: Kernel computation in morphological associative memories for gray-scale image recollection. Proc. 5th IASTED Int. Conf. on Signal and Image Processing, Honolulu HI (2003) 450–455.
22. Urcid, G., Nieves-V., J.A. Reyes-G, C.A.: Enhanced fuzzy autoassociative morphological memory for binary pattern recall. Proc. 4th IASTED Int. Conf. on Signal Processing, Pattern Recognition, and Applications, Innsbruck, Austria (2007) 116–121.
23. Urcid, G., Ritter, G.X.: Noise Masking for Pattern Recall using a Single Lattice Matrix Associative Memory. In: Kaburlasos, V.G. & Ritter, G.X.(eds) Computational Intelligence based on Lattice Theory, **67**. Springer, Heidelberg, Germany (2007) 79–98.
24. Urcid, G., Valdiviezo-N., J.C., Generation of lattice independent vector sets for pattern recognition applications. SPIE Proc., Mathematics of Data/Image Pattern Recognition, Compression, Coding, and Encryption X with Applications,**6700**, San Diego CA (2007) 67000C :1–12.
25. Valle, M.E., Sussner, P.: A general framework for fuzzy morphological associative memories. Fuzzy Sets and Systems, **159**(7) (2008) 747–768.
26. Wang, M., Chen, S.: Enhanced fuzzy morphological auto-associative memory based on empirical kernel map. IEEE Trans. on Neural Networks, **16**(3) (2005) 557–564.