Observations on Morphological Associative Memories and the Kernel Method

Peter Sussner
Institute of Mathematics, Statistics, and Scientific Computation
State University of Campinas
Campinas, CEP13081-970, SP, Brazil
email: sussner@ime.unicamp.br
phone: (55) 19/788-5959
fax: (55) 19/289-5808

1 Abstract

The ability of human beings to retrieve information on the basis of associated cues continues to elicit great interest among researchers. Investigations of how the brain is capable to make such associations from partial information have led to a variety of theoretical neural network models that act as associative memories. Several researchers have had significant success in retrieving complete stored patterns from noisy or incomplete input pattern keys by using morphological associative memories. Thus far morphological associative memories have been employed in two different ways: a direct approach which is suitable for input patterns containing either dilative or erosive noise and an indirect one for arbitrarily corrupted input patterns which is based on kernel vectors. In a recent paper [22], we suggested how to select these kernel vectors and we deduced exact statements on the amount of noise which is permissible for perfect recall.

In this paper, we establish the proofs for all our claims made about the choice of kernel vectors and perfect recall in kernel method applications. Moreover, we provide arguments for the success of both approaches beyond the experimental results presented up to this point.

Keywords. *Associative memories, morphological neural networks, kernel method, kernel vectors.*

2 Introduction

The concept of morphological neural networks grew out of the theory of image algebra developed by G.X. Ritter [20]. A sub-algebra of image algebra can be viewed as the mathematical background not only for morphological image processing but also for morphological neural networks [16, 6]. A number of researchers devised morphological neural networks for very specialized applications. J.L. Davidson employed morphological neural networks in order to solve template identification and target classification problems [5, 4]. Suarez-Araujo applied morphological neural networks to compute homothetic auditory and visual invariances [21] Another interesting network consisting of a morphological net and a classical feedforward network used for feature extraction and classification was designed by Won, Gader, and Coffield [24, 25].

The properties of morphological neural networks differ drastically from those of traditional neural network models. These differences are due to the fact that traditional neural network operations consist of linear operations followed by an application of nonlinear activation functions whereas in morphological neural computing the next state of a neuron or in performing the next layer neural network computation involves the nonlinear operation of adding neural values and their synaptic strengths followed by forming the maximum of the results. A fairly comprehensive and rigorous basis for computing with morphological neural networks appeared in [17].

One of the first goals achieved in the development of morphological neural networks was the establishment of a morphological associative memory network (MAM). In its basic form, this model of an associative memory resembles the well-known correlation memory or linear associative memory [9]. As is the case in correlation recording, the morphological associative memory provides for a simple method to add new associations. Correlation recording requires orthogonality of the key vectors in order to exhibit perfect recall of the fundamental associations. The morphological auto-associative memory does not restrict the domain of the key vectors in any way. Thus, as many associations as desired can be encoded into the memory [19]. In the binary case, the limit is 2^n, where n is the length of memory. In comparison, McEliece et al. showed that the (asymptotic) capacity of the Hopfield associative memory is $\frac{n}{2\log n}$ if with high probability the unique fundamental memory is to be recovered, except for a vanishingly small fraction of fundamental memories [13]. Evidently, the information storage capacity (number of bits which can be stored and recalled associatively) of the morphological auto-associative memory also exceeds the respective number for certain linear matrix associative memories which was calculated by Palm [15, 23].

The discrete correlation-recorded Hopfield net is one of the most popular methods for auto-association of binary patterns [7, 8, 1, 12]. Unlike the Hopfield network, which is a recurrent neural network, our model provides the final result in one pass through the network without any significant amount of training. We previously used a number of experiments to demonstrate the MAM's efficiency to deal with either dilative or erosive changes to the input patterns including incomplete patterns [18]. Thus, the morphological associative memory model reveals almost all the desired general characteristics of an associative memory with one notable exception: This network's inability to deal with patterns which include erosive as well as dilative noise at the same time.

We suggested the kernel method as a possible solution to this dilemma. The kernel method depends on the choice of a number of suitable vectors called kernel vectors. In a recent paper, we presented a characterization of kernel vectors which greatly simplifies their choice in the binary case [22]. We were furthermore able to answer the following question: What is the amount of corruption of the key vectors which is admissible in order to maintain perfect recall when using the kernel method? This paper provides the necessary proofs for these results. Furthermore, it includes various observations on the direct approach as well as on the kernel method which confirm our previous assumptions about morphological associative memories.

3 Computational Basis for Morphological Neural Networks

Artificial neural network models are specified by the network topology, node characteristics, and training or learning rules. The underlying algebraic system used in these models is the set of real numbers $\mathbb{R}$ together with the operations of addition and multiplication and the laws governing these operations. This algebraic system, known as a *ring*, is commonly denoted by $(\mathbb{R}, +, \cdot)$.

The basic computations occurring in the proposed morphological network are based on the algebraic lattice structure $(\mathbb{R}_{\pm\infty}, \vee, \wedge, +, +')$. The algebra of matrices over $\mathbb{R}_{\pm\infty}$ which has found widespread applications in the engineering sciences provides for an elegant way to express the total input effect on a morphological neural network layer [18, 19]. The set $\mathbb{R}_{\pm\infty}$ contains the real numbers together with the upper and lower bounds ∞ and $-\infty$. This set forms a bounded lattice ordered group. The symbols $\vee$ and $\wedge$ denote the binary operations of maximum and minimum, respectively. For the purposes of this paper, it is unnecessary to distinguish between the operations $+$ and $+'$ and it suffices to consider operations in the substructure $(\mathbb{R}, \vee, \wedge, +)$.

The focus of this paper is on artificial associative memories. Many popular models of associative memories allow for a formulation using matrices [9, 10, 2, 7, 14, 11, 3]. The model of associative memories described in this paper can also be expressed using products of matrices. For an $m \times p$ matrix A and a $p \times n$ matrix B with entries from $\mathbb{R}$, the matrix $C = A \boxdot\!\!\vee B$, also called the *max product* of A and B, is defined below. The matrix $D = A \boxdot\!\!\wedge B$, the *min product* of A and B, is defined in a similar fashion.

$$c_{ij} = \bigvee_{k=1}^{p} a_{ik} + b_{kj} \ , \ \ d_{ij} = \bigwedge_{k=1}^{p} a_{ik} + b_{kj} \ . \tag{1}$$

Some additional comments concerning lattice based operations are pertinent when discussing morphological network computations. When using the lattice $(\mathbb{R}, \vee, \wedge, +)$, the maximum (or minimum) of two matrices replaces the usual matrix addition of linear algebra. Here the i, jth entry of the matrix $C = A \vee B$ is given by $c_{ij} = a_{ij} \vee b_{ij}$. Similarly, the minimum of two matrices $C = A \wedge B$ is defined by $c_{ij} = a_{ij} \wedge b_{ij}$.

The algebraic structure $(\mathbb{R}, \vee, \wedge, +)$ provides for an elegant duality between matrix operations. For any real number r , we define its *additive conjugate* r^* as $r^* = -r$. Therefore,

$$(r^*)^* = r \quad \text{and} \quad r \wedge u = (r^* \vee u^*)^* \quad \forall r, u \in R \tag{2}$$

Now if $A = (a_{ij})_{m \times n}$ is an $m \times n$ matrix with $a_{ij} \in \mathbb{R}_{\pm\infty}$, then the *conjugate matrix* A^* of A is the $n \times m$ matrix $A^* = (b_{ij})_{n \times m}$ defined by $b_{ij} = [a_{ji}]^*$, where $[a_{ji}]^*$ is the additive conjugate of a_{ji}. It follows that

$$A \wedge B = (A^* \vee B^*)^* \ \text{ and } \ A \boxed{\wedge} B = (B^* \boxed{\vee} A^*)^* \ . \tag{3}$$

This implies that a morphological neural net using the operation $\boxed{\vee}$ can always be reformulated in terms of the operation $\boxed{\wedge}$, and vice versa, by using the duality relations expressed in Equations 3. Therefore, every statement about morphological neural networks induces a dual statement which simply arises by replacing each $\wedge$ symbol with a $\vee$ symbol and vice versa, and by reversing each inequality. During the course of this paper, we will often leave it to the reader to perform this simple transformation.

4 Introduction to Morphological Associative Memories

An associative memory is an input-output system that describes a relation $R \subseteq \mathbb{R}^m \times \mathbb{R}^n$. If $(\mathbf{x}, \mathbf{y}) \in R$, i.e. if the input $\mathbf{x}$ produces the output $\mathbf{y}$, then the associative memory is said to *store* or *record* the *memory association* $(\mathbf{x}, \mathbf{y})$. Unlike conventional computer associative memories, neural network models serving as associative memories are generally capable of retrieving a complete output pattern $\mathbf{y}$ even if the input pattern $\mathbf{x}$ is corrupted or incomplete. The purpose of *auto-associative* memories is the retrieval of $\mathbf{x}$ from corrupted or incomplete versions $\tilde{\mathbf{x}}$. If an artificial associative memory stores associations $(\mathbf{x}, \mathbf{y})$, where $\mathbf{x}$ cannot be viewed as a corrupted or incomplete version of $\mathbf{y}$, then we speak of a *hetero-associative* memory.

The earliest neural network approach to associative memories is the *linear associative memory* or *correlation memory* [9]. Association of patterns $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$ is achieved by means of a matrix-vector product $\mathbf{y} = W \cdot \mathbf{x}$. If we suppose that the goal is to store k vector pairs $(\mathbf{x}^1, \mathbf{y}^1), \ldots, (\mathbf{x}^k, \mathbf{y}^k)$, where $\mathbf{x}^\xi \in \mathbb{R}^n$ and $\mathbf{y}^\xi \in \mathbb{R}^m$ for all $\xi = 1, \ldots, k$, then W is an $m \times n$ matrix given by the following outer product rule

$$\sum_{\xi=1}^{k} \mathbf{y}^\xi \cdot (\mathbf{x}^\xi)' . \tag{4}$$

If X denotes the matrix whose columns are the input patterns $\mathbf{x}^1, \ldots, \mathbf{x}^k$ and if Y denotes the matrix whose columns are the output patterns $\mathbf{y}^1, \ldots, \mathbf{y}^k$, then this equation can be written in the simple form $Y \cdot X'$. If the input patterns $\mathbf{x}^1, \ldots, \mathbf{x}^k$ are orthonormal, then

$$W \cdot \mathbf{x}^i = \left(\mathbf{y}^1 \cdot (\mathbf{x}^1)' + \cdots + \mathbf{y}^k \cdot (\mathbf{x}^k)'\right) \cdot \mathbf{x}^i = \mathbf{y}^i. \tag{5}$$

Thus, we have perfect recall of the output patterns $\mathbf{y}^1, \ldots, \mathbf{y}^k$. If $\mathbf{x}^1, \ldots, \mathbf{x}^k$ are not orthonormal (which they are not in most realistic cases), then filtering processes using threshold functions become necessary in order to retrieve the desired output pattern.

There are two basic approaches to record k vector pairs $(\mathbf{x}^1, \mathbf{y}^1), \ldots, (\mathbf{x}^k, \mathbf{y}^k)$ using a morphological associative memory. The first approach consists of constructing an $m \times n$ matrix W_{XY} as $W_{XY} = Y \boxed{\wedge} X^*$. The second, dual approach consists of constructing an $m \times n$ matrix M_{XY} of the form $M_{XY} = Y \boxed{\vee} X^*$. If the matrix W_{XY} receives a vector $\mathbf{x}$ as input, the product $W_{XY} \boxed{\vee} \mathbf{x}$ is formed. Dually, if the matrix M_{XY} receives a vector $\mathbf{x}$ as input, the product $M_{XY} \boxed{\wedge} \mathbf{x}$ is formed.

As the reader may have already noticed, there are surprising similarities between the morphological associative memories described above and the linear associative memory. These similarities become even more apparent when using an outer product notation for the matrices W_{XY} and M_{XY}.

$$W_{XY} = \bigwedge_{\xi=1}^{k} \left[\mathbf{y}^{\xi} \boxed{\wedge} \left(-\mathbf{x}^{\xi}\right)'\right] \text{ and } M_{XY} = \bigvee_{\xi=1}^{k} \left[\mathbf{y}^{\xi} \boxed{\vee} \left(-\mathbf{x}^{\xi}\right)'\right] \tag{6}$$

Since $\mathbf{y} \boxed{\wedge} \mathbf{x}' = \mathbf{y} \boxed{\vee} \mathbf{x}'$ holds for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}$, we can simplify the notation by denoting morphological outer vector products of this form by $\mathbf{y} \times \mathbf{x}'$.

In view of these equations it is evident that perfect recall is always guaranteed in the case $k = 1$. Two fundamental theorems on morphological associative memories have been established which are concerned with necessary and sufficient conditions for perfect recall of uncorrupted and corrupted patterns [18]. The first theorem which can be found below answers the existence question of perfect recall for sets of pattern pairs. The second theorem establishes bounds for the amount of distortion of the exemplar patterns $\mathbf{x}^{\xi}$ for which perfect recall can be assured. With Theorem 2 we provide a slightly generalized version of this second theorem. This version will turn out to be useful for the upcoming discussions in this paper.

Theorem 1 *$W_{XY} \boxed{\vee} X = Y$ if and only if for each $\xi = 1, \ldots k$ each row of the matrix $W_{XY} - \left[\mathbf{y}^{\xi} \times \left(-\mathbf{x}^{\xi}\right)'\right]$ contains a zero entry. Similarly, $M_{XY} \boxed{\wedge} X = Y$ if and only if for each $\xi = 1, \ldots k$ each row of the matrix $M_{XY} - \left[\mathbf{y}^{\xi} \times \left(-\mathbf{x}^{\xi}\right)'\right]$ contains a zero entry.*

If $X = Y$ (i.e., $\mathbf{y}^{\xi} = \mathbf{x}^{\xi}$, for $\xi = 1, \ldots, k$), we obtain the *morphological auto-associative memories W_{XX}* and M_{XX}. The fact that these memories are able to store and to recall an arbitrary number of patterns is given by a corollary which states that $W_{XX} \boxed{\vee} X = X$ and $M_{XX} \boxed{\wedge} X = X$ for arbitrary matrices X. The reader may want to verify the first equation in the example below.

$$X = \begin{pmatrix} 4 & 3 & 1 \\ 8 & 7 & 2 \\ 2 & 0 & 1 \end{pmatrix}, \quad W_{XX} = \begin{pmatrix} 0 & -4 & 0 \\ 1 & 0 & 1 \\ -3 & -7 & 0 \end{pmatrix}. \tag{7}$$

Apart from unlimited storage capacity, auto-associative morphological memories have a number of other desirable properties such as the ones presented in the following theorems and corollaries.

Theorem 2 *Let $\tilde{\mathbf{x}}^{\gamma}$ denote a distorted version of the pattern $\mathbf{x}^{\gamma}$ and $i \in \{1, \ldots, m\}$ be an arbitrary, but fixed index. The equality $(W_{XY} \boxed{\vee} \tilde{\mathbf{x}}^{\gamma})_i = y_i^{\gamma}$ holds if and only if*

$$\tilde{x}_j^{\gamma} \leq x_j^{\gamma} \vee \left(\bigvee_{\xi \neq \gamma} \left[y_i^{\gamma} - y_i^{\xi} + x_j^{\xi} \right] \right) \quad \forall\, j = 1, \ldots, n \tag{8}$$

and there exists an index $j_i \in \{1, \ldots, n\}$ such that

$$\tilde{x}_{j_i}^{\gamma} = x_{j_i}^{\gamma} \vee \left(\bigvee_{\xi \neq \gamma} \left[y_i^{\gamma} - y_i^{\xi} + x_{j_i}^{\xi} \right] \right). \tag{9}$$

Proof. If W denotes W_{XY}, then the statement $(W \boxed{\vee} \tilde{\mathbf{x}}^{\gamma})_i = y_i^{\gamma}$ is equivalent to the following two conditions:

(a) The inequality $w_{ij} + \tilde{x}_j^{\gamma} \leq y_i^{\gamma}$ holds for all $j = 1, \ldots, n$.

(b) There exists an index $j_i \in \{1, \ldots, n\}$ such that $w_{ij_i} + \tilde{x}_{j_i}^{\gamma} = y_i^{\gamma}$.

The following sequence of equivalences relates condition **(a)** to Equation 8. The proof of the second part of Theorem 2 is similar.

$$w_{ij} + \tilde{x}_j^{\gamma} \leq y_i^{\gamma} \quad \forall\, j = 1, \ldots, n$$

$$\Leftrightarrow \bigwedge_{\xi=1}^{k} \left(y_i^{\xi} - x_j^{\xi} \right) + \tilde{x}_j^{\gamma} \leq y_i^{\gamma} \quad \forall\, j = 1, \ldots, n$$

$$
\begin{aligned}
&\Leftrightarrow\ \ \tilde{x}_j^\gamma \le y_i^\gamma - \bigwedge_{\xi=1}^{k} \left(y_i^\xi - x_j^\xi\right) \quad \forall\, j = 1, \ldots, n \\
&\Leftrightarrow\ \ \tilde{x}_j^\gamma \le y_i^\gamma + \bigvee_{\xi=1}^{k} \left(x_j^\xi - y_i^\xi\right) \quad \forall\, j = 1, \ldots, n \\
&\Leftrightarrow\ \ \tilde{x}_j^\gamma \le \bigvee_{\xi=1}^{k} \left(y_i^\gamma - y_i^\xi + x_j^\xi\right) \quad \forall\, j = 1, \ldots, n \\
&\Leftrightarrow\ \ \tilde{x}_j^\gamma \le x_j^\gamma \vee \bigvee_{\xi \ne \gamma} \left(y_i^\gamma - y_i^\xi + x_j^\xi\right) \quad \forall\, j = 1, \ldots, n\,.
\end{aligned}
$$

Q.E.D.

Corollary 1 *The equation $W_{XX} \boxed{\vee}\, \tilde{\mathbf{x}}^\gamma = \mathbf{x}^\gamma$ holds if and only if $\tilde{\mathbf{x}}^\gamma \le \mathbf{x}^\gamma$ and for each row index $i \in \{1, \ldots, m\}$ there exists a column index $j_i \in \{1, \ldots, n\}$ such that*

$$
\tilde{x}_{j_i}^\gamma = x_{j_i}^\gamma \vee \left(\bigvee_{\xi \ne \gamma} \left[x_i^\gamma - x_i^\xi + x_{j_i}^\xi\right] \right). \tag{10}
$$

Similarly, the equation $M_{XX} \boxed{\wedge}\, \tilde{\mathbf{x}}^\gamma = \mathbf{x}^\gamma$ holds if and only if $\tilde{\mathbf{x}}^\gamma \ge \mathbf{x}^\gamma$ and for each row index $i \in \{1, \ldots, m\}$ there exists a column index $j_i \in \{1, \ldots, n\}$ such that

$$
\tilde{x}_{j_i}^\gamma = x_{j_i}^\gamma \wedge \left(\bigwedge_{\xi \ne \gamma} \left[x_i^\gamma - x_i^\xi + x_{j_i}^\xi\right] \right). \tag{11}
$$

Proof. We restrict ourselves to showing the first statement of the theorem. The proof of the second, dual statement is similar.

Since $W_{XX} \boxed{\vee}\, \tilde{\mathbf{x}}^\gamma = \mathbf{x}^\gamma$, Equation 8 is valid for all $i = 1, \ldots, m$ in the special case $Y = X$ according to Theorem 2. It suffices to show that the right hand side of Equation 8 equals x_j^γ for all $j = 1, \ldots, n$ whenever $Y = X$. Clearly, each of these equations depending on j can be established using bounds from above and from below as follows

$$
\begin{aligned}
x_j^\gamma &\le x_j^\gamma \vee \bigvee_{\xi \ne \gamma} \left[x_i^\gamma - x_i^\xi + x_j^\xi\right] \quad \forall\, i = 1, \ldots, m \\
&\le x_j^\gamma \vee \bigvee_{\xi \ne \gamma} \left[x_j^\gamma - x_j^\xi + x_j^\xi\right] \\
&\le x_j^\gamma \vee \bigvee_{\xi \ne \gamma} x_j^\gamma = x_j^\gamma\,.
\end{aligned} \tag{12}
$$

Q.E.D.

We say that distorted version $\tilde{\mathbf{x}}$ of the pattern $\mathbf{x}$ has undergone an *erosive change* or that $\tilde{\mathbf{x}}$ is an *eroded version* of $\mathbf{x}$ whenever $\tilde{\mathbf{x}} \le \mathbf{x}$ and a *dilative change* or that $\tilde{\mathbf{x}}$ is a *dilated version* of $\mathbf{x}$ whenever $\tilde{\mathbf{x}} \ge \mathbf{x}$. Using this terminology, Corollary 1 implies the following: If $W_{XX} \boxed{\vee}\, \tilde{\mathbf{x}}^\gamma = \mathbf{x}^\gamma$ then $\tilde{\mathbf{x}}^\gamma$ must be an eroded version of $\mathbf{x}^\gamma$. Similarly, if $M_{XX} \boxed{\wedge}\, \bar{\mathbf{x}}^\gamma = \mathbf{x}^\gamma$ then $\bar{\mathbf{x}}^\gamma$ must be a dilated version of $\mathbf{x}^\gamma$.

We provide an example from [18]. Consider the ten pattern images $\mathbf{p}^1, \ldots, \mathbf{p}^{10}$ shown in Figure 1. Using the standard row-scan method, each pattern image $\mathbf{p}^\xi$ can be converted into a pattern vector $\mathbf{x}^\xi = \left(x_1^\xi, \ldots, x_{324}^\xi\right)$ by defining

$$x_{18(i-1)+j}^\xi = \begin{cases} 1 & if \;\; \mathbf{p}^\xi(i,j) = 1 \;\; (= black\ pixel) \\ 0 & if \; \mathbf{p}^\xi(i,j) = 0 \;\; (= white\ pixel). \end{cases}$$

We used the ten pattern vectors $(\mathbf{x}^1, \ldots, \mathbf{x}^{10})$ in constructing the morphological memories W_{XX} and M_{XX}. As expected each individual pattern vector $\mathbf{x}^\xi$ was perfectly recalled in a single application of either W_{XX} or M_{XX}.

Figure 1:
The ten patterns used in constructing the morphological memories. The output of either auto-associative morphological memory W_{XX} or M_{XX} is identical to the input patterns.

The same ten pattern vectors served as bipolar inputs to a discrete Hopfield net operating in snychronous mode. (In this setting, the pattern vectors adopt the values 1 and -1 instead of 1 and 0.) In the table below, we measured the Hamming Distance d between the resulting image after convergence and the original image. Clearly, the Hopfield network did not succeed in recalling any of the original patterns.

letter	"**A**"	"**a**"	"**B**"	"**b**"	"**C**"	"**c**"	"**X**"	"**x**"	"**E**"	"**e**"
d	138	128	157	99	148	140	136	116	161	128

Table 1:
The bottom row represents the Hamming Distance between the pattern in the top row and the output of the Hopfield network when applied to this pattern.

For the above Boolean patterns, a change in pattern values from $\mathbf{p}(i,j) = 1$ to $\mathbf{p}(i,j) = 0$ represents an erosive change. Figure 2 shows three experimental examples where the pattern represented by the letter "**X**" was intentionally corrupted by use of large erosive changes. In contrast to the Hopfield net, which failed to recall the exemplar pattern in each case, the morphological auto-associative memory W_{XX} provides perfect recall in all three cases. Similar observations can be made when processing a pattern corrupted by severe dilative changes using the morphological auto-associative memory M_{XX}.

As an additional example, we randomly eliminated black pixels of each letter image with probability p. We considered a sample of 100 corrupted images of each letter and recorded the number of successful retrievals of the original image. The results are shown in the table below. We conducted the same experiment using a discrete Hopfield. In this case, the success rate was 0% for all patterns and all p.

letter	"**A**"	"**a**"	"**B**"	"**b**"	"**C**"	"**c**"	"**X**"	"**x**"	"**E**"	"**e**"
$p = 0.1$	100	80	100	99	100	89	100	90	100	96
$p = 0.2$	96	63	100	97	100	75	96	81	100	92
$p = 0.3$	92	42	100	94	100	64	87	69	99	88

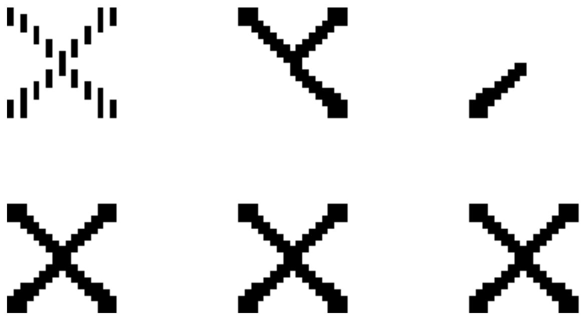

Figure 2:
The top row shows the corrupted input patterns and the bottom row the corresponding output patterns of the morphological memory W_{XX}.

Table 2:
Experimental results after random elimination of black pixels with probablity p. The table shows the number of images in a sample of 100 images which were perfectly recalled in an application of W_{XX}.

Quite naturally, the following question arises: What is the reason for these excellent experimental results? The following obvious corollary of Theorem 2 provides more insight into this question.

Corollary 2 *Let* $\tilde{\mathbf{x}}^\gamma$ *denote an eroded version of the pattern* $\mathbf{x}^\gamma$ *and* $i \in \{1, \ldots, m\}$ *be an arbitrary, but fixed index. The equality* $(W_{XX} \boxtimes \tilde{\mathbf{x}}^\gamma)_i = x^\gamma)_i$ *holds if and only if there exists an index* j_i *such that*

$$\tilde{x}^{\gamma}_{j_i} = x^{\gamma}_{j_i} \vee \left(\bigvee_{\xi \neq \gamma} \left[x^{\gamma}_i - x^{\xi}_i + x^{\xi}_{j_i} \right] \right). \tag{13}$$

Our experiments seem to indicate that the morphological auto-associative memory W_{XX}, when applied to binary patterns, is extremely *robust* in recalling patterns that are distorted by *erosive* noise. (Similarly, the morphological auto-associative memory M_{XX} is extremely *robust* in recalling patterns that are distorted by *dilative* noise.) In terms of Corollary 2, this statement means that for X, $\tilde{X} \in \{0,1\}^{m\times k}$ and $\tilde{\mathbf{x}}^\gamma \leq \mathbf{x}^\gamma$ the following situation occurs with high probability: For each $i \in \{1, \ldots, m\}$ there exists an index j_i such that Equation 13 holds. Let us distinguish between the following two cases with respect to the index i.

- $\tilde{x}^\gamma_i = x^\gamma_i$: Choosing $j_i = i$ gives

$$x^{\gamma}_i \vee \left(\bigvee_{\xi \neq \gamma} \left[x^{\gamma}_i - x^{\xi}_i + x^{\xi}_i \right] \right) = x^{\gamma}_i \vee \bigvee_{\xi \neq \gamma} x^{\gamma}_i = \tilde{x}^{\gamma}_i = x^{\gamma}_i \,. \tag{14}$$

- $\tilde{x}^\gamma_i = 0 < 1 = x^\gamma_i$: In this case $\bigvee_{\xi\neq\gamma} \left[x^{\gamma}_i - x^{\xi}_i + x^{\gamma}_j \right] = \bigvee_{\xi\neq\gamma} \left[1 - x^{\xi}_i + x^{\gamma}_j \right]$ for all $j = 1, \ldots, n$. Most likely, the set $\Psi = \left\{ \xi \,|\, x^{\xi}_i = 0 \right\}$ is not empty. Under this assumption we have

$$\bigvee_{\xi \neq \gamma} \left[x^{\gamma}_i - x^{\xi}_i + x^{\gamma}_j \right] = \bigvee_{\xi \in \Psi} \left[1 + x^{\gamma}_j \right] \geq 1 \,. \tag{15}$$

This inequality implies that Equation 13 holds whenever there exists an index j_i such that

$$x^{\xi}_{j_i} = 0 \ \forall \xi \in \Psi \neq \emptyset \text{ and } \tilde{x}^{\gamma}_{j_i} = x^{\gamma}_{j_i} = 1 \,. \tag{16}$$

Informally speaking, we might say that $\mathbf{x}^{\gamma}$ must have another entry j_i apart from the entry i with $x^{\gamma}_{j_i} = 1$ which has not been eroded, i.e. $\tilde{x}^{\gamma}_{j_i} = 1$ and which provides at least the same amount of information as the entry i, i.e. $x^{\xi}_{j_i} = 0$ for all $\xi \in \Psi$. If X is a collection of artificially structured images such as the letter images in Figure 1 and $x^{\gamma}_i = 1 > 0 = \tilde{x}^{\gamma}_i$ for some indices γ, i, then there are multiple pixels j_i with $x^{\gamma}_{j_i} = 1$ and which offer at least the same information as pixel i. Probably, not all of these black pixels have been eroded. Therefore, $(W_{XX} \boxdot \tilde{\mathbf{x}}^{\gamma})_i = x^{\gamma}_i$ with high probability.

It has to be mentioned, however, that every pattern $\hat{\mathbf{x}}^{\gamma}$ containing some dilative change cannot be recalled using W_{XX}, since by Corollary 1

$$\hat{x}^{\gamma}_i > x^{\gamma}_i \quad \text{for some } i \Rightarrow W_{XX} \boxdot \hat{\mathbf{x}}^{\gamma} \neq \mathbf{x}^{\gamma}. \tag{17}$$

Likewise, every pattern $\hat{\mathbf{x}}^{\gamma}$ containing some erosive change cannot be recalled using M_{XX}, since according to the dual version of Corollary 1

$$\hat{x}^{\gamma}_i < x^{\gamma}_i \quad \text{for some } i \Rightarrow M_{XX} \boxtimes \hat{\mathbf{x}}^{\gamma} \neq \mathbf{x}^{\gamma}. \tag{18}$$

5 Conditions for Kernels

The previous section indicates that morphological associative memories outperform "conventional" associative memories such as correlation recording and the Hopfield net in many aspects. In fact, the morphological associative memory model reveals almost all the desired general characteristics of an associative memory with one notable exception: the network's inability to deal with patterns which include erosive and dilative noise at the same time.

The memory W_{XX} is suitable for application to patterns containing erosive noise and the memory M_{XX} is suitable for application to patterns containing dilative noise. Therefore, an intuitive idea to process noisy versions $\tilde{\mathbf{x}}^{\gamma}$ of $\mathbf{x}^{\gamma}$ containing both erosive and dilative changes is to use a combination of W_{XX} and M_{XX}. Specifically, the output of $M_{XX} \boxtimes \tilde{\mathbf{x}}^{\gamma}$ is multiplied (using $\boxdot$) by W_{XX} or, dually, the output of $W_{XX} \boxdot \tilde{\mathbf{x}}^{\gamma}$ is multiplied (using $\boxtimes$) by M_{XX}. Let us investigate why this approach usually does not work.

Suppose that $\tilde{\mathbf{x}}^{\gamma}$ is a corrupted version of the binary pattern $\mathbf{x}^{\gamma}$ such that $\tilde{\mathbf{x}}^{\gamma} \not\geq \mathbf{x}^{\gamma}$ and $\tilde{\mathbf{x}}^{\gamma} \not\geq \mathbf{x}^{\gamma}$. We will attempt to point out that $(M_{XX} \boxtimes \tilde{\mathbf{x}}^{\gamma})_i = 0$ in most cases, and that therefore $W_{XX} \boxdot (M_{XX} \boxtimes \tilde{\mathbf{x}}^{\gamma})$ will not yield $\mathbf{x}^{\gamma}$, the desired result. Let m_{ij} denote the entries of M_{XX} in this discussion.

- $\tilde{x}^{\gamma}_i = 0 < 1 = x^{\gamma}_i$: We observe that

$$(M_{XX} \boxtimes \tilde{\mathbf{x}}^{\gamma})_i \leq m_{ii} + \tilde{x}^{\gamma}_i = 0 < 1 = x^{\gamma}_i \,. \tag{19}$$

- $i \in I = \{i \mid \tilde{x}^{\gamma}_i = 1 > 0 = x^{\gamma}_i\}$ Let us consider the dilated version $\hat{\mathbf{x}}^{\gamma}$ of $\mathbf{x}^{\gamma}$ which is given by $\hat{x}^{\gamma}_i = \tilde{x}^{\gamma}_i$ for all $i \in I$ and $\hat{x}^{\gamma}_i = x^{\gamma}_i$ for all $i \notin I$. Since the memory M_{XX} is robust in processing dilated versions of $\mathbf{x}^{\gamma}$, we have $M_{XX} \boxtimes \tilde{\mathbf{x}}^{\gamma} = \mathbf{x}^{\gamma}$ in most cases. In particular, $(M_{XX} \boxtimes \hat{\mathbf{x}}^{\gamma})_i = x^{\gamma}_i = 0$ for $i \in I$. The inequality $\tilde{\mathbf{x}}^{\gamma} \leq \hat{\mathbf{x}}^{\gamma}$ implies that

$$(M_{XX} \boxtimes \tilde{\mathbf{x}}^{\gamma})_i \leq (M_{XX} \boxtimes \hat{\mathbf{x}}^{\gamma})_i = x^{\gamma}_i = 0 \ \ \forall i \in I \,. \tag{20}$$

- $\tilde{x}^{\gamma}_i = 1 = x^{\gamma}_i$: If $(M_{XX} \boxtimes \tilde{\mathbf{x}}^{\gamma})_i = x^{\gamma}_i$, then the two conditions of the dual version of Theorem 2 are satisfied in the special case $Y = X$. The first condition states that

$$\tilde{x}^{\gamma}_j \geq x^{\gamma}_j \wedge \left(\bigwedge_{\xi \neq \gamma} \left[x^{\gamma}_i - x^{\xi}_i + x^{\xi}_j \right] \right) \quad \forall\, j = 1, \ldots, n \,. \tag{21}$$

Note that the inequality 21 is guaranteed to hold for all j such that $\tilde{x}_j^\gamma = 1$. Hence, the indices j which satisfy $\tilde{x}_j^\gamma = 0 < 1 = x_j^\gamma$ constitute the critical ones. If J corresponds to the set of eroded pixels, i.e. $J = \left\{ j \mid \tilde{x}_j^\gamma = 0 < 1 = x_j^\gamma \right\}$, we obtain the following set of equivalences.

$$\begin{aligned}
&\bigwedge_{\xi \neq \gamma} \left[x_i^\gamma - x_i^\xi + x_j^\xi \right] \leq 0 \quad \forall j \in J \\
\Leftrightarrow\ & 1 + \bigwedge_{\xi \neq \gamma} \left[x_j^\xi - x_i^\xi \right] \leq 0 \quad \forall j \in J \\
\Leftrightarrow\ & \bigwedge_{\xi \neq \gamma} \left[x_j^\xi - x_i^\xi \right] \leq -1 \quad \forall j \in J \\
\Leftrightarrow\ & \exists \xi_j \ : \ x_i^{\xi j} = 1 \text{ and } x_j^{\xi j} = 0 \ .
\end{aligned} \tag{22}$$

The last statement implies that if there are a multitude of indices j corresponding to eroded pixels, then an application of M_{XX} will most likely cause $\tilde{x}_i^\gamma = x_i^\gamma = 1$ to vanish, especially if $x_i^\xi = 1$ for only a few patterns x^ξ. The indices i which satisfy $x_i^\xi = 1$ for a small number of patterns x^ξ are the most important ones when trying to retrieve x^γ by means of W_{XX}.

Inspite of these difficulties, a modified approach can be applied to create a morphological associative recall memory that is robust in the presence of random noise (i.e. both dilative and erosive noise), even in the general situation when $X \neq Y$ and X and Y are not boolean. A memory M is defined as a memory which associates each input pattern $\mathbf{x}^\xi$ with an intermediate pattern $\mathbf{z}^\xi$. Furthermore, a memory W is defined such that each pattern $\mathbf{z}^\xi$ is associated with the corresponding output pattern $\mathbf{y}^\xi$. In other words, we obtain the following equations:

$$W \boxed{\vee} (M \boxed{\wedge} \mathbf{x}^\xi) = W \boxed{\vee} \mathbf{z}^\xi = \mathbf{y}^\xi \tag{23}$$

Under certain conditions depending on Z, the matrix M_{ZZ} can serve as M and the matrix W_{ZY} can serve as W. In this case, the $n \times k$ matrix $Z = (\mathbf{z}^1, \mathbf{z}^2, \ldots, \mathbf{z}^k)$ is called a *kernel* for (X, Y). Furthermore, for properly chosen vectors $\mathbf{z}^\xi$, we also have $M_{ZZ} \boxed{\wedge} \tilde{\mathbf{x}}^\xi = \mathbf{z}^\xi$ for most corrupted versions $\tilde{\mathbf{x}}^\xi$ of $\mathbf{x}^\xi$. As an example suppose that X is the matrix whose columns correspond to the letter images of Figure 1. The sparse images below each original letter image in Figure 5 are possible choices for kernel vectors in the sense of the definition below in the auto-associative case.

Let us recall the formal definition of kernel:

Definition *An $n \times k$ matrix $Z = (\mathbf{z}^1, \mathbf{z}^2, \ldots, \mathbf{z}^k)$ is called a* kernel *for (X, Y) if and only if the following two conditions are satisfied:*

1. *$M_{ZZ} \boxed{\wedge} X = Z$.*
2. *$W_{ZY} \boxed{\vee} Z = Y$.*

Although the definition indicates if a given $n \times k$ matrix represents a kernel, conditions 1. and 2. are not helpful when faced with the practical problem of selecting kernel vectors. Ritter et al. suggested [18] that for binary valued patterns a sparse representation $\mathbf{z}^\gamma$ of $\mathbf{x}^\gamma$ is a kernel vector if

$$\mathbf{z}^\gamma \wedge \mathbf{x}^\xi = \mathbf{0} \quad \forall \gamma \neq \xi. \tag{24}$$

Since the original patterns $\mathbf{x}^\xi$ will often exhibit significant overlay, it will not be possible to satisfy Equation 24 in most cases. The goal of this section consists in generalizing Equation 24 in such a way that kernels can be constructed easily. As a first step in reaching this goal, we present the following two lemmas which a direct consequences of the fundamental Theorems 1 and 2. The matrices X, Y, and Z are arbitrarily valued.

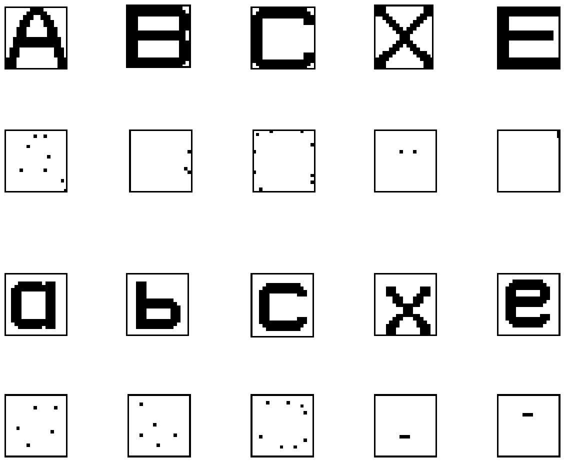

Figure 3:
An example of kernel images as sparse representations of letters. A particular kernel image corresponds to the letter image directly above the kernel image.

Lemma 1 *An $n \times k$ matrix $Z \leq X$ satisfies the equation $M_{ZZ} \boxed{\wedge} X = Z$ if and only if for all $\gamma \in \{1, \ldots, k\}$ and for all $i \in \{1, \ldots, m\}$ there exist indices $j \in \{1, \ldots, n\}$ such that*

$$x_j^\gamma + z_i^\xi \leq z_i^\gamma + z_j^\xi \;\; \forall \xi = 1, \ldots, k\,. \tag{25}$$

Proof. Let us choose an arbitrary index γ. Since $\mathbf{z}^\gamma \leq \mathbf{x}^\gamma$, an application of the dual version of Corollary 1 yields that $M_{ZZ} \boxed{\wedge} \mathbf{x}^\gamma = \mathbf{z}^\gamma$ if and only if for each row index $i \in \{1, \ldots, m\}$ there exists a column index $j_i \in \{1, \ldots, n\}$ such that

$$x_{j_i}^\gamma \leq z_{j_i}^\gamma \wedge \left(\bigwedge_{\xi \neq \gamma} \left[z_i^\gamma - z_i^\xi + z_{j_i}^\xi \right] \right). \tag{26}$$

Now it merely remains to show that Equations 26 and 25 are equivalent.

$$\begin{aligned} & x_{j_i}^\gamma \leq z_{j_i}^\gamma \wedge \bigwedge_{\xi \neq \gamma} \left[z_i^\gamma - z_i^\xi + z_{j_i}^\xi \right] \\ \Leftrightarrow\ & x_{j_i}^\gamma \leq \bigwedge_{\xi = 1}^{k} \left[z_i^\gamma - z_i^\xi + z_{j_i}^\xi \right] \\ \Leftrightarrow\ & x_{j_i}^\gamma \leq z_i^\gamma - z_i^\xi + z_{j_i}^\xi \;\; \forall \xi = 1, \ldots, k \\ \Leftrightarrow\ & x_{j_i}^\gamma + z_i^\xi \leq z_i^\gamma + z_{j_i}^\xi \;\; \forall \xi = 1, \ldots, k\,. \end{aligned}$$

Q.E.D.

Lemma 2 (Cor. 2.1 of [18]) *The equation $W_{ZY} \boxed{\vee} Z = Y$ holds if and only if for each $\gamma \in \{1, \ldots, k\}$ and for each $i \in \{1, \ldots, m\}$ there exists an index $j \in \{1, \ldots, n\}$ (depending on γ and i) such that*

$$z_j^\gamma - y_i^\gamma = \bigvee_{\xi=1}^{k} (z_j^\xi - y_i^\xi)\,. \tag{27}$$

We now restrict our attention to the boolean case, where X, Y, and Z are binary valued. First, we attempt to characterize the binary patterns Z which satisfy $M_{ZZ} \boxed{\wedge}\, X = Z$.

Theorem 3 *Let X and Z be binary patterns with $X \geq Z$. If $\mathbf{z}^\gamma \wedge \mathbf{z}^\xi = \mathbf{0}$ and $\mathbf{z}^\gamma \not\leq \mathbf{x}^\xi$ for all γ, ξ such that $\gamma \neq \xi$, then*

$$M_{ZZ} \boxed{\wedge}\, X = Z\ . \tag{28}$$

Proof. By Lemma 1 it is enough to show that for all $\gamma \in \{1, \ldots, k\}$ and for all $i \in \{1, \ldots, m\}$ there exist indices $j \in \{1, \ldots, n\}$ such that Equation 25 holds for all $\xi \in \{1, \ldots, k\}$. Let us consider arbitrary, but fixed indices γ, i and choose $\xi_\circ$ such that $z_i^{\xi_\circ}$ is maximal. Since $\mathbf{x}^\gamma \not\geq \mathbf{z}^{\xi_\circ}$ by assumption, there exists an index j with $0 = x_j^\gamma < z_j^{\xi_\circ} = 1$. Hence,

$$x_j^\gamma + z_i^{\xi_\circ} \leq z_j^{\xi_\circ} \leq z_i^\gamma + z_j^{\xi_\circ} \tag{29}$$

This choice of j works for all $\xi = 1, \ldots, k$ since $z_i^\xi = 0$ for all $\xi \neq \xi_\circ$ due to the fact that $z_i^{\xi_\circ}$ is maximal and $\mathbf{z}^{\xi_\circ} \wedge \mathbf{z}^\xi = \mathbf{0}$ for all $\xi \neq \xi_\circ$.

Q.E.D.

Theorem 4 *Let Y and Z be arbitrary binary patterns. If $\mathbf{z}^\gamma \wedge \mathbf{z}^\xi = \mathbf{0}$ and $\mathbf{z}^\gamma \not\leq \mathbf{z}^\xi$ for all γ, ξ such that $\gamma \neq \xi$, then*

$$W_{ZY} \boxed{\vee}\, Z = Y\ . \tag{30}$$

Proof. Let us fix an arbitrary tuple of indices i and γ. By Lemma 2, it is enough to show that there exists an index $j \in \{1, \ldots, n\}$ (depending on γ and i) such that $z_j^\gamma - y_i^\gamma = \bigvee_{\xi=1}^{k} (z_j^\xi - y_i^\xi)$ or equivalently,

$$z_j^\gamma - y_i^\gamma \geq z_j^\xi - y_i^\xi \ \ \forall\, \xi \neq \gamma\ . \tag{31}$$

Since $\mathbf{z}^\gamma \not\leq \mathbf{z}^\xi$ for all $\xi \neq \gamma$, there is an index j with $z_j^\gamma = 1$. This equality implies that $z_j^\xi = 0$ for all $\xi \neq \gamma$ because $\mathbf{z}^\gamma \wedge \mathbf{z}^\xi = \mathbf{0}$. Now the desired inequality emerges as follows.

$$z_j^\gamma - y_i^\gamma = 1 - y_i^\gamma \geq 0 \geq -y_i^\gamma = z_j^\xi - y_i^\xi \ \ \forall\, \xi \neq \gamma\ . \tag{32}$$

Q.E.D.

Theorem 5 *Let X, Y, Z be binary patterns with $Z \leq X$. If*

$$\mathbf{z}^\gamma \wedge \mathbf{z}^\xi = \mathbf{0} \ \text{ and } \ \mathbf{z}^\xi \not\leq \mathbf{x}^\gamma \ \ \forall\, \gamma \text{ and } \xi \text{ with } \gamma \neq \xi\ , \tag{33}$$

then Z is a kernel for (X, Y).

Proof. By Theorem 3, $M_{ZZ} \boxed{\wedge}\, X = Z$. The fact that $\mathbf{z}^\xi \not\leq \mathbf{x}^\gamma$ for $\gamma \neq \xi$ implies that $\mathbf{z}^\xi \not\leq \mathbf{z}^\gamma$ for $\gamma \neq \xi$ given that $X \geq Z$. The latter argument assures that Theorem 4 can be applied yielding $W_{ZY} \boxed{\vee}\, Z = Y$. Thus, Z is a kernel for (X, Y) by definition.

Q.E.D.

Given a particular problem of associating binary patterns $\mathbf{x}^\xi$ with binary pattern $\mathbf{y}^\xi$, the kernel with respect to (X, Y) can now be easily chosen by means of Theorem 5. The reader may wish to test the kernel images in Figure 5 for these conditions. Note that Equation 33 generalizes Ritter's original assumptions about kernel vectors found in Equation 24.

6 Conditions for Perfect Recall of Corrupted Patterns

In view of the fact that the results of the previous section induce a practical method for selecting kernels in an association problem, this section poses the following question: Given a collection of kernel vectors whose choice is based on Theorem 5, what is the amount of corruption of the input patterns which is admissible in order to maintain perfect recall? In other words, given a corrupted version $\tilde{\mathbf{x}}^\gamma$ of $\mathbf{x}^\gamma$, under what conditions is

$$W_{ZY} \boxed{\vee} (M_{ZZ} \boxed{\wedge} \tilde{\mathbf{x}}^\gamma) = \mathbf{y}^\gamma. \tag{34}$$

Theorem 5 again comes to the rescue when looking for an answer to this question. Note that the equation $W_{ZY} \boxed{\vee} (M_{ZZ} \boxed{\wedge} \tilde{\mathbf{x}}^\xi) = \mathbf{y}^\xi$ holds for all $\xi \in \{1, \ldots, k\}$, if Z is a kernel for $(\tilde{X}, Y)$, where $\tilde{X}$ denotes the matrix $(\tilde{\mathbf{x}}^1, \tilde{\mathbf{x}}^2, \ldots, \tilde{\mathbf{x}}^k)$

Theorem 6 *If Z is a kernel for (X, Y) satisfying the conditions 33 then we have perfect recall for a distorted version $\tilde{\mathbf{x}}^\gamma$ of $\mathbf{x}^\gamma$, in other words $W_{ZY} \boxed{\vee} (M_{ZZ} \boxed{\wedge} \tilde{\mathbf{x}}^\gamma) = \mathbf{y}^\gamma$, if*

$$\tilde{\mathbf{x}}^\gamma \geq \mathbf{z}^\gamma \text{ and } \tilde{\mathbf{x}}^\gamma \not\geq \mathbf{z}^\xi \;\; \forall \xi \neq \gamma. \tag{35}$$

A comparison of Figure 5 and the following Figure 6 shows that the distorted versions of the letters A, B, and X represented in Figure 6 satisfy Equation 35.

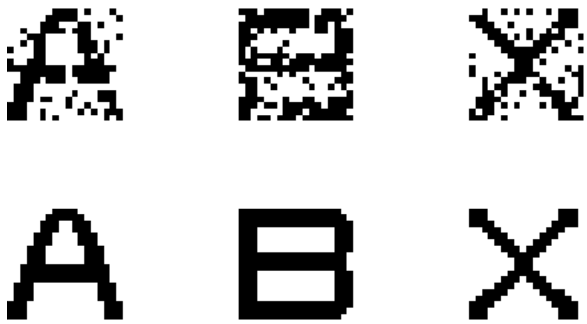

Figure 4:
An application of the kernel method. Each letter was corrupted by randomly reversing each bit with a probability of 0.15. The associative memory $\{$*input* $\rightarrow M_{ZZ} \rightarrow W_{ZY} \rightarrow$ *output*$\}$ was trained using the ten exemplars shown in Figure 1. Presenting the memory with the corrupted patterns of the letters A, B, and X resulted in perfect recall (lower row).

Suppose now that Z is a kernel for (X, Y) of the form 33 and suppose that one of the two sufficient conditions for perfect recall is violated. If $\tilde{\mathbf{x}}^\gamma \geq \mathbf{z}^\xi$ for some $\xi \in \{1, \ldots, k\}$ then $M_{ZZ} \boxed{\wedge} \tilde{\mathbf{x}}^\gamma \geq M_{ZZ} \boxed{\wedge} \mathbf{z}^\xi = \mathbf{z}^\xi$. Assuming that $\mathbf{z}^\xi \not\leq \mathbf{z}^\gamma$, we obtain that $W_{ZY} \boxed{\vee} (M_{ZZ} \boxed{\wedge} \tilde{\mathbf{x}}^\gamma) \neq \mathbf{y}^\gamma$.

As mentioned earlier in this section, if $\tilde{\mathbf{x}}^\gamma \ngeq \mathbf{z}^\gamma$ then $M_{ZZ} \boxed{\wedge} \tilde{\mathbf{x}}^\gamma = \mathbf{0}$ in most cases. Since $\mathbf{0} \leq \mathbf{z}^\xi$ for all $\xi \in \{1, \ldots, k\}$, the product $W_{ZY} \boxed{\vee} (M_{ZZ} \boxed{\wedge} \tilde{\mathbf{x}}^\gamma)$ is then bounded from above by $\mathbf{y}^\xi$ for all ξ, and thus by $\bigwedge_{\xi=1}^{k} \mathbf{y}^\xi \neq \mathbf{y}^\gamma$. A similar argumentation can be applied if $M_{ZZ} \boxed{\wedge} \tilde{\mathbf{x}}^\gamma \leq \mathbf{z}^\xi$ holds only for some $\xi \neq \gamma$.

7 Concluding Remarks

Morphological Associative Memories (MAMs) are a recent development in the theory of artificial neural networks. Interesting results on MAMs such as infinite storage capacity, the ability to recognize partially occluded patterns, and the ease of incorporating additional pattern associations into the memory have already been established. A combination of morphological associative memories M and W forming an input-output system $\{\, input \;\rightarrow\; M \;\rightarrow\; W \;\rightarrow\; output \,\}$ has been used to deal with input patterns which are corrupted by dilative and by erosive noise at the same time. The construction of the MAMs M and W depends on kernel vectors whose selection is greatly simplified by the results of this paper. Moreover, we determined conditions for perfect recall in kernel method applications. The results of this paper provide a basis for further research on MAMs. In an upcoming paper, we will estimate and increase the probability of perfect recall in MAM applications.

References

[1] Y. Abu-Mostafa and J. St. Jacques. Information capacity of the Hopfield model. *IEEE Transactions on Inf. Theory*, 7:1–11, 1985.

[2] J.S. Albus. A new approach to manipulator control: The cerebellar model articulation controller. *Journal of Dynamic Systems Measurement and Control, Transactions of the ASME*, 97:220–227, 1975.

[3] J.A. Austin. Adam: A distributed associative memory for scene analysis. In *Proceedings of the IEEE First International Conference on Neural Networks*, volume IV, page 285, San Diego, 1987.

[4] J.L. Davidson. Simulated annealing and morphological neural networks. In *Image Algebra and Morphological Image Processing III*, volume 1769 of *Proceedings of SPIE*, pages 119–127, San Diego, CA, July 1992.

[5] J.L. Davidson and F. Hummer. Morphology neural networks: An introduction with applications. *IEEE Systems Signal Processing*, 12(2):177–210, 1993.

[6] J.L. Davidson and G.X. Ritter. A theory of morphological neural networks. In *Digital Optical Computing II*, volume 1215 of *Proceedings of SPIE*, pages 378–388, July 1990.

[7] J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. volume 79 of *Proceedings of the National Academy of Sciences*, pages 2554–2558, U.S.A., April 1982.

[8] J.J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. volume 81 of *Proceedings of the National Academy of Sciences*, pages 3088–3092, U.S.A., May 1984.

[9] T. Kohonen. Correlation matrix memory. *IEEE Transactions on Computers*, C-21:353–359, 1972.

[10] T. Kohonen and M. Ruohonen. Representation of associated data by computers. *IEEE Transactions on Computers*, C-22:701–702, 1973.

[11] B. Kosko. Adaptive bidirectional associative memories. 26:4947–4960, 1987.

[12] R.P. Lippmann. An introduction to computing with neural nets. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-4:4–22, 1987.

[13] R. McEliece and et. al. The capacity of Hopfield associative memory. *Transactions on Information Theory*, 1:33–45, 1987.

[14] K. Okajima, S. Tanaka, and S. Fujiwara. A heteroassociative memory with feedback connection. In *Proceedings of the IEEE First International Conference on Neural Networks*, volume II, pages 711–718, San Diego, 1987.

[15] G. Palm. On associative memory. *Biological Cybernetics*, 36:19–31, 1980.

[16] G.X. Ritter and J. L. Davidson. Recursion and feedback in image algebra. In *SPIE's 19th AIPR Workshop on Image Understanding in the 90's*, Proceedings of SPIE, McLean, Va., October 1990.

[17] G.X. Ritter and P. Sussner. An introduction to morphological neural networks. In *Proceedings of the 13th International Conference on Pattern Recognition*, pages 709–717, Vienna, Austria, 1996.

[18] G.X. Ritter, P. Sussner, and J.L. Diaz de Leon. Morphological associative memories. *IEEE Transactions on Neural Networks*, 9(2):281–293, March 1998.

[19] G.X. Ritter, P. Sussner, and W.B. Hacker. Associative memories with infinite storage capacity. In *InterSymp' 97, 9th International Conference on Systems Research Informatics and Cybernetics*, Baden-Baden, Germany, 1997. Invited Plenary Paper.

[20] G.X. Ritter, J.N. Wilson, and J.L. Davidson. Image algebra: An overview. *Computer Vision, Graphics, and Image Processing*, 49(3):297–331, March 1990.

[21] C.P. Suarez-Araujo. Novel neural network models for computing homothetic invariances: An image algebra notation. *Journal of Mathematical Imaging and Vision*, 7(1):69–83, 1997.

[22] P. Sussner. Kernels for morphological associative memories. In *Proceedings of the International ICSA/IFAC Symposium on Neural Computation*, Vienna, September 1998.

[23] D.J. Willshaw, O.P. Buneman, and H.C. Longuet-Higgins. Non-holographic associative memory. *Nature*, 222:960–962, 1969.

[24] Y. Won and P.D. Gader. Morphological shared weight neural network for pattern classification and automatic target detection. volume 4 of *Proceedings of the 1995 IEEE International Conference on Neural Networks*, pages 2134–2138, Perth, Australia, November 1995.

[25] Y. Won, P.D. Gader, and P. Coffield. Morphological shared-weight networks with applications to automatic target recognition. *IEEE Transactions on Neural Networks*, 8(5):1195 –1203, September 1997.