Enhanced FMAM Based on Empirical Kernel Map

Min Wang and Songcan Chen*

***Abstract*—The existing morphological auto-associative memory models based on the morphological operations, typically including Morphological auto-Associative Memories (auto-MAM) proposed by Ritter *et al.* and our Fuzzy Morphological auto-Associative Memories (auto-FMAM), have many attractive advantages such as unlimited storage capacity, one-shot recall speed and good noise-tolerance to single erosive or dilative noise. However, they suffer from the extreme vulnerability to noise of mixing erosion and dilation, resulting in great degradation on recall performance. To overcome this shortcoming, we focus on FMAM and propose an Enhanced FMAM (EFMAM) based on the empirical kernel map. Although it is simple, EFMAM can significantly improve the auto-FMAM with respect to the recognition accuracy under hybrid-noise and computational effort. Experiments conducted on the thumbnail-sized faces (28×23 and 14×11) scaled from the ORL database show the average accuracies of 92%, 90% and 88% with 40 classes under 10%, 20% and 30% randomly generated hybrid-noises respectively, which are far higher than the auto-FMAM (67%, 46%, 31%) under the same noise levels.**

***Index Terms*—Associative memory, Empirical kernel map, Face recognition, Fuzzy mathematics, Morphological neural networks**

I. Introduction

In recent years, a number of different morphological neural network models and applications have emerged [1]–[9]. Morphological neural networks can be characterized as neural networks which calculate a maximum or a minimum of a sum at each node. The focus of this paper is morphological auto-associative memory. The existing models include Morphological auto-Associative Memories (auto-MAM) [4] proposed by Ritter *et al* and our Fuzzy Morphological auto-Associative Memories (auto-FMAM) [8]. MAM is based on the lattice algebra $(R,\vee,\wedge,+)$ or morphological operation, where the symbols $\vee$ and $\wedge$ denote the binary operations of maximum and minimum, respectively. MAM behaves more like human associative memories than the traditional semi-linear models such as the Hopfield net [4]. Once a pattern has been memorized, recall is instantaneous when the MAM is presented with the pattern. In the absence of noise, an auto-MAM will always provide perfect recall for any number of patterns programmed into its memory! The auto-MAM W_{XX} is extremely robust in recalling patterns that are distorted due to erosive changes, while auto-MAM M_{XX} is extremely robust in recalling patterns that are distorted due to dilative changes [4].

Originated from the basic ideas of MAM, the fuzzy morphological associative memory (FMAM) uses two basic morphological operations $(\vee,\cdot)$, $(\wedge,\cdot)$ instead of fuzzy operation $(\vee,\wedge)$ in fuzzy associative memory [8]. FMAM solves MAM's fuzzy rules memory problem. Under certain conditions, FMAM can be viewed as a new encoding way of fuzzy associative memory such that it can embody fuzzy operators and the concepts of fuzzy membership value and fuzzy rules [8]. Besides, auto-FMAM has the same attractive advantages as auto-MAM, such as unlimited storage capacity, one-shot recall speed and good noise-tolerance to either erosive or dilative noise.

The above advantageous properties make auto-MAM and auto-FMAM appropriate for face recognition, especially for the large scale real-world face databases. However, it is impossible for an actual face image to only exhibit erosive or dilative noise. When the images exhibit both erosive and dilative noise, both auto-MAM and auto-FMAM suffer from the extreme vulnerability, which brings the difficulty to distinguish among a number of people and thus weakens the practicality of auto-MAM and auto-FMAM.

It should be mentioned that perfect recall of Boolean noisy patterns using the idea of kernels was settled [4]. But the definition of kernel given by Ritter *et al* is essentially different from ours. Our kernel is a map: $X \times X \to R_+$, while Ritter's kernel definition is given as follows:

Let $X=(\mathbf{x}^1,\ldots,\mathbf{x}^k)$, $Y=(\mathbf{y}^1,\ldots,\mathbf{y}^k)$ be a set of pattern pairs composed by input $\mathbf{x}^i$ and its desired output $\mathbf{y}^i$, an $m\times k$ matrix $Z=(\mathbf{z}^1,\mathbf{z}^2,\ldots,$

Manuscript received February 26, 2004; revised September 9, 2004. This work was supported in part by the National Science Foundations of China and of Jiangsu under Grant Nos. 60271017 and BK2002092, Jiangsu Natural Science Key Project under Grant BK2004001, the "QingLan" Project Foundation of Jiangsu Province and in part by the Returnee Foundation of China Scholarship Council.

M. Wang is with the Computer Science and Engineering Department, Nanjing University of Aeronautics and Astronautics, Nanjing, 210016 P.R.C. (e-mail: wm_wangmin@yahoo.com.cn).

* S.C. Chen (corresponding author) is with the Computer Science and Engineering Department, Nanjing University of Aeronautics and Astronautics, Nanjing, 210016 P.R.C. (phone: +86-25-8489-2805; fax: +86-25-8489-2400; e-mail: s.chen@ nuaa.edu.cn).

$\mathbf{z}^k$) is called a kernel for (X,Y) if and only if there exists a memory matrix W such that $W\boxtimes(M_{ZZ}\boxtimes X)=Y$, where m is the dimension of $\mathbf{z}^k$, $M_{ZZ}=\bigvee_{l=1}^{k}(\mathbf{z}^l\boxtimes(-\mathbf{z}^l)')$, $\boxtimes$ and $\boxtimes$ denote fuzzy composite operation $(\wedge, +)$ and $(\vee, +)$, respectively. If $X=Y$, Z is a kernel for X.

An additional development that uses the notion of dual kernels to enhance the error-correction capability of binary auto-associative morphological memories has been introduced [10]. For the small-dimension toy problem, the new MAM models exhibit better error-correction capabilities than M_{XX} and W_{XX} and a reduced number of spurious memories which can be easily described in terms of the fundamental memories.

It is easy to see that both Ritter and Sussner managed to design kernel from patterns themselves. However, the above kernel methods are problem-dependent. That's to say, the kernels varies with the change of Boolean input patterns set and thus their determination is generally difficult and also experience-dependent. And at present the recording or encoding strategies based on such kernels are not designed yet in order to deal with arbitrary noise. Therefore, their practicality is still not ideal enough.

Thus, our task is to develop a new morphological associative recall memory, which is robust in the presence of random noise (i.e., both dilative and erosive random noise) and help to distinguish among a number of people represented by examples. According to the pattern recognition theory, first of all, we should find an appropriate representation of objects. One of the possible representations is based on similarity or dissimilarity relations between objects. When properly defined, it might be advantageous for solving class identification problems.

The widely-used kernel [11]–[13] can be thought of as a nonlinear similarity measure that corresponds to the dot product in the associated feature space. In this paper, we will introduce a kind of kernel called empirical kernel map [14] into auto-FMAM and develop an Enhanced FMAM (EFMAM). Due to achieving double noise tolerance without destroying the original associative memory structure and encoding strategies, EFMAM not only inherits the good properties of auto-FMAM but also solves the hybrid-noise problem. Experimental results show that EFMAM could achieve far higher accuracy than auto-FMAM under the same hybrid-noise level.

It should be mentioned that the fundamental idea of constructing EFMAM can be easily and straightforwardly applied to MAM. Thus, we only focus on the EFMAM in this paper.

The rest of this paper is organized as follows. A brief introduction to FMAM and its disadvantages are given in Section II. Our EFMAM is detailed in Section III. Section IV presents the experimental results of the proposed method on different datasets and comparisons of EFMAM and auto-FMAM. It is found that in case of noise, the EFMAM correctly recognizes 95% of the thumbnail-sized face images scaled from ORL face database [15] with 40 classes. The conclusions are given in Section V.

II. FMAM And Its Disadvantages

For completeness of presentation, we firstly introduce our FMAM in this section. It should be noted that although we only give the properties of auto-FMAMs, the following theorems can be easily generalized to hetero-FMAMs.

The basic computations used in FMAM are based on the algebraic lattice structure $(R_+, \vee, \wedge, \cdot)$ $(R_+ = (0,\infty))$. If the input vector $\boldsymbol{x}^l=(x_1^l,\cdots,x_m^l)'$ is defined in R^m, by using some transformation, for example, $\exp(x)$ (acting on each component of $\boldsymbol{x}$), the input vectors can be transformed into R_+^m. Therefore, the output vectors $\boldsymbol{x}^l$ must be in R_+^m. Set $X=(\boldsymbol{x}^1,\ldots,\boldsymbol{x}^k)$, with each pair of (X, X), the two new morphological $m\times m$ memories A_{XX} and B_{XX} are as follows:

$$A_{XX}=\bigwedge_{l=1}^{k}(\boldsymbol{x}^l\boxtimes(\boldsymbol{x}^l)^{-1}) \tag{1}$$

$$B_{XX}=\bigvee_{l=1}^{k}(\boldsymbol{x}^l\boxtimes(\boldsymbol{x}^l)^{-1}) \tag{2}$$

where

$$(\boldsymbol{x}^l)^{-1}=(1/x_1^l,\cdots,1/x_n^l)', x_i^l>0,\forall i \tag{3}$$

$$\boldsymbol{x}^l\boxtimes(\boldsymbol{x}^l)^{-1}=\boldsymbol{x}^l\boxtimes(\boldsymbol{x}^l)^{-1}=\begin{pmatrix} x_1^l/x_1^l & \cdots & x_1^l/x_m^l \\ \vdots & \ddots & \vdots \\ x_m^l/x_1^l & \cdots & x_m^l/x_m^l \end{pmatrix} \tag{4}$$

where $\boxtimes$ and $\boxtimes$ denote fuzzy composite operation $(\wedge, \cdot)$ and $(\vee, \cdot)$ often used in fuzzy set theory, respectively.

Definition 1. A matrix $W=(w_{ij})_{m\times m}$ is called a $\boxtimes$–perfect memory for (X, X) if and only if $W\boxtimes X=X$, and is called a $\boxtimes$–perfect memory for (X, X) if and only if $W\boxtimes X=X$.

Theorem 1. For any X, defined as above, auto-FMAM always exists, and A_{XX} is the least upper bound of all $\boxtimes$–perfect recall memories and B_{XX} is the greatest lower bound of all $\boxtimes$–perfect recall memories for (X,X). The proof is presented in Appendix.

This theorem shows that A_{XX} is the least upper bound of all $\boxtimes$–perfect recall memories and B_{XX} is the greatest lower bound of all $\boxtimes$–perfect recall memories for (X, X). Moreover, it should be pointed out that there is no restriction on the number of input patterns. This means that A_{XX} and B_{XX} can perfectly recall any number of stored input patterns, i.e. A_{XX} and B_{XX} has unlimited storage capacity. Furthermore, theorem 2 will show that A_{XX} and B_{XX} can converge in one step for any input pattern z.

Theorem 2. If $A_{XX} \boxtimes \boldsymbol{z} = \boldsymbol{w}$ and $B_{XX} \boxtimes \boldsymbol{z} = \boldsymbol{u}$, then $A_{XX} \boxtimes \boldsymbol{w} = \boldsymbol{w}$ and $B_{XX} \boxtimes \boldsymbol{u} = \boldsymbol{u}$. The proof is presented in Appendix.

Definition 2. If $\bar{x}_i^l \le x_i^l, \forall i = 1, \cdots, m$, then $\bar{\boldsymbol{x}}^{\boldsymbol{l}}$ is called the erosive version of $\boldsymbol{x}^{\boldsymbol{l}}$; If $\bar{x}_i^l \ge x_i^l, \forall i = 1, \cdots, m$, then $\bar{\boldsymbol{x}}^{\boldsymbol{l}}$ is called the dilative version of $\boldsymbol{x}^{\boldsymbol{l}}$.

The following two theorems show that A_{XX} and B_{XX} can tolerate single dilative and single erosive noise well.

Theorem 3. Let $X=(\boldsymbol{x}^1, \cdots, \boldsymbol{x}^k)$, suppose $\bar{\boldsymbol{x}}^{\boldsymbol{l}}$ denote a noisy version of the input pattern $\boldsymbol{x}^{\boldsymbol{l}}$, then $A_{XX} \boxtimes \bar{\boldsymbol{x}}^{\boldsymbol{l}} = \boldsymbol{x}^{\boldsymbol{l}}$ if and only if

$$\bar{x}_j^l \le x_j^l \vee \bigwedge_{i=1}^{m} [\bigvee_{p \ne l} ((x_i^l / x_i^p) x_j^p)] \tag{5}$$

and for each row index $i \in \{1, \cdots, m\}$ there exists a column index $j_i \in \{1, \cdots, m\}$ such that

$$\bar{x}_{j_i}^l = x_{j_i}^l \vee [\bigvee_{p \ne l} ((x_i^l / x_i^p) x_{j_i}^p)] \tag{6}$$

The proof is presented in Appendix.

Theorem 4. Let $X=(\boldsymbol{x}^1, \cdots, \boldsymbol{x}^k)$, suppose $\bar{\boldsymbol{x}}^{\boldsymbol{l}}$ denote a noisy version of the input pattern $\boldsymbol{x}^{\boldsymbol{l}}$, then $B_{XX} \boxtimes \bar{\boldsymbol{x}}^{\boldsymbol{l}} = \boldsymbol{x}^{\boldsymbol{l}}$ if and only if

$$\bar{x}_j^l \ge x_j^l \wedge \bigvee_{i=1}^{m} [\bigwedge_{p \ne l} ((x_i^l / x_i^p) x_j^p)] \tag{7}$$

and for each row index $i \in \{1, \cdots, m\}$ there exists a column index $j_i \in \{1, \cdots, m\}$ such that

$$\bar{x}_{j_i}^l = x_{j_i}^l \wedge [\bigwedge_{p \ne l} ((x_i^l / x_i^p) x_{j_i}^p)] \tag{8}$$

The proof is presented in Appendix.

In sum, the auto-FMAMs encoded respectively with A_{XX} and B_{XX} have unlimited storage capacity, one-shot recall speed and good noise tolerance.

However, there are still some application-related problems with the auto-FMAMs A_{XX} and B_{XX}. Among them, the most prominent is the incapability of A_{XX} to dilative noise while B_{XX} to erosive noise. Theorem 3 and 4 state that if $\bar{x}_i^l > x_i^l, \forall i = 1, \cdots, m$, then, in general, the conclusion $(A_{XX} \boxtimes \bar{\boldsymbol{x}}^{\boldsymbol{l}})_{\boldsymbol{i}} = x_i^l$ does not hold, and while if $\bar{x}_i^l < x_i^l, \forall i = 1, \cdots, m$, then we have a high probability that $(B_{XX} \boxtimes \bar{\boldsymbol{x}}^{\boldsymbol{l}})_{\boldsymbol{i}} \ne x_i^l$. Hence, dilative noise destroys perfect recall for A_{XX} while erosive noise destroys perfect recall for B_{XX}.

In fact, for pattern recognition problem, there is no need for us to get the perfect recall. Observing the process of human recognition and categorization, human beings usually take the nearest similarity as the recognition result. That is to say, the crucial factor is similarities, which can be considered as an inherent connection between perception and higher-level knowledge [16]–[18]. Thus, in the next section, we will introduce a nonlinear similarity measure called empirical kernel into FMAM and induce an enhanced FMAM.

III. Enhaced Fmam Based On Empirical Kernel Map

A. Empirical Kernel Map

There are many ways to discriminate between objects represented by (dis)similarities. A traditional one is the nearest neighbor method (NN) [19]. However, it suffers from a number of limitations, e.g., high computational complexity, a potential loss of accuracy when a small set of prototypes is used, and sensitivity to noise. The novelty of EFMAM lies in interpreting kernel here as a representation of a similarity space.

To construct EFMAM on similarities, the training set T of size n (having n objects) will be used. T is a set of prototypes covering all classes present. In the learning process, the recognizer is built on the $n \times n$ similarity matrix $K(T, T)$, relating all training objects to themselves. The information on a set S of s test objects is provided in terms of their similarities to T, i.e., as an $s \times n$ matrix $K(S,$

T). In particular, *K*(*S*, *T*) is treated as a description of a space where each dimension corresponds to the similarities to a prototype.

In general, *K*(***x***, *T*) defines a vector consisting of *n* similarities found between the object ***x*** and all the objects in the training set *T*, i.e., if *T*={$\boldsymbol{p}_1,\ldots, \boldsymbol{p}_n$} then *K*(***x***, *T*)=[*K*(***x***, $\boldsymbol{p}_1$),…, *K*(***x***, $\boldsymbol{p}_n$)]^t, where *t* denotes a transpose for matrix or vector. Therefore, $K(\cdot, T)$ is seen as a mapping onto an *n*-dimensional similarity space. In this convention, neither ***x*** nor *T* refers to points in a feature space. Instead they refer to the objects themselves. Such a feature map sometimes called the empirical kernel map [20].

B. EFMAM

The architecture of EFMAM is shown in Fig. 1. EFMAM includes an empirical kernel map and an auto-FMAM. Since there is auto-FMAM here, as many associations as desired can be encoded into the memory; one step convergence and perfect recall of Boolean and gray-level noisy patterns using the idea of kernels will also be settled.

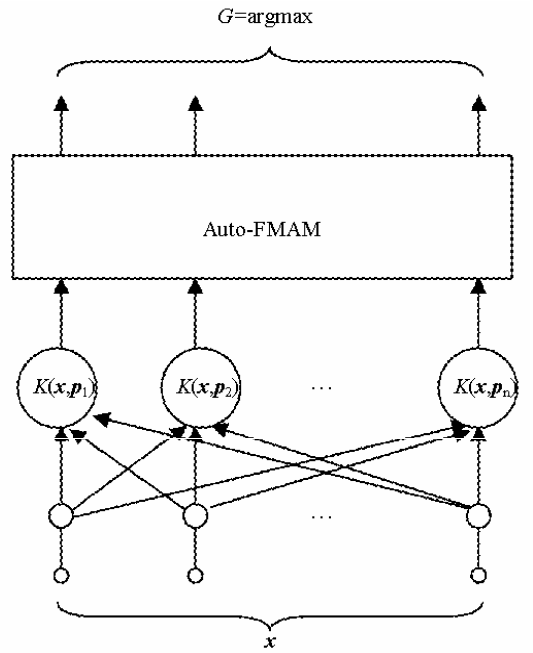

Fig. 1. The architecture of EFMAM.

The input to the recognizer is the vector ***x*** and each output of the *n* empirical kernel map nodes represents the similarity between the object ***x*** and a certain object $\boldsymbol{p}_r$ in the training set *T*, for all r=1, 2,…, *n*. Since we feed into auto-FMAM the *n*-dimension similarity vector between the object ***x*** and all the objects in the training set *T*, in the testing stage, we also get the similarity vector from auto-FMAM. Thus, a max arbitrator is defined after auto-FMAM.

The training of EFMAM is a two-stage process. The first stage is trained using an empirical kernel map and the second stage is then trained using the auto-FMAM algorithm.

It is important to remark that the similarity measure is optional. In this paper, the similarity measure between two vectors ***x*** and ***y*** is defined as follows:

$$K(\boldsymbol{x}, \boldsymbol{y}) = e^{-\beta D(\boldsymbol{x}, \boldsymbol{y})}$$

where $D(\boldsymbol{x}, \boldsymbol{y}) = (\sum_{i=1}^{m} \left| x_i^c - y_i^c \right|^{\boldsymbol{a}})^{\boldsymbol{b}}$, *m* is the dimension of vectors ***x***, ***y***. In our case, β=*a*=*b*=*c*=1.

EFMAM has at least six advantages. Firstly, besides the auto-FMAM, empirical kernel map itself is robust to noise. Thus, EFMAM possesses double noise-tolerance: one is due to the empirical kernel map itself and the other due to the auto-FMAM itself. So, EFMAM is expected to work better than auto-FMAM under the hybrid-noise cases. Secondly, due to the introduction of the kernels, EFMAM can deal with great-dimension pattern problem. If the training set is small, it has the third advantage that only a small set of similarities has to be computed for the following auto-FMAM, while it may still keep the original accuracy. Fourthly, in this paper, since we take exponent-type similarity measure, there is no need for us to linearly transform the input vectors into (0,∞) like auto-FMAM. Thus, the EFMAM memories do not restrict the domain of the input vectors in any way. The experiments in the next section also demonstrate the equivalence between EFMAM A_{XX} and B_{XX}. It becomes the fifth advantage, especially benefit for the real application. Because, in fact, we are not given in priori which type of noises corrupts the unknown images. Therefore, what only we can do is to try both auto-FMAM A_{XX} and B_{XX} then select the better one, but for EFMAM we need to

compute only one memory. The last advantage is that we can more easily incorporate inherent relationships among patterns into the empirical kernels to facilitate problem solving [20]. All those above make EFMAM more suitable for application problem than MAM and FMAM.

IV. EXPERIMENTS

A number of experiments are conducted to compare the performance of the EFMAM and FMAM. They are designed to observe and analyze the behavior of these recognizers in relation to different kinds and levels of noise.

A. Toy Dataset – Characters

The patterns represented the entire alphabet of capital and lower case letters together with some Chinese characters. Some of them are shown in Fig. 2.

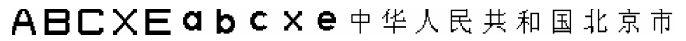

Fig. 2. The twenty patterns used in constructing the EFMAM and the FMAM.

Each pattern image is an 18×18 pixels Boolean image. Concatenating the columns of the image together, using the intensity of each pixel as a single item, each pattern image $\boldsymbol{p}^{\varepsilon}$ can be converted into a pattern vector $\boldsymbol{x}^{\varepsilon} = (x_1^{\varepsilon}, \cdots, x_{324}^{\varepsilon})$ by defining

$$x_{i+18(j-1)}^{\varepsilon} = \begin{cases} 1, \boldsymbol{p}^{\varepsilon}(i,j) = 255 (= white) \\ 0, \boldsymbol{p}^{\varepsilon}(i,j) = 0 (= black) \end{cases}$$

Using the FMAM A_{XX} and B_{XX} for these patterns does not result in perfect recall even if a hybrid-noise pattern corrupted with a little probability is presented to the memory. The second and third rows of Fig. 3 respectively show the corresponding outputs of A_{XX} and B_{XX} when presented with the images in the top row. Obviously, the patterns "c", "X" and "和" all converged to the configurations not represented by any of those patterns in Fig. 2, especially for the third row. The reason for this is the extreme sensitivity of A_{XX} and B_{XX} to hybrid-noise.

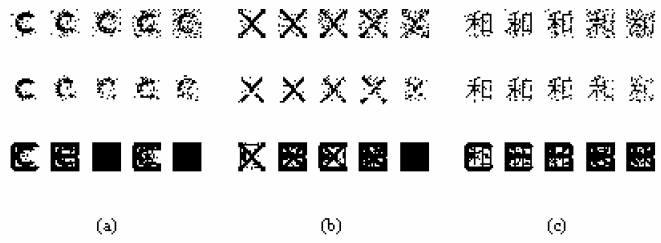

Fig. 3. The top row is the hybrid-noise version of the input patterns under erosive and dilative noise (from left to right, the noise is more and more heavy), the second row corresponds to output patterns of A_{XX} and the bottom row to output patterns of B_{XX}.

The recognition accuracies for the test patterns with different levels of salt & pepper noise are shown in Table I.

It is obvious that either A_{XX} or B_{XX} of EFMAM can distinguish patterns under different levels of hybrid-noise very well. In contrast, the performance of FMAM decreases greatly with the noise increasing.

Table I. The recognition accuracies for the test patterns with different levels of salt & pepper noise (Resolution 18×18)

Level	EFMAM		FMAM	
(%)	A_{XX}	B_{XX}	A_{XX}	B_{XX}
5	1.00	1.00	1.00	0.40
10	1.00	1.00	0.95	0.20
15	1.00	1.00	0.90	0.10
20	1.00	1.00	0.85	0.10
25	1.00	1.00	0.85	0.10
30	1.00	1.00	0.80	0.05
35	1.00	1.00	0.80	0.05
40	1.00	1.00	0.80	0.05
45	1.00	1.00	0.75	0.05
50	1.00	1.00	0.70	0.05

B. Real Dataset – ORL

In order to assess the application performance of the proposed method, experiments were carried using the ORL database, which includes 40 distinct subjects and ten different images per subject. These subjects are of different gender, age and race. For some of the subjects, the images were taken at different times. There are variations in facial expression (open/closed eyes, smiling/nonsmiling) and facial details (glasses/no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright frontal position, with tolerance for some tilting and rotation of up to about 20 degrees. There is some variation in scale of up to about 10%. Examples of images for two subjects are shown in Fig. 4.

Fig. 4. Sample images obtained from ORL database.

Here our main concern is to recognize thumbnail-sized face image, which takes less storage memory and recognition time. These excellent properties lead to its widely application such as display and retrieval on mobile device. But the original images are grayscale with a resolution of 112×92. So we scaled the 112×92 size images to 14×11 and 28×23, respectively. For the training and testing of the recognizer, the grayscale was transformed to lie within [0, 1] by being divided by 255.

All experiments were performed using 5/4/3 training images and 5/6/7 test images per person for a total of 200/160/120 training images and 200/240/280 test images. There was no overlap between the training and test sets. Since the recognition performance will be affected by the selection of training images, the reported results were obtained by training 20 times for each recognizer with different training examples (random selection of 5/4/3 images from ten per each subject) and selecting the average error over all the results. The recognition accuracies for the test images under 14×11 and 28×23 resolutions without noise are shown in Table II.

Table II. The recognition accuracies for the test images without noise (Resolutions 14×11 and 28×23)

Resolution	14×11			28×23		
Number of test images	200	240	280	200	240	280
EFMAM A_{XX}	0.95	0.93	0.89	0.96	0.94	0.90
EFMAM B_{XX}	0.95	0.93	0.89	0.96	0.94	0.90
FMAM A_{XX}	0.94	0.92	0.89	0.95	0.93	0.89
FMAM B_{XX}	0.71	0.64	0.56	0.41	0.34	0.28

Obviously, on the whole, the performance of EFMAM is better than that of FMAM under no noise.

The above experiment without involving noise addition is just a typical and conventional one on ORL database. However, for associative memory models, such experiments are not sufficient. What is more important is to examine their error-correction

capability and compare the recognition performance of the EFMAM proposed in this paper with that of FMAM under different types of random noises, especially the case of both dilative and erosive noises added. In the following experiments we firstly corrupted any of the test images from ORL database with 10%, 20% and 30% randomly generated noise, respectively, which exhibits a major difference to the conventional experiment without regard to noise addition. There are three types of noises: salt noise (i.e., patterns pixels turned white randomly, pixel by pixel), pepper noise (i.e., patterns pixels turned black randomly, pixel by pixel) and salt & pepper noise. One test image and its corresponding three different noisy versions are shown in Fig. 5.

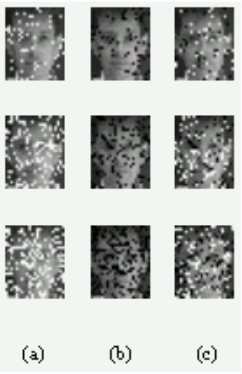

Fig. 5. Three kinds of noisy images of one test image obtained from ORL database: (a) salt noise (b) pepper noise (c) salt & pepper noise added. The images of each row were obtained from the origin with the same corrupted probability. Three probabilities are 0.1, 0.2 and 0.3 from top to bottom.

The recognition accuracies for the test images under 14×11 and 28×23 resolutions with different noises are shown in Tables III-V, where either results of EFMAM A_{XX} and B_{XX} is labeled as EFMAM. The reason why we can make such a label is that the recognition accuracies are equivalent for EFMAM A_{XX} and B_{XX}.

Table III. The recognition accuracies for the test images with different levels of salt noise (Resolutions 14×11 and 28×23, Number of test images are 200, 240 and 280, respectively)

Resolution	Level (%)	EFMAM			FMAM					
					200		240		280	
		200	240	280	A_{XX}	B_{XX}	A_{XX}	B_{XX}	A_{XX}	B_{XX}
14×11	10	0.94	0.91	0.87	0.59	0.72	0.50	0.66	0.38	0.58
	20	0.90	0.88	0.84	0.34	0.70	0.27	0.65	0.18	0.57
	30	0.81	0.79	0.73	0.19	0.70	0.17	0.65	0.11	0.58
28×23	10	0.96	0.93	0.90	0.28	0.43	0.20	0.38	0.13	0.27
	20	0.93	0.90	0.86	0.14	0.45	0.11	0.38	0.08	0.30
	30	0.88	0.86	0.79	0.10	0.46	0.09	0.40	0.08	0.32

Table IV. The recognition accuracies for the test images with different levels of pepper noise (Resolutions 14×11 and 28×23, Number of test images are 200, 240 and 280, respectively)

Resolution	Level (%)	EFMAM			FMAM					
					200		240		280	
		200	240	280	A_{XX}	B_{XX}	A_{XX}	B_{XX}	A_{XX}	B_{XX}
14×11	10	0.94	0.92	0.89	0.94	0.02	0.92	0.02	0.89	0.02
	20	0.91	0.89	0.84	0.93	0.02	0.89	0.02	0.86	0.02
	30	0.83	0.80	0.75	0.87	0.02	0.83	0.02	0.79	0.02
28×23	10	0.96	0.94	0.91	0.96	0.02	0.93	0.02	0.88	0.02
	20	0.95	0.91	0.88	0.94	0.02	0.92	0.02	0.88	0.02
	30	0.90	0.88	0.83	0.93	0.02	0.90	0.02	0.86	0.02

Table V. The recognition accuracies for the test images with different levels of salt & pepper noise (Resolutions 14×11 and 28×23, Number of test images are 200, 240 and 280, respectively)

Resolution	Level (%)	EFMAM			FMAM					
		200	240	280	200		240		280	
					A_{XX}	B_{XX}	A_{XX}	B_{XX}	A_{XX}	B_{XX}
14×11	10	0.95	0.92	0.89	0.76	0.02	0.68	0.02	0.57	0.02
	20	0.93	0.91	0.87	0.55	0.02	0.49	0.02	0.34	0.02
	30	0.91	0.89	0.85	0.40	0.02	0.32	0.02	0.21	0.02
28×23	10	0.96	0.94	0.90	0.50	0.02	0.37	0.02	0.24	0.02
	20	0.96	0.93	0.90	0.26	0.02	0.18	0.02	0.10	0.02
	30	0.95	0.93	0.89	0.14	0.02	0.10	0.02	0.08	0.02

Although Table IV reveals that EFMAM is, in recognition performance, equivalent or even slightly inferior to FMAM A_{XX} under the same condition for pepper noise case, Table III and V demonstrate that EFMAM achieves far higher recognition accuracy than FMAM for salt noise and salt & pepper noise cases. For example, the average accuracies of EFMAM are 92%, 90% and 88% with the hybrid-noises of 10%, 20% and 30% respectively, far higher than FMAM (67%, 46%, 31%) under the same noise cases, which indicates that EFMAM is quite suitable for recognizing so-called thumbnail face images. For the same recognizer, the recognition accuracy becomes smaller and smaller as the noise increases. But such influence on recognition performance of EFMAM is less than that of FMAM. In other words, EFMAM is much more robust to noise, especially the hybrid noise. Comparing the experimental results of the two resolutions, we can also easily find that the greater the resolution, the better the recognition accuracies of EFMAM. Hence, it is reasonably expected that EFMAM can also work well for the original resolution of ORL database (112×92).

V. Conclusions

The basic goal of the artificial associative memories is the retrieval of complete stored patterns from noisy or incomplete input pattern keys. However, the existing morphological associative memory models, including MAM and our previous FMAM, have a common disadvantage: the extreme vulnerability to hybrid-noise. In this paper, we propose an enhanced FMAM based on empirical kernel map. The fundamental idea of EFMAM is to introduce an empirical kernel map into FMAM to achieve double noise-tolerance. Although the improvement trick itself is simple, experiments reveal that EFMAM does increase FMAM on accuracy greatly without sacrificing its structure simplicity and computational effort much. Therefore, we believe that EFMAM is an alternative solution to hybrid-noise problems.

Appendix

Proof of Theorem 1

According to Definition 1, W is $\otimes$–perfect for (X, X) if and only if $W \otimes \boldsymbol{x}^l = \boldsymbol{x}^l$, $l = 1,2,\ldots, k$. Similarly, W is $\overline{\otimes}$–perfect for (X, X), then

$$(W \otimes \boldsymbol{x}^l)_i = x_i^l, i = 1,2,\cdots,m; l = 1,2,\cdots,k \tag{9}$$

$$\text{i.e. } \bigvee_{j=1}^{m}(w_{ij} \cdot x_j^l) = x_i^l, i = 1,2,\cdots,m; l = 1,2,\cdots,k \tag{10}$$

If W is $\overline{\otimes}$–perfect for (X, X), then

$$(W \overline{\otimes} \boldsymbol{x}^l)_i = x_i^l, i = 1,2,\cdots,m; l = 1,2,\cdots,k \tag{11}$$

$$\text{i.e. } \bigwedge_{j=1}^{m}(w_{ij} \cdot x_j^l) = x_i^l, i = 1,2,\cdots,m; l = 1,2,\cdots,k \tag{12}$$

From (1)–(4), it follows that

$$A_{XX} \otimes \boldsymbol{x}^l \le \begin{pmatrix} x_1^l / x_1^l & \cdots & x_1^l / x_m^l \\ \vdots & \ddots & \vdots \\ x_m^l / x_1^l & \cdots & x_m^l / x_m^l \end{pmatrix} \otimes \boldsymbol{x}^l = \boldsymbol{x}^l$$

$$= \begin{pmatrix} x_1^l / x_1^l & \cdots & x_1^l / x_m^l \\ \vdots & \ddots & \vdots \\ x_m^l / x_1^l & \cdots & x_m^l / x_m^l \end{pmatrix} \otimes \boldsymbol{x}^{\boldsymbol{l}} \le B_{XX} \otimes \boldsymbol{x}^{\boldsymbol{l}} \quad (13)$$

$l = 1,2,\ldots, k$. Equivalently,

$$A_{XX} \otimes X \le X \le B_{XX} \otimes X \quad (14)$$

Now, let us prove $A_{XX} \otimes X = X, B_{XX} \otimes X = X$. In terms of (10), for any $\otimes$–perfect W for (X, X), we have $w_{ij} \cdot x_j^l \le x_i^l, l = 1,2,\cdots,k; i = 1,2,\cdots,m$

Since x_i^l is always positive, so $w_{ij} \le x_i^l / x_j^l$

i.e. $$w_{ij} \le \bigwedge_{l=1}^{k} (x_i^l / x_j^l) \quad (15)$$

According to (1), the following immediately holds $W \le A_{XX}$ (16)

In terms of (14) and Definition 1, we have $X = W \otimes X \le A_{XX} \otimes X \le X$

Hence, $A_{XX} \otimes X = X$. Similarly, we also have $B_{XX} \otimes X = X, B_{XX} \le V$ (17)

for any $\otimes$–perfect for (X, X).

Up to now, we have proved that A_{XX} is the least upper bound of all $\otimes$–perfect recall memories and B_{XX} is the greatest lower bound of all $\otimes$–perfect recall memories for (X,X). Next, let us prove that auto-FMAM always exists.

$$A_{XX} = \bigwedge_{\boldsymbol{l}=1}^{\boldsymbol{k}} (\boldsymbol{x}^{\boldsymbol{l}} \otimes (\boldsymbol{x}^{\boldsymbol{l}})^{-1})$$

$$\Leftrightarrow ((\boldsymbol{x}^{\boldsymbol{l}} \otimes (\boldsymbol{x}^{\boldsymbol{l}})^{-1}) - A_{XX})_{ij} \ge 0 \text{ ,and } ((\boldsymbol{x}^{\boldsymbol{l}} \otimes (\boldsymbol{x}^{\boldsymbol{l}})^{-1}) - A_{XX})_{ii} = 0$$

$$\Leftrightarrow \bigwedge_{\boldsymbol{j}=1}^{\boldsymbol{m}} ([\boldsymbol{x}^{\boldsymbol{l}} \otimes (\boldsymbol{x}^{\boldsymbol{l}})^{-1}] - A_{XX})_{ij} = 0, i = 1,2,\cdots,m; l = 1,2,\cdots,k$$

$$\Leftrightarrow \bigwedge_{j=1}^{m} (x_i^l / x_j^l - a_{ij}) = 0, i = 1,2,\cdots,m; l = 1,2,\cdots,k$$

$$\Leftrightarrow \bigwedge_{j=1}^{m} (x_i^l - a_{ij} \cdot x_j^l) = 0, i = 1,2,\cdots,m; l = 1,2,\cdots,k$$

$$\Leftrightarrow x_i^l + \bigvee_{j=1}^{m} (-a_{ij} \cdot x_j^l) = 0, i = 1,2,\cdots,m; l = 1,2,\cdots,k$$

$$\Leftrightarrow x_i^l - \bigvee_{j=1}^{m} (a_{ij} \cdot x_j^l) = 0, i = 1,2,\cdots,m; l = 1,2,\cdots,k$$

$$\Leftrightarrow x_i^l - (A_{XX} \otimes \boldsymbol{x}^{\boldsymbol{l}})_{\boldsymbol{i}} = 0, \ i = 1,2,\cdots,m; l = 1,2,\cdots,k$$

$$\Leftrightarrow (A_{XX} \otimes \boldsymbol{x}^{\boldsymbol{l}})_{\boldsymbol{i}} = x_i^l, i = 1,2,\cdots,m; l = 1,2,\cdots,k$$

i.e. A_{XX} is $\otimes$–perfect memory for (X, X).

Similarly, it is also true for B_{XX}.

Proof of Theorem 2

Suppose $A_{XX} \otimes \boldsymbol{z} = \boldsymbol{w}$, since $a_{ii} = 1, i = 1,2,\cdots,n$, $(A_{XX} \otimes \boldsymbol{w})_{\boldsymbol{i}} = \underset{j=1}{\overset{n}{\vee}} a_{ij} \cdot w_j \geq a_{ii} \cdot w_i = w_i$

so, $\boldsymbol{w} \leq A_{XX} \otimes \boldsymbol{w}$ (18)

In terms of the definition of a_{ij} in A_{XX}, we have $a_{ip} = \underset{l=1}{\overset{k}{\wedge}}(x_i^l / x_p^l), a_{pj} = \underset{l=1}{\overset{k}{\wedge}}(x_p^l / x_j^l)$,

thus $a_{ip} \cdot a_{pj} = \underset{l=1}{\overset{k}{\wedge}}(x_i^l / x_p^l) \cdot \underset{l=1}{\overset{k}{\wedge}}(x_p^l / x_j^l)$

$$\leq \underset{l=1}{\overset{k}{\wedge}}(x_i^l / x_p^l \cdot x_p^l / x_j^l)$$

$$= \underset{l=1}{\overset{k}{\wedge}}(x_i^l / x_j^l) = a_{ij} \quad (19)$$

Since $w_i = \underset{j=1}{\overset{n}{\vee}}(a_{ij} \cdot z_j)$, in terms of (19), we immediately have $w_i \geq \underset{j=1}{\overset{n}{\vee}}(a_{ip} \cdot a_{pj} \cdot z_j), p = 1,2,\cdots,n$

Thus, $w_i \geq \underset{p=1}{\overset{n}{\vee}} \underset{j=1}{\overset{n}{\vee}}(a_{ip} \cdot a_{pj} \cdot z_j) = \underset{p=1}{\overset{n}{\vee}}(a_{ip} \cdot \underset{j=1}{\overset{n}{\vee}}(a_{pj} \cdot z_j))$

$$= \underset{p=1}{\overset{n}{\vee}}(a_{ip} \cdot w_p) = (A_{XX} \otimes \boldsymbol{w})_{\boldsymbol{i}}$$

i.e. $A_{XX} \otimes \boldsymbol{w} \leq \boldsymbol{w}$ (20)

According to (18) and (20), we have $A_{XX} \otimes \boldsymbol{w} = \boldsymbol{w}$

Similarly, we can easily prove $B_{XX} \otimes \boldsymbol{u} = \boldsymbol{u}$

Proof of Theorem 3

If $A_{XX} \otimes \bar{\boldsymbol{x}}^{\boldsymbol{l}} = \boldsymbol{x}^{\boldsymbol{l}}$, then

$$x_i^l = (A_{XX} \otimes \bar{\boldsymbol{x}}^{\boldsymbol{l}})_{\boldsymbol{i}} = \underset{j=1}{\overset{m}{\vee}}(a_{ij} \cdot \bar{x}_j^l)$$

$$\geq a_{ij} \cdot \bar{x}_j^l, i, j = 1,\cdots,m \quad (21)$$

Thus, in terms of (1) and (21), we have

$$\bar{x}_j^l \leq \underset{i=1}{\overset{m}{\wedge}}(x_i^l / a_{ij}), j = 1,\cdots,m$$

$$\Leftrightarrow \bar{x}_j^l \leq \underset{i=1}{\overset{m}{\wedge}}(x_i^l / (\underset{p=1}{\overset{k}{\wedge}}(x_i^p / x_j^p))), j = 1,\cdots,m$$

$$\Leftrightarrow \bar{x}_j^l \leq \underset{i=1}{\overset{m}{\wedge}}(x_i^l \vee \underset{p\neq l}{\vee}((x_i^l / x_i^p) \cdot x_j^p))), j = 1,\cdots,m$$

$$\Leftrightarrow \bar{x}_j^l \leq x_j^l \vee \underset{i=1}{\overset{m}{\wedge}}[\underset{p\neq l}{\vee}((x_i^l / x_i^p) x_j^p)],$$

$$j = 1,\cdots,m; p = 1,\cdots,l-1,l+1,\cdots,k$$

This shows that (5) holds. From (5), we immediately have $\bar{x}_j^l \leq x_j^l \vee [\underset{p\neq l}{\vee}((x_i^l / x_i^p) x_j^p)]$ (22)

where $j = 1,\cdots,m, p = 1,\cdots,l-1,l+1,\cdots,k.$ Now, let us assume that there exists a row index $i \in \{1,\cdots,m\}$ such that

$$\bar{x}_j^l < x_j^l \vee [\underset{p\neq l}{\vee}((x_i^l / x_i^p) x_j^p)],$$

$$j = 1,\cdots,m; p = 1,\cdots,l-1,l+1,\cdots k$$

then

$$(A_{XX} \otimes \bar{\boldsymbol{x}}^{\boldsymbol{l}})_{\boldsymbol{i}} = \underset{j=1}{\overset{m}{\vee}}(a_{ij} \cdot \bar{x}_j^l)$$

$$< \underset{j=1}{\overset{m}{\vee}}(a_{ij} \cdot (x_j^l \vee [\underset{p\neq l}{\vee}((x_i^l / x_i^p) x_j^p)]))$$

$$= \underset{j=1}{\overset{m}{\vee}}(a_{ij} \cdot (\underset{p=1}{\overset{k}{\vee}}(x_i^l / x_i^p)x_j^p)) = \underset{j=1}{\overset{m}{\vee}}(x_i^l \cdot (\underset{p=1}{\overset{k}{\vee}}(a_{ij} \cdot x_j^p / x_i^p)))$$

$$= x_i^l \text{ (in terms of (1))}$$

So, $(A_{XX} \otimes \overline{\boldsymbol{x}}^{\boldsymbol{l}})_{\boldsymbol{i}} < \boldsymbol{x}_{\boldsymbol{i}}^{\boldsymbol{l}}$ which contradicts the hypothesis that $A_{XX} \otimes \boldsymbol{x}^{\boldsymbol{l}} = \boldsymbol{x}^{\boldsymbol{l}}$.

In other words, (6) follows.

In terms of the above, it is easily to see $\overline{x}_j^l \le x_j^l \vee \underset{i=1}{\overset{m}{\wedge}}[\underset{p \ne l}{\vee}((x_i^l / x_i^p)x_j^p)]$ holds if and only if

$$x_i^l \ge a_{ij} \cdot \overline{x}_j^l, i, j = 1, \cdots, m,$$

i.e. $x_i^l \ge \underset{j=1}{\overset{n}{\vee}}(a_{ij} \cdot \overline{x}_j^l), i = 1, \cdots, m,$

i.e. $x_i^l \ge (A_{XX} \otimes \overline{\boldsymbol{x}})_{\boldsymbol{i}}, i = 1, \cdots, m,$

i.e. $A_{XX} \otimes \overline{\boldsymbol{x}}^{\boldsymbol{l}} \le \boldsymbol{x}^{\boldsymbol{l}}, l = 1,2, \cdots, k$ (23)

Now, consider

$$(A_{XX} \otimes \overline{\boldsymbol{x}}^{\boldsymbol{l}})_{\boldsymbol{i}}$$

$$= \underset{j=1}{\overset{n}{\vee}}(a_{ij} \cdot \overline{x}_j^l) \ge a_{ij_i} \cdot \overline{x}_{j_i}^l$$

$$= a_{ij_i} \cdot (x_{j_i}^l \vee [\underset{p \ne l}{\vee}((x_i^l / x_i^p)x_{j_i}^p)]) \text{ (in terms of (6))}$$

$$= x_i^l$$

which means that $A_{XX} \otimes \overline{\boldsymbol{x}}^{\boldsymbol{l}} \ge \boldsymbol{x}^{\boldsymbol{l}}$ (24)

In terms of (23) and (24), we immediately have $A_{XX} \otimes \overline{\boldsymbol{x}}^{\boldsymbol{l}} = \boldsymbol{x}^{\boldsymbol{l}}$.Thus, this theorem is proved.

Proof of Theorem 4

If $B_{XX} \otimes \overline{\boldsymbol{x}}^{\boldsymbol{l}} = \boldsymbol{x}^{\boldsymbol{l}}$, then

$$x_i^l = (B_{XX} \otimes \overline{\boldsymbol{x}}^{\boldsymbol{l}})_{\boldsymbol{i}}$$

$$= \underset{j=1}{\overset{m}{\wedge}}(b_{ij} \cdot \overline{x}_j^l) \le b_{ij} \cdot \overline{x}_j^l, i, j = 1, \cdots, m \quad (25)$$

Thus, in terms of (2) and (25), we have

$$\overline{x}_j^l \ge \underset{i=1}{\overset{m}{\vee}}(x_i^l / b_{ij}), j = 1, \cdots, m$$

$$\Leftrightarrow \overline{x}_j^l \ge \underset{i=1}{\overset{m}{\vee}}(x_i^l / (\underset{p=1}{\overset{k}{\vee}}(x_i^p / x_j^p))), j = 1, \cdots, m$$

$$\Leftrightarrow \overline{x}_j^l \ge \underset{i=1}{\overset{m}{\vee}}(x_i^l \wedge \underset{p \ne l}{\wedge}((x_i^l / x_i^p) \cdot x_j^p))), j = 1, \cdots, m$$

$$\Leftrightarrow \overline{x}_j^l \ge x_j^l \wedge \underset{i=1}{\overset{m}{\vee}}[\underset{p \ne l}{\wedge}((x_i^l / x_i^p)x_j^p)],$$

$$j = 1, \cdots, m; p = 1, \cdots, l-1, l+1, \cdots, k$$

This shows that (7) holds. From (7), we immediately have $\overline{x}_j^l \ge x_j^l \wedge [\underset{p \ne l}{\wedge}((x_i^l / x_i^p)x_j^p)]$ (26)

where $j = 1, \cdots, m, p = 1, \cdots, l-1, l+1, \cdots, k.$ Now, let us assume that there exists a row index $i \in \{1, \cdots, m\}$ such that

$$\overline{x}_j^l > x_j^l \wedge [\underset{p \ne l}{\wedge}((x_i^l / x_i^p)x_j^p)],$$

$$j = 1, \cdots, m; p = 1, \cdots, l-1, l+1, \cdots k$$

then

$$(B_{XX} \otimes \bar{\boldsymbol{x}}^{\boldsymbol{l}})_{\boldsymbol{i}} = \underset{j=1}{\overset{m}{\wedge}}(b_{ij} \cdot \bar{x}_j^l)$$

$$> \underset{j=1}{\overset{m}{\wedge}}(b_{ij} \cdot (x_j^l \wedge [\underset{p \neq l}{\wedge}((x_i^l / x_i^p)x_j^p)]))$$

$$> \underset{j=1}{\overset{m}{\wedge}}(b_{ij} \cdot (x_j^l \wedge [\underset{p \neq l}{\wedge}((x_i^l / x_i^p)x_j^p)]))$$

$$= \underset{j=1}{\overset{m}{\wedge}}(b_{ij} \cdot (\underset{p=1}{\overset{k}{\wedge}}(x_i^l / x_i^p)x_j^p)) = \underset{j=1}{\overset{m}{\wedge}}(x_i^l \cdot (\underset{p=1}{\overset{k}{\wedge}}(b_{ij} \cdot x_j^p / x_i^p)))$$

$$= x_i^l \quad \text{(in terms of (2))}$$

So, $(B_{XX} \otimes \bar{\boldsymbol{x}}^{\boldsymbol{l}})_{\boldsymbol{i}} > x_i^l$ which contradicts the hypothesis that $B_{XX} \otimes \boldsymbol{x}^{\boldsymbol{l}} = \boldsymbol{x}^{\boldsymbol{l}}$.

In other words, (8) follows.

In terms of the above, it is easily to see $\bar{x}_j^l \geq x_j^l \wedge \underset{i=1}{\overset{m}{\vee}}[\underset{p \neq l}{\wedge}((x_i^l / x_i^p)x_j^p)]$ holds if and only if

$$x_i^l \leq b_{ij} \cdot \bar{x}_j^l, i, j = 1, \cdots, m,$$

i.e. $x_i^l \leq \underset{j=1}{\overset{n}{\wedge}}(b_{ij} \cdot \bar{x}_j^l), i = 1, \cdots, m,$

i.e. $x_i^l \leq (B_{XX} \otimes \bar{\boldsymbol{x}})_{\boldsymbol{i}}, i = 1, \cdots, m,$

i.e. $B_{XX} \otimes \bar{\boldsymbol{x}}^{\boldsymbol{l}} \leq \boldsymbol{x}^{\boldsymbol{l}}, l = 1,2, \cdots, k$ (27)

Now, consider

$$(B_{XX} \otimes \bar{\boldsymbol{x}}^{\boldsymbol{l}})_{\boldsymbol{i}}$$

$$= \underset{j=1}{\overset{n}{\wedge}}(b_{ij} \cdot \bar{x}_j^l) \leq b_{ij_i} \cdot \bar{x}_{j_i}^l$$

$$= b_{ij_i} \cdot (x_{j_i}^l \wedge [\underset{p \neq l}{\wedge}((x_i^l / x_i^p)x_{j_i}^p)]) \text{ (in terms of (8))}$$

$$= x_i^l$$

which means that $B_{XX} \otimes \bar{\boldsymbol{x}}^{\boldsymbol{l}} \leq \boldsymbol{x}^{\boldsymbol{l}}$. (28)

In terms of (27) and (28), we immediately have $B_{XX} \otimes \bar{\boldsymbol{x}}^{\boldsymbol{l}} = \boldsymbol{x}^{\boldsymbol{l}}$.Thus, this theorem is proved.

REFERENCES

[1] P. D. Gader, M. A. Khabou and A. Koldobsky, "Morphological regularization neural networks," *Pattern Recognition*, vol. 33, pp. 935–944, June 2000.

[2] L. F. C. Pessoa and P. Maragos, "Neural networks with hybrid morphological/rank/linear nodes: a unifying framework with applications to handwritten character recognition," *Pattern Recognition*, vol. 33, pp. 945–960, June 2000.

[3] B. Raducanu, M. Grana and F. X. Albizuri, "Morphological scale spaces and associative memories: results on robustness and practical applications," *Journal of Mathematical Imaging and Vision*, vol. 19, pp. 113–131, Sept. 2003.

[4] G. X. Ritter, P. Sussner and J. L. Diaz-de-Leon, "Morphological associative memories," *IEEE Trans. Neural Networks*, vol. 9, pp. 281–293, March 1998.

[5] G. X. Ritter and P. Sussner, "An introduction to morphological neural networks," in *Proc. Int. Conf. Pattern Recognition*, 1996, vol. 4, pp. 709–717.

[6] P. Sussner, "Generalizing operations of binary morphological autoassociative memories using fuzzy set theory," *Journal of Mathematical Imaging and Vision*, vol. 19, pp. 81–93, Sept. 2003.

[7] Y. Won, P. D. Gader and P. C. Coffield, "Morphological shared-weight neural network with applications to automatic target detection," *IEEE Trans. Neural Networks*, vol. 8, pp. 1195–1203, Sept. 1997.

[8] M. Wang, S.T. Wang and X.J. Wu, "Initial results on fuzzy morphological associative memories," *Acta Electronica Sinica* (in Chinese), vol. 31, pp. 690–693, May 2003.

[9] G. X. Ritter and G. Urcid, "Lattice algebra approach to single-neuron computation," *IEEE Trans. Neural Networks*, vol. 14, pp. 282–295, March 2003.

[10] P. Sussner, "Binary autoassociative morphological memories based on the kernel and dual kernel methods," in *Proc. Int. Joint Conf. Neural Networks*, 2003, vol. 1, pp. 236–241.

[11] J. Lu, K. N. Plataniotis and A. N. Venetsanopoulos, "Face recognition using kernel direct discriminant analysis algorithms," *IEEE Trans. Neural Networks*, vol. 14, pp. 117–126, Jan. 2003.

[12] M. M. Van Hulle, "Entropy-based kernel mixture modeling for topographic map formation," *IEEE Trans. Neural Networks*, vol. 15, pp. 850–858, July 2004.

[13] R. A. Morejon and J. C. Principe, "Advanced search algorithms for information-theoretic learning with kernel-based estimators," *IEEE Trans. Neural Networks*, vol. 15, pp. 874–884, July 2004.

[14] B. Scholkopf, J. Weston, E. Eskin, C. Leslie and W. S. Noble, "A kernel approach for learning from almost orthogonal patterns," *ECML*, *LNAI* 2430, pp. 511–528, 2002.

[15] AT&T Laboratories, Cambridge. Available: http://www.cam-orl.co.uk/facedatabase.html.
[16] R. L. Goldstone, "Similarity," in R.A.Wilson, F.C.Keil Eds. MIT Encyclopedia of the Cognitive Sciences, Cambridge, MA: MIT Press, 1999.
[17] S. Edelman, *Representation and recognition in vision*. Cambridge: MIT Press, 1999.
[18] C. M. Wharton, K. J. Holyoak, P. E. Downing, T. E. Lange and T. D. Wickens, "The story with reminding: memory retrieval is influenced by analogical similarity," in *Proc. Annual Conf. Cognitive Science Society*, 1992, pp. 588–593.
[19] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inform. Theory*, vol. 13, pp. 21–27, Jan. 1967.
[20] B. Scholkopf and A. J. Smola, *Learning with kernels*, Cambridge, MA: MIT Press, 2002.

Min Wang received the M.S. degree in computer science from East China Shipbuilding Institute, Zhenjiang, China, in 2003. She is currently working toward the Ph.D. degree in Dept. of Computer Science & Engineering, Nanjing University of Aeronautics & Astronautics, China. Her research interests are mainly in neural computing, pattern recognition and image processing.

Songcan Chen received the M.S. degree in computer science from Shanghai Jiaotong University in 1985, and Ph.D degree in electronics engineering from Nanjing University of Aeronautics and Astronautics (NUAA) in 1997. Currently, he is a professor of Pattern Recognition and Artificial Intelligence in Department of Computer Science and Engineering of NUAA. His research interests include pattern recognition, bioinformatics, medical image processing and neural networks. He has published more than 50 journal papers on IEEE Trans. SMC-B, Pattern Recognition, Pattern Recognition Letters, Artificial Intelligence in Medicine, Neural Processing Letters and Applied Mathematics & Computation etc.