Sparsely Connected Autoassociative Fuzzy Implicative Memories for the Reconstruction of Large Gray-Scale Images

Marcos Eduardo Valle
Department of Mathematics, University of Londrina
Londrina, Paraná, Brazil
Email: valle@uel.br

Abstract— *Autoassociative fuzzy implicative memories (AFIMs) are models that exhibit optimal absolute storage capacity and an excellent tolerance with respect to incomplete or eroded patterns. Thus, they can be effectively used for the reconstruction of gray-scale images. In practice, however, applications of AFIMs are confined to images of small size due to computational limitations. This paper introduces a class of sparsely connected AFIMs (SCAFIMs) that circumvent this computational overhead and, therefore, can be used for the reconstruction of large images. We show that SCAFIMs exhibit optimal absolute storage capacity and tolerance with respect to incomplete or eroded patterns. Furthermore, we compare the performance of SCAFIMs with their corresponding fully connected AFIM both theoretically and by means of computational experiments.*

Keywords— Fuzzy associative memory, reconstruction of large gray-scale images, sparse encoding, storage capacity, tolerance with respect to noise.

1 Introduction

Associative memories (AMs) are models inspirated in the human brain ability to recall by association [1, 2, 3, 4]. Here, a partial or approximate representation of a stored item is used to recall the full item. An example would be recalling a poem by knowing its first words as an initial clue. Applications of AM models include classification, prediction, control, and pattern recognition [5, 6, 7, 8, 9].

We speak of a *fuzzy associative memory* (FAM) if the AM model is used for the association of fuzzy sets [5, 8, 10, 11]. The *max-min* and *max-product FAMs* of Kosko, the *generalized FAMs* (GFAMs) of Chung and Lee, and the *implicative fuzzy associative memories* (IFAMs) are instances of FAMs [5, 8, 12, 13]. We would like to recall, however, that the FAMs of Kosko as well as the GFAMs usually fail to perfectly recall a memorized item due to crosstalk between the stored patterns [8, 12]. Therefore, despite their successful applications in a variety of problems [5, 12], the FAMs of Kosko and the GFAMs are not recommended for the reconstruction of gray-scale images.

In contrast, IFAMs exhibit *optimal absolute storage capacity* in the autoassociative case [13]. In other words, one can store has many patterns as desired in an *autoassociative fuzzy implicative memory* (AFIM). In addition, AFIMs exhibit one step-convergence and an excellent tolerance with respect incomplete or eroded patterns [8, 13]. As a consequence, AFIMs can be effectively used for the reconstruction of gray-scale images [8, 14].

Due to computational limitations, however, AFIMs can be used only for the storage and recall of images of small size. Specifically, such as several other AMs, including the *complex-valued neural network* [15, 16], the multi-state model of Costantini *et al.* [17, 18], and the *morphological associative memories* [14, 19], AFIMs require the storage of a large amount of synaptic junctions if used to process large images. For example, approximately 512 gigabytes of memory space would be consumed by an AFIM that is used for the storage of gray-scale images of size 512×512. In order to circumvent this computational overhead, we propose to remove a large amount of synaptic junctions of an AFIM. The resulting model is referred to as a *sparsely connected AFIM* (SCAFIM). We show that SCAFIMs inherit the optimal absolute storage capacity and the tolerance with respect to incomplete or eroded patterns of AFIMs. Therefore, they can be used for the reconstruction of large gray-scale images.

The paper is organized as follows. The next section reviews the basic concepts of AFIMs. SCAFIMs are introduced in Section 3. Some theoretical results concerning the storage capacity and noise tolerance of SCAFIMs, as well as two strategies for the construction of a SCAFIM from a (fully connected) AFIM, are also given in Section 3. Section 4 provides computational experiments. The paper finishes with the concluding remarks in Section 5.

2 A Brief Review on Autoassociative Fuzzy Implicative Memories

Associative memories (AMs) are input-output systems that store a set of associations $\{(\mathbf{x}^\xi, \mathbf{y}^\xi) : \xi = 1, \ldots, k\}$, called *fundamental memory set* [2, 3]. Formally, an AM corresponds to a mapping G such that $G(\mathbf{x}^\xi) = \mathbf{y}^\xi$ for every $\xi = 1, \ldots, k$. In addition, an AM model should be endowed with a certain tolerance with respect to noise. In other words, the mapping G should be such that $G(\tilde{\mathbf{x}}^\xi)$ equals $\mathbf{y}^\xi$ even for noisy or incomplete versions $\tilde{\mathbf{x}}^\xi$ of $\mathbf{x}^\xi$.

An associative memory can be classified as *autoassociative* or *heteroassociative* [3]. We have an *autoassociative memory* if $\mathbf{y}^\xi = \mathbf{x}^\xi$ for all $\xi = 1, \ldots, k$. We speak of an heteroassociative memory otherwise. This paper focus on the autoassociative case. The famous *Hopfield network* is an example of autoassociative memory for binary patterns [2, 4, 20].

We speak of a *fuzzy associative memory* (FAM) if the mapping G is given by a fuzzy neural network and the patterns $\mathbf{x}^\xi$ and $\mathbf{y}^\xi$ represent finite fuzzy sets for every $\xi = 1, \ldots, k$ [10, 11]. Applications of FAMs include backing up a truck and trailer, target tracking, and forecasting the average monthly streamflow of a large hydroelectric plant [5, 8]. This paper focus only on the class of *autoassociative fuzzy implicative memories* (AFIMs). The reader interested in a comprehensive survey on FAM models is invited to read [8].

ISBN: 978-989-95079-6-8

2.1 Autoassociative Fuzzy Implicative Memories

AFIMs are single-layer feedforward fuzzy neural networks equipped with neurons that compute the maximum of continuous triangular norms (t-norms) [13, 21]. Specifically, given a synaptic weight matrix $W \in [0,1]^{n\times n}$ and a threshold vector $\boldsymbol{\theta} \in [0,1]^n$, the output $\mathbf{y} \in [0,1]^m$ of an AFIM is given by

$$\mathbf{y} = \mathcal{W}(\mathbf{x}) = (W \circ \mathbf{x}) \vee \boldsymbol{\theta}\,, \tag{1}$$

where $\mathbf{x} \in [0,1]^n$ is the input pattern and the symbol "$\circ$" denotes a max-T product[1]. Recall that the max-T product of two matrices $A \in [0,1]^{n\times k}$ and $B \in [0,1]^{k\times n}$, denoted by $C = A \circ B \in [0,1]^{n\times m}$, is defined as follows [10]:

$$c_{ij} = \bigvee_{\xi=1}^{k} T(a_{i\xi}, b_{\xi j})\,. \tag{2}$$

The synaptic weight matrix W and the threshold vector $\boldsymbol{\theta}$ of an AFIM model are computed by means of a recording recipe called *implicative fuzzy learning* (IFL). Formally, given a fundamental memory set $\{\mathbf{x}^1, \ldots, \mathbf{x}^k\}$, where each $\mathbf{x}^\xi \in [0,1]^n$, IFL defines $W \in [0,1]^{n\times n}$ and $\boldsymbol{\theta} \in [0,1]^n$ as follows [13, 21, 11]:

$$[W \,|\, \boldsymbol{\theta}] = \bigwedge \left\{ [A \,|\, \boldsymbol{\beta}] : (A \circ \mathbf{x}^\xi) \vee \boldsymbol{\beta} \leq \mathbf{x}^\xi,\ \forall\, \xi \in \mathcal{K} \right\}. \tag{3}$$

Here, $[W \,|\, \boldsymbol{\theta}]$ represents the $n \times (n+1)$ fuzzy matrix obtained by concatenating W and $\boldsymbol{\theta}$, and $\mathcal{K} = \{1, \ldots, k\}$.

It is important to note that IFL makes optimal use of the synaptic weights and thresholds of the AM model given by (1). In fact, if there exist $A \in [0,1]^{m\times n}$ and $\boldsymbol{\beta} \in [0,1]^n$ such that $(A \circ \mathbf{x}^\xi) \vee \boldsymbol{\beta} = \mathbf{x}^\xi$ for all $\xi \in \mathcal{K}$, then W and $\boldsymbol{\theta}$ given by (3) also satisfies $(W \circ \mathbf{x}^\xi) \vee \boldsymbol{\theta} = \mathbf{x}^\xi$ for all $\xi \in \mathcal{K}$. Furthermore, the inequalities $A \leq W$ and $\boldsymbol{\beta} \leq \boldsymbol{\theta}$ hold true.

Observe, in particular, that the identity matrix $I \in [0,1]^{n\times n}$ and the vector of zeros $\mathbf{0} = [0, \ldots, 0]^T \in [0,1]^n$ are such that $(I \circ \mathbf{x}^\xi) \vee \mathbf{0} = \mathbf{x}^\xi$ for every original pattern $\mathbf{x}^\xi$. Thus, the synaptic weight matrix W and the threshold vector $\boldsymbol{\theta}$ given by (3) satisfy $I \leq W$ and $(W \circ \mathbf{x}^\xi) \vee \boldsymbol{\theta} = \mathbf{x}^\xi$ for all $\xi \in \mathcal{K}$. As a consequence, one can store as many patterns as desired in an AFIM model [13].

The following proposition establishes a relationship between the output of an AFIM and the fixed points of the synaptic weight matrix W [13, 8, 11]. Recall that $\mathbf{z} \in [0,1]^n$ is a fixed point of $W \in [0,1]^{n\times n}$ if and only if $W \circ \mathbf{z} = \mathbf{z}$. We denote the set of all fixed points of W by $F(W)$, i.e., $F(W) = \{\mathbf{z} \in [0,1]^n : W \circ \mathbf{z} = \mathbf{z}\}$.

Proposition 1. *Given an input pattern $\mathbf{x} \in [0,1]^n$, the output $\mathcal{W}(\mathbf{x}) = (W \circ \mathbf{x}) \vee \boldsymbol{\theta}$ of an AFIM is the smallest fixed point $\mathbf{z}$ of W such that $\mathbf{z} \geq \mathbf{x}$ and $\mathbf{z} \geq \boldsymbol{\theta}$, i.e.,*

$$\mathcal{W}(\mathbf{x}) = \bigwedge \{\mathbf{z} \in F(W) : \mathbf{z} \geq (\mathbf{x} \vee \boldsymbol{\theta})\}\,. \tag{4}$$

Proposition 1 shows that AFIMs exhibit one-step convergence. Moreover, it gives useful insights on the noise tolerance of AFIMs. For example, Proposition 1 says that an AFIM recalls an original pattern $\mathbf{x}^\xi$ only if the input $\mathbf{x}$ is smaller than $\mathbf{x}^\xi$. In other words, AFIMs are suited for the reconstruction of patterns corrupted by erosive noise, but are incapable of handling dilative noise. Recall that a distorted version $\tilde{\mathbf{x}}^\xi$ of the original pattern $\mathbf{x}^\xi$ has undergone an *erosive change* if $\tilde{\mathbf{x}}^\xi \leq \mathbf{x}^\xi$ and a *dilative change* if $\tilde{\mathbf{x}}^\xi \geq \mathbf{x}^\xi$ [19]. Further information on the noise tolerance of AFIMs can be obtained by investigating $F(W)$. A complete characterization of the set of fixed points of a general class of FAMs can be found in [22].

Let us conclude this section by recalling that the synaptic weight matrix $W = (w_{ij}) \in [0,1]^{n\times n}$ and the threshold vector $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_n]^T \in [0,1]^n$ given by IFL can be easily computed by means of the following equations for every $i, j = 1, \ldots, n$:

$$w_{ij} = \bigwedge_{\xi=1}^{p} I_T(x_j^\xi, x_i^\xi) \quad \text{and} \quad \theta_i = \bigwedge_{\xi=1}^{p} x_i^\xi\,. \tag{5}$$

Here, the operator $I_T : [0,1] \times [0,1] \to [0,1]$ denotes the R-implication associated with the t-norm that is used in the recall phase [23, 10]. Recall that the R-implication associated with a continuous t-norm T is defined as follows:

$$I_T(x,y) = \bigvee \{z \in [0,1] : T(x,z) \leq y\}\,. \tag{6}$$

The following equations present the R-implications associated with the *minimum*, *product*, and *Lukasiewicz t-norm*, respectively [23, 10, 13]:

$$I_M(x,y) = \begin{cases} 1, & x \leq y \\ y, & x > y \end{cases} \qquad \text{(Gödel)} \tag{7}$$

$$I_P(x,y) = \begin{cases} 1, & x \leq y \\ y/x, & x > y \end{cases} \qquad \text{(Goguen)} \tag{8}$$

$$I_L(x,y) = 1 \wedge (1 - x + y) \qquad \text{(Lukasiewicz)} \tag{9}$$

3 Autoassociative Fuzzy Implicative Memories for Large Gray-Scale Images

First of all, recall that a gray-scale image $\mathbf{a}$ of size $M \times N$ can be identified with a finite fuzzy set $\mathbf{x} = [x_1, \ldots, x_n]^T$ of length $n = NM$. The fuzzy set $\mathbf{x}$ is obtained by confining the gray-scale values to the unit interval and by arranging the pixels in a column vector using the standard column-scan method. As a consequence, AFIMs can be used for the storage and retrieval of gray-scale images.

Note, however, that an AFIM needs a synaptic weight matrix $W \in [0,1]^{n\times n}$ and a threshold vector $\boldsymbol{\theta} \in [0,1]^n$. In other words, the model requires computation and storage of $n(n+1)$ values. The following example reveals that the required computational resources restrain the applications of AFIMs to small gray-scale images.

Example 1. Consider a gray-scale image of size 512×512 with 256 shades of gray. This image corresponds to a finite fuzzy set that can be represented by a vector of length $n = 512^2 = 262144$. As a consequence, an AFIM requires the computation and storage of approximately 6.9×10^9 values corresponding to the synaptic weights and thresholds. If these values are represented in a computer using 64-bit double precision floating points, then the model allocates approximately 512 gigabytes of memory space. Similar considerations reveal that AFIMs allocate approximately 2 and 32 gigabytes of memory spaces for the storage and recall of images of size 128×128 and 256×256, respectively.

[1] In this paper, the symbols "$\bigvee$" and "$\bigwedge$" denote the supremum (or maximum) and infimum (or minimum) operations, respectively.

ISBN: 978-989-95079-6-8

Concluding, due to computer memory limitations, applications of AFIMs are usually bounded to images of size less than or approximately equal to 128×128. The following subsections provide strategies that allow for the storage and recall of large gray-scale images in AFIM based models.

3.1 Sparsely Connected AFIMs

Several studies suggest that neurons in the parts of the human brain that exhibit functional properties of associative memory are connected to few other neurons [24, 25, 26]. Motivated by this biological remark, we will remove a considerable amount of synaptic junctions of an AFIM model, i.e., we will introduce zeros in the synaptic weight matrix given by (5). However, in view of the following theorem[2], we will not delete synaptic weights in the main diagonal of W.

Theorem 1. *Consider an arbitrary fundamental memory set $\{\mathbf{x}^1, \ldots, \mathbf{x}^k\}$, where each $\mathbf{x}^\xi \in [0,1]^n$. If $V \in [0,1]^{n\times n}$ and $\boldsymbol{\vartheta} \in [0,1]^n$ satisfy the inequalities $I \leq V \leq W$ and $\boldsymbol{\vartheta} \leq \boldsymbol{\theta}$, where $W \in [0,1]^{n\times n}$ and $\boldsymbol{\theta} \in [0,1]^n$ are given by IFL, then the following equation holds true for every $\xi = 1, \ldots, k$:*

$$\mathbf{x}^\xi = (V \circ \mathbf{x}^\xi) \vee \boldsymbol{\vartheta}\,. \tag{10}$$

In view of Theorem 1, we define a *sparsely connected AFIM* (SCAFIM) as an autoassociative max-T FAM with threshold $\boldsymbol{\vartheta} \leq \boldsymbol{\theta}$ and a synaptic weight matrix $I \leq V \leq W$ that has very few nonzero elements. Since V has a sparse structure, SCAFIMs usually do not require large computational resources. Thus, they are suited for the storage and recall of large gray-scale images.

By definition, SCAFIMs exhibit optimal absolute storage capacity such as the fully connected AFIMs. However, in contrast to the latter, SCAFIMs do not necessarily exhibit one-step convergence. Therefore, in analogy to the famous *Hopfield network*, we may employ SCAFIMs with feed-back [2, 20]. In this case, given an input pattern $\mathbf{x} \in [0,1]^n$, we define the following sequence where $\mathbf{x}(0) = \mathbf{x}$:

$$\mathcal{V}_t(\mathbf{x}) = \mathbf{x}(t) = \big[V \circ \mathbf{x}(t-1)\big] \vee \boldsymbol{\vartheta}\,, \quad \forall\, t = 1, 2, \ldots \tag{11}$$

The resulting model is referred to as *dynamic or recursive SCAFIM*. Note that $\mathcal{V}_t(\mathbf{x})$ represents the output of a dynamic SCAFIM after t steps. In particular, $\mathcal{V}_1(\mathbf{x})$ corresponds to the output of a single-step model.

The following theorem shows that the sequence $\mathcal{V}_t(\mathbf{x})$ given by (11) converges for every input pattern $\mathbf{x} \in [0,1]^n$. Furthermore, Theorem 2 below relates the limit of the sequence $\mathcal{V}_t(\mathbf{x})$, denoted by $\mathcal{V}_*(\mathbf{x})$, with the fixed points of the synaptic weight matrix V. Recall that the set of fixed points of V is defined as $F(V) = \{\mathbf{z} \in [0,1]^n : V \circ \mathbf{z} = \mathbf{z}\}$.

Theorem 2. *Given a fundamental memory set $\{\mathbf{x}^1, \ldots, \mathbf{x}^k\}$, define $\boldsymbol{\vartheta}$ and V such that $\boldsymbol{\vartheta} \leq \boldsymbol{\theta}$ and $I \leq V \leq W$, where W and $\boldsymbol{\theta}$ are given by IFL. Then, for every input pattern $\mathbf{x} \in [0,1]^n$, the sequence given by* (11) *is monotonically increasing and converges to the smallest fixed point $\mathbf{z}$ of V such that $\mathbf{z} \geq \mathbf{x}$ and $\mathbf{z} \geq \boldsymbol{\vartheta}$, i.e.,*

$$\mathcal{V}_*(\mathbf{x}) = \lim_{t\to\infty} \mathbf{x}(t) = \bigwedge \{\mathbf{z} \in F(V) : \mathbf{z} \geq (\mathbf{x} \vee \boldsymbol{\vartheta})\}. \tag{12}$$

[2]We would like to point out that we intend to publish the proofs of Theorems 1, 2, and 3 in an upcoming journal paper.

Note that Theorem 2 is analogous to Proposition 1. Thus, such as AFIMs, SCAFIMs cannot recall an original pattern $\mathbf{x}^\xi$ if the input $\mathbf{x}$ is greater than $\mathbf{x}^\xi$. More importantly, both Proposition 1 and Theorem 2 relate noise tolerance with the set of fixed points of the synaptic weight matrix. The following theorem establishes a relationship between $F(V)$ and $F(W)$. As a consequence, we are able to relate the noise tolerance of a certain SCAFIM and its corresponding fully connected AFIM. Theorem 3 below also relates the set of fixed points of V with the set of fixed points of a matrix U that has a smaller number of non-zeros entries than V.

Theorem 3. *Let $F(U), F(V)$ and $F(W)$ denote the set of fixed points of matrices $U, V, W \in [0,1]^{n\times n}$ with respect to a certain max-T product. If $I \leq U \leq V \leq W$, then $F(W) \subseteq F(V) \subseteq F(U)$.*

Remark 1. Theorems 2 and 3 tell us that the noise tolerance of dynamic SCAFIMs with respect to erosive noise diminishes - because the number of fixed points increases - as we remove more and more synaptic junctions of a fully connected AFIM. In contrast, a dynamic SCAFIMs exhibit better noise tolerance with respect to dilative noise than its corresponding AFIM. Let us clarify this remark by means of an illustrative example.

Example 2. Let W and $\boldsymbol{\theta}$ denote the synaptic weight matrix and threshold vector of an AFIM. Suppose that U and V are sparsely connected synaptic weight matrices that satisfy the inequalities $I \leq U \leq V \leq W$, i.e., U has fewer non-zeros elements than V. Figure 1 depicts the sets of fixed points $F(W)$, $F(V)$, and $F(U)$. Here, we represented the set $[0,1]^n$ by a line from $\mathbf{0} = [0, 0, \ldots, 0]^T$ to $\mathbf{1} = [1, 1, \ldots, 1]^T$ although this set is not *totally ordered*[3]. The fixed points of W, V, and U are represented by short vertical lines. Note that, in agreement with Theorem 3, we have depicted $F(W) \subseteq F(V) \subseteq F(U)$.

Suppose that $\mathbf{x}^\xi$ represents an original pattern and $\check{\mathbf{x}}$ corresponds to an eroded version of $\mathbf{x}^\xi$. Let $\mathcal{W}$ denote the AFIM model whose synaptic weight matrix is W and threshold vector $\boldsymbol{\theta}$ is shown in Figure 1. By Proposition 1, $\mathcal{W}(\check{\mathbf{x}}) = \mathbf{x}^\xi$ because $\mathbf{x}^\xi$ is the smallest fixed point of W greather than $\check{\mathbf{x}}$ and $\boldsymbol{\theta}$. Similarly, let $\mathcal{V}_*$ and $\mathcal{U}_*$ denote dynamic SCAFIMs with threshold vector $\boldsymbol{\vartheta} = \boldsymbol{\theta}$ and synaptic weight matrices V and U, respectively. In contrast to the AFIM model, the dynamic SCAFIMs $\mathcal{V}_*$ and $\mathcal{U}_*$ fail to recall the original pattern $\mathbf{x}^\xi$ because there are fixed points in $F(V)$ and $F(U)$ between $\check{\mathbf{x}} \vee \boldsymbol{\vartheta}$ and $\mathbf{x}^\xi$. Furthermore, observe that $\mathcal{V}_*$ outperforms the SCAFIM $\mathcal{U}_*$ because U has a fixed point between $\check{\mathbf{x}} \vee \boldsymbol{\vartheta}$ and $\mathcal{V}_*(\check{\mathbf{x}})$.

Conversely, let us suppose that $\hat{\mathbf{x}}$ represents a corrupted version of $\mathbf{x}^\xi$ contaminated by a small dilative noise. In view of the relationship $F(W) \subseteq F(V) \subseteq F(U)$, the dynamic SCAFIM $\mathcal{U}_*$ and the AFIM $\mathcal{W}$ yield the best and the worst results, respectively.

3.2 Strategies for the Construction of SCAFIMs

This subsection briefly answers the question as to derive an SCAFIM model from an AFIM. First, note that the threshold

[3]Recall that $(X, \leq)$ is *totally ordered*, also called *linearly ordered*, if for any $x, y \in X$, either $x \leq y$ or $y \leq x$. Note, however, that there are patterns $\mathbf{x}, \mathbf{y} \in [0,1]^n$ such that neither inequalities $\mathbf{x} \leq \mathbf{y}$ nor $\mathbf{y} \leq \mathbf{x}$ hold true. Thus, Figure 1 is merely illustrative.

ISBN: 978-989-95079-6-8

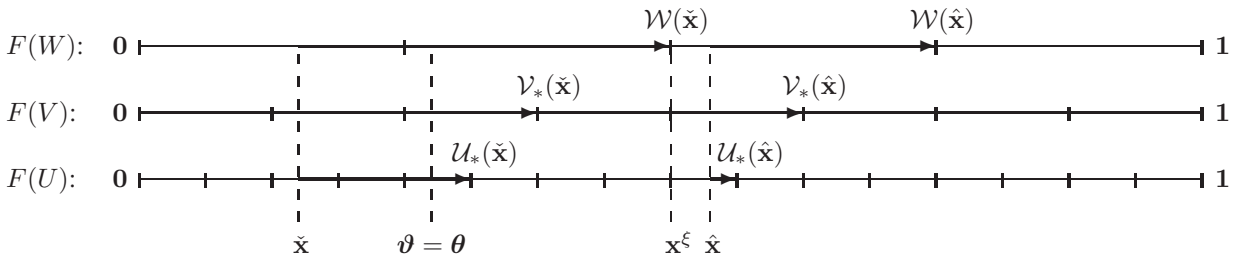

Figure 1: Illustrative example of the noise tolerance of an AFIM $\mathcal{W}$ and two dynamic SCAFIMs $\mathcal{V}_*$ and $\mathcal{U}_*$. Here, the synaptic weight matrices satisfy the inequalities $I \leq U \leq V \leq W$.

vector of an SCAFIM can be defined either as $\boldsymbol{\vartheta} = \boldsymbol{\theta}$ or $\boldsymbol{\vartheta} = \mathbf{0}$, depending on computational resources available. Let us consider $\boldsymbol{\vartheta} = \boldsymbol{\theta}$ in this paper. As far as we are concerned, the synaptic junctions of an AFIM can be removed in a strength-based manner or in a structured-based manner.

In the strength-based strategy, we delete synaptic weights by taking into account their strength. Specifically, we define the sparsely connected synaptic weight matrix $V \in [0,1]^{n \times n}$ as follows for some $\alpha \in (0,1]$:

$$v_{ij} = \begin{cases} w_{ij} & \text{if } w_{ij} \geq \alpha \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

Note that the synaptic junctions with strength less than α are removed from the fully connected AFIM.

Alternatively, we say that we removed synaptic junctions in a structurated-based manner if we fix the non-zeros entries of the synaptic weight matrix before its computation. For example, we can set that the synaptic weight matrix V has a band or a block diagonal structure. We would like to point out that this strategy generalizes the widely used procedure of partitioning large gray-scale images in order to obtain several small-sized AM models [15, 16, 17].

4 Computational Experiments

This section provides some computational experiments concerning the reconstruction of noisy or incomplete gray-scale images. In order to compare the performance of SCAFIMs with the fully connected AFIM model, and due to page constrain, we only consider gray-scale images of small size. Specifically, let us consider the images of size 64×64 with 256 shades of gray displayed in the first row of Figure 2. For each of these image, we generated a vector $\mathbf{x}^\xi$ of length $n = 64 \times 64 = 4096$ using the standard column-scan method. Furthermore, we divided all entries of $\mathbf{x}^\xi$ by 255 in order to obtain patterns in the hypercube $[0,1]^{4096}$.

We stored the eight original patterns $\mathbf{x}^\xi$ in the Lukasiewicz AFIM, i.e., the AFIM based on the Lukasiewicz t-norm. We would like to recall that the Lukasiewicz AFIM outperformed several AM models, including the FAMs of Kosko and the GFAM based on the Lukasiewicz t-norm, in an experiment using faces images from the database of AT&T Laboratories, Cambridge [8, 13].

Next, we constructed six SCAFIM models, denoted by $\mathcal{V}^a, \mathcal{V}^b, \ldots, \mathcal{V}^f$, all of them with threshold vector $\boldsymbol{\vartheta} = \boldsymbol{\theta}$. The corresponding sparsely connected synaptic weight matrices $V^a, \ldots, V^f$ were derived as follows from W given by IFL: V^a has 16 blocks of size 256×256, V^b has 64 blocks of size 64×64, and V^c is a band matrix with bandwidth equals to 19. The synaptic weight matrices V^d, V^e, and V^f were obtained by means of (13) with $\alpha = 0.893, 0.961$, and 1, respectively. We would like to point out that we have chosen α and the bandwith of V^c in order to obtain pairs V^a and V^d, V^b and V^e, V^c and V^f with approximately the same number of non-zeros entries. Furthermore, the output of the SCAFIM $\mathcal{V}^b$ corresponds to the gray-scale image that is obtained by adopting the following procedure: Store each column of an original image in a separate Lukasiewicz AFIM. This results in 64 AFIMs, each of which has a 64×64 synaptic weight matrix that corresponds to a block of V^b. The recalled image is obtained by feeding each AFIM with its corresponding column of a given image and by concatenating the outputs.

The first row of Table 1 provides the density (i.e., percentage of non-zeros entries) of the synaptic weight matrices W, $V^a, \ldots, V^f$. Table 1 also provides the memory space required by the seven AM models if double precision numbers are used to represent the synaptic weights and threshold values. We would like to point out that we used the standart row-column-value representation to store the sparsely connected synaptic weight matrices in the computer.

Afterward, we confirmed that the eight original patterns represent fixed points of $W, V^a, \ldots, V^f$, i.e., the Lukasiewicz AFIM as well as the six SCAFIMs exhibit optimal absolute storage capacity.

In order to compare the noise tolerance of the seven AM models, we first introduced as input an image corrupted by *pepper noise* with probability 0.3. The second row of Figure 2 shows the corrupted image and the corresponding recalled patterns. Table 1 contains the *base-10 logarithm* applied to the *arithmetic mean* of the *relative normalized mean squared error* (RNMSE) in 100 experiments. Recall that the RNMSE is given by the following equation where x_i, $\tilde{x}_i$, and y_i denote the i-th entry of the original, noisy, and recalled patterns, respectively:

$$RNMSE = \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(x_i - \tilde{x}_i)^2}\,. \tag{14}$$

We would like to point out that we have corrupted an original pattern that was randomly selected in each trial. Moreover, the six dynamic SCAFIMs were iterated until convergence. The arithmetic means of the number of iterations required for convergence are shown between parentheses in Table 1.

We also conducted similar experiments using *salt noise* with probability 0.01, *salt and pepper noise* with probability 0.05, *additive Gaussian noise* with mean 0 and variance 0.01,

ISBN: 978-989-95079-6-8

	$\mathcal{W}$	$\mathcal{V}^a$	$\mathcal{V}^b$	$\mathcal{V}^c$	$\mathcal{V}^d$	$\mathcal{V}^e$	$\mathcal{V}^f$
Density of the AM model	100%	6.25%	1.56%	0.46%	6.3%	1.42%	0.48%
Required memory space	134.2 MB	16.8 MB	4.2 MB	1.3 MB	17.0 MB	4.2 MB	1.3 MB
Tolerance w.r.t. pepper noise	-3.05 (1)	-2.47 (1)	-1.93 (1)	-1.78 (2.3)	-2.37 (2.6)	-1.27 (2.7)	-0.62 (1)
Tolerance w.r.t. salt noise	1.85 (1)	1.55 (1)	1.13 (1)	0.97 (5.3)	1.80 (8.9)	1.51 (9.7)	0.96 (1)
Tol. w.r.t. salt & pepper noise	13.84 (1)	9.53 (1)	5.00 (1)	3.76 (3.7)	13.08 (7.6)	8.01 (6.7)	3.17 (1)
Tol. w.r.t. add. gaussian noise	6.40 (1)	3.50 (1)	2.18 (1)	1.64 (4.1)	4.97 (7.5)	2.78 (7.3)	1.35 (1)
Tolerance w.r.t. speckle noise	7.64 (1)	4.27 (1)	2.67 (1)	1.80 (4.8)	5.47 (8.7)	2.17 (8.1)	1.09 (1)

Table 1: Comparison of the noise tolerance, density of synaptic junctions, and required computational resources of Lukasiewicz AFIM and six dynamic SCAFIMs.

and *multiplicative gaussian noise (speckle)* with mean 1 and variance 0.04. Table 1 provides the base-10 logarithm of the mean of the RNMSE in 100 experiments. Again, the mean of the number of iterations required for convergence is shown between parentheses. Figure 2 shows instances of the noisy images and corresponding recalled patterns.

Note that, in accordance with Remark 1, the noise tolerance with respect to erosive (pepper) noise diminishes - whereas the tolerance with respect to dilative (salt) and mixed (gaussian and speckle) noise increase - as we remove more and more synaptic junctions from a fully connected AFIM.

The following compares the two strategies for the construction of dynamic SCAFIMs. Let us begin by considering $\mathcal{V}^a$ and $\mathcal{V}^d$. Both SCAFIMs have approximately the same number of synaptic junctions. Thus, they required almost the same amount of memory space. However, the SCAFIM $\mathcal{V}^a$ outperformed $\mathcal{V}^d$ with respect to erosive, dilative, and mixed noise. Furthermore, $\mathcal{V}^a$ is preferable than $\mathcal{V}^d$ in computational terms. First, because V^a converged to the fixed point with only one iteration whereas V^b required in mean 7 iterations. Secondly, in the structurated-based strategy, we only compute the synaptic junctions that are really stored in the computer. In contrast, in the strength-based strategy, we compute all synaptic junctions but we keep only those with strength greater than or equal to α. In addition, we can determine exactly the amount of bytes that the SCAFIM $\mathcal{V}^a$ will require before the computation of V^a, but we can only estimate the memory space that V^d will consume. We reach a similar conclusion by comparing the SCAFIMs $\mathcal{V}^b$ and $\mathcal{V}^e$.

Finally, let us compare $\mathcal{V}^c$ and $\mathcal{V}^f$. Note that both models consumed almost the same amount of computational resources. Moreover, they exhibited similar tolerance with respect to dilative or mixed noise. However, the SCAFIM $\mathcal{V}^c$ outperformed $\mathcal{V}^f$ with respect to erosive noise, but the former usually required more than one iteration to converge[4].

5 Concluding Remarks

This paper introduces the class of SCAFIMs. In few words, a SCAFIM is obtained by removing a considerable amount of synaptic junctions of an AFIM model. The novel models inherit the optimal storage capacity of AFIMs but do not necessarily exhibit one step convergence. Nevertheless, we showed that dynamic SCAFIMs, i.e., SCAFIMs with feedback, yield a monotonically increasing sequence of patterns that converges to a fixed point of the underlying (sparse) synaptic weight matrix. Thus, such as the AFIMs, SCAFIMs exhibit tolerance with respect to incomplete or eroded patterns. Furthermore, we pointed out that the noise tolerance of SCAFIMs decreases as the density of synaptic junctions. In contrast, the tolerance with respect to dilative and mixed noise increases as the density of synaptic junctions decreases.

In addition, we briefly investigated two strategies for the construction of SCAFIMs: the structurated-based and the strength-based strategies. In computational terms, the former is usually preferable than the latter due to the following reasons. First, because we compute only the synaptic junctions that will be really stored in the computer. Secondly, because we can determine exactly the amount of bytes that the sparse matrix will require. Finally, computational experiments revealed that the structurated-based strategy usually yields better noise tolerance than the strength-based strategy. However, further research is needed on how to choose the best strategy for the construction of SCAFIMs.

[4] We would like to point out that the SCAFIM $\mathcal{V}^f$ corresponds to the *Gaines fuzzy morphological associative memory* (FMAM) that was introduced and investigated in [22]. This model always converges to a fixed point with only one step.

References

[1] J. Anderson. *An Introduction to Neural Networks*. MIT Press, MA, 1995.

[2] M. Hassoun, editor. *Associative Neural Memories: Theory and Implementation*. Oxford University Press, Oxford, U.K., 1993.

[3] T. Kohonen. *Self-Organization and Associative Memory*. Springer Verlag, 1984.

[4] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, Upper Saddle River, NJ, 1999.

[5] B. Kosko. *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Prentice Hall, Englewood Cliffs, NJ, 1992.

[6] B. Zhang, H. Zhang, and S. Ge. Face recognition by applying wavelet subband representation and kernel associative memory. *IEEE Transactions on Neural Networks*, 15(1):166–177, January 2004.

[7] D. Zhang and W. Zuo. Computational intelligence-based biometric technologies. *IEEE Computational Intelligence Magazine*, 2(2):26–36, May 2007.

[8] P. Sussner and M. Valle. Fuzzy associative memories and their relationship to mathematical morphology. In W. Pedrycz, A. Skowron, and V. Kreinovich, editors, *Handbook of Granular Computing*, chapter 33, pages 733–754. John Wiley and Sons, Inc., New York, 2008.

[9] S. Teddy, C. Quek, and E. Lai. PSECMAC: A novel self-organizing multiresolution associative memory architecture.

ISBN: 978-989-95079-6-8

Figure 2: First row displays the original gray-scale images of size 64×64 with 256 shades of gray. The first column of the following three rows shows patterns corrupted by *pepper noise* with probability 0.3, *salt noise* with probability 0.01, and *salt and pepper noise* with probability 0.05, respectively. The other seven columns (from left to right) exhibit the corresponding patterns that were recalled by the Lukasiewicz AFIM $\mathcal{W}$ and the SCAFIMs $\mathcal{V}^a$, $\mathcal{V}^b$, $\mathcal{V}^c$, $\mathcal{V}^d$, $\mathcal{V}^e$, and $\mathcal{V}^f$, respectively.

IEEE Transactions on Neural Networks, 19(4):689–712, April 2008.

[10] W. Pedrycz and F. Gomide. *Fuzzy Systems Engineering: Toward Human-Centric Computing*. Wiley-IEEE Press, New York, 2007.

[11] M. Valle and P. Sussner. A general framework for fuzzy morphological associative memories. *Fuzzy Sets and Systems*, 159(7):747–768, 2008.

[12] F. Chung and T. Lee. On fuzzy associative memory with multiple-rule storage capacity. *IEEE Transactions on Fuzzy Systems*, 4(3):375–384, 1996.

[13] P. Sussner and M. Valle. Implicative fuzzy associative memories. *IEEE Transactions on Fuzzy Systems*, 14(6):793–807, 2006.

[14] P. Sussner and M. Valle. Grayscale morphological associative memories. *IEEE Transactions on Neural Networks*, 17(3):559–570, May 2006.

[15] D. Lee. Improvements of complex-valued hopfield associative memory by using generalized projection rules. *IEEE Transactions on Neural Networks*, 17(5):1341–1347, September 2006.

[16] M. Müezzinoğlu, C. Güzeliş, and J. Zurada. A new design method for the complex-valued multistate Hopfield associative memory. *IEEE Transactions on Neural Networks*, 14(4):891–899, July 2003.

[17] G. Costantini, D. Casali, and R. Perfetti. Neural associative memory storing gray-coded gray-scale images. *IEEE Transactions on Neural Networks*, 14(3):703–707, May 2003.

[18] G. Costantini, D. Casali, and R. Perfetti. Associative memory design for 256 gray-level images using a multilayer neural network. *IEEE Transactions on Neural Networks*, 17(2):519–522, march 2006.

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

[20] J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79:2554–2558, April 1982.

[21] M. Valle, P., and F. Gomide. Introduction to implicative fuzzy associative memories. In *Proceedings of the IEEE International Joint Conference on Neural Networks*, pages 925 – 931, Budapest, Hungary, 2004.

[22] M. Valle and P. Sussner. Storage and recall capabilities of fuzzy morphological associative memories with adjunction-based learning. 2007. Submitted for publication.

[23] H. Nguyen and E. Walker. *A First Course in Fuzzy Logic*. Chapman & Hall/CRC, Boca Raton, 2 edition, 2000.

[24] H. Bosch and F. Kurfess. Information storage capacity of incompletely connected associative memories. *Neural Networks*, 11(5):869–876, 1998.

[25] B. Olshausen and D. Field. Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14(4):481–487, August 2004.

[26] R. Quiroga, L. Reddy, G. Kreiman, C. Koch, and I. Fried. Invariant visual representation by single neurons in the human brain. *Nature*, 435(7045):1102–1107, June 2005.

ISBN: 978-989-95079-6-8